THE AUTONOMY AND FUNCTION OF EUKARYOTIC SIGNAL SEQUENCES

Segregation of Variants of Preprochymosin and Prelysozyme in *Xenopus* Oocytes and *In Vitro*

Fiona Rachel Strachan
(BSc. Hons. University College of Wales, Aberystwyth)

Submitted for degree of PhD.

Department of Biological Sciences,
University of Warwick

August 1986

Dedication

This thesis is dedicated with love and great respect to two people:
Mrs. Marie Witty, aged 88, who showed me the distinction between wisdom and knowledge,
and
Irena Horak, whose tragic and untimely death in 1984 at the age of 25, caused me to reappraise my outlook on life.

-CONTENTS-

List of Figures & Tables

Acknowledgements

The financial support of the Medical Research Council is acknowledged in providing my research studentship grant.

There are numerous people without whose help and encouragement the work described in this thesis and its presentation would not have been possible. Particular acknowledgement is made of a few of these people but my thanks extend to them all.

My supervisor Alan Colman, for all his help from the inception of the project to the submission of this thesis. John Armstrong for his careful and constructive criticism of the draft version of my thesis. Other past and present members of the Developmental Biology group for their discussion and practical assistance, scientific and otherwise; especially Paul Krieg, Linda Tabe, Douglas Drummond, Dave Jackson, Liz Wallis and Giorgio Valle; also Christine Newton, whose efficient running of the laboratories was much apppreciated. Thanks to Gareth Cross, Linda Tabe, Ann Horrell and Jane Lewis for reminding me of 'life beyond research and thesis-writing'!

My parents Anne and John, and my sister Diana for the constancy of their support, encouragement and interest throughout, but especially during the writing-up period. Tony Page for his generous help and advice over the years. My brother David who proposed and implemented a joint investment in a word processor for producing the thesis; and discerned and resolved many, but unfortunately not all, of the idiosyncrasies of 'The Machine'. The generous and extended loan of drawing instruments by Charlotte Morrison is also gratefully acknowledged.

Declaration

I declare that this thesis was composed by myself and has not previously been submitted for any degree. Where results have been obtained in collaboration with others this is acknowledged in the text, otherwise the work reported was carried out by myself.

Summary

Cloned complementary DNAs encoding the secretory proteins chick prelysozyme and calf preprochymosin were inserted downstream from various viral promoters in modified recombinant "shuttle" vectors. The microinjection of these constructs into the nuclei of *Xenopus laevis* oocytes resulted in the efficient expression of lysozyme and prochymosin proteins which were segregated into membranes and secreted by the oocyte. The signal sequences of the proteins were correctly processed as judged from molecular weight estimations. Injection of DNA encoding prochymosin without its signal sequence resulted in the synthesis of a prochymosin protein which was localised in the oocyte cytoplasm; whereas when DNA encoding mature chymosin was injected no proteins were detected by immunoprecipitation with prochymosin antisera. The same cDNAs were subsequently cloned into SP6 vectors and synthetic, capped RNA was prepared. Following cytoplasmic injection of SP6 RNA into oocytes the same compartmentation of the proteins was observed but again no chymosin protein was detected following injection of RNA encoding mature chymosin, although translation of this RNA *in vitro* produced a protein with the expected molecular weight of chymosin which was precipitated by prochymosin antisera.

The expression of preprochymosin messenger RNA, following cytoplasmic injection into oocytes and also using *in vitro* translation systems, showed the mRNA encoded two preprochymosin proteins specifically precipitated by prochymosin antisera. In the oocyte both forms were processed, segregated and secreted; whilst *in vitro* the precursors were cleaved on translocation within dog pancreatic microsomes where they became resistant to digestion by exogenous proteases. The translation of preprochymosin mRNA *in vitro* has previously been reported to produce only one major polypeptide on gel electrophoresis of products precipitated by antiprochymosin sera. The origin and nature of the two electrophoretically distinct species is not certain; but it was noted that the protein product of the cloned preprochymosin cDNA showed the same mobility on SDS-polyacrylamide gels as the faster migrating species encoded by the mRNA.

Two hybrid genes were constructed encoding proteins in which the signal sequence of prelysozyme was replaced with different N-terminal regions of preprochymosin. C_6L contained a fragment from the preprochymosin cDNA which encoded the signal sequence and the first six amino acids of prochymosin; this was fused to codons 8 to 129 of mature lysozyme. The second construct $C_{62}L$ carried a larger portion of preprochymosin, up to codon 62 of prochymosin, with the same C-terminal region of lysozyme. These fusions showed poor and variable expression following nuclear injection of the hybrid genes contained in the shuttle vector. However cytoplasmic injection of the corresponding SP6 RNAs demonstrated that both fusion proteins were synthesized in the oocyte and segregated into membranes, but did not get secreted from the oocyte. The distribution of $C_{62}L$ protein within the oocyte corresponded with that observed for the majority of other secretory proteins including preprochymosin, with most protein fractionating with vesicles. In contrast C_6L displayed the anomalous fractionation previously observed for lysozyme, with approximately equal amounts of the processed protein fractionating with the cytoplasm and the membranes. Relative to the precursor polypeptides produced by *in vitro* translation of the SP6 RNAs for C_6L and $C_{62}L$, the respective proteins expressed in oocytes each showed an increased mobility on electrophoresis consistent with the cleavage of the signal peptide. The same processing of the preC_6L and pre$C_{62}L$ proteins was observed when *in vitro* translation was carried out in the presence of pancreatic microsomes. The observed compartmentation and processing of these hybrid proteins indicates that a eukaryotic signal sequence functions autonomously in initiating the translocation of secretory proteins, but that other properties of the protein conformation are necessary to achieve subsequent secretion.

Abbreviations

Listed below are the abbreviations used in this thesis which are not classified as 'accepted' in the 'Instructions to authors' of The Biochemical Journal (volume 225, 1985).

BCIG (X-gal)	5-bromo-4-chloro-3-indolyl-β-galactoside
bisacrylamide	N,N'-methylenebisacrylamide
bp	base pair
BSA	bovine serum albumin
CIP	calf intestinal phosphatase
DHFR	dihydrofolate reductase
DTT	dithiothreitol
ER	endoplasmic reticulum
IPTG	isopropyl-β-D-thiogalactopyranoside
LTR	long terminal repeat
MBP	maltose binding protein
OTC	ornithine transcarbamylase
PAGE	polyacrylamide gel electrophoresis
PMSF	phenylmethylsulphonylflouride
RER	rough endoplasmic reticulum
SDS	sodium dodecyl sulphate (sodium lauryl sulphate)
SRP	signal recognition particle
ssRuBPC	ribulose bisphosphate carboxylase small subunit
SV40	Simian Virus 40
TCA	trichloroacetic acid
TEMED	N,N,N',N'-tetramethylethylenediamine
TK	thymidine kinase
VSV-G	Vesicular Stomatitis Virus Glycoprotein

Buffers - full composition given in section II.A

TE	Tris/EDTA
TEA	Tris/EDTA/Acetate
TBE	Tris/Borate/EDTA

-INTRODUCTION-

I.A. INTRACELLULAR SORTING OF PROTEINS

I.A.1. Preamble

It is evident when one considers the diverse location of proteins in a eukaryotic cell that there is a necessity for mechanisms which direct proteins to their correct subcellular compartment. Means are required to distinguish proteins destined for export or for specific organelles (*ie.* mitochondria, endoplasmic reticulum, Golgi body, lysosomes) from those which remain in the cytoplasm. In this process of protein segregation each transported protein must cross (translocate) at least one membrane. The sorting of proteins to particular compartments results in organelles with a characteristic protein complement, both in terms of content and membrane composition.

In this section I will be discussing current views of the mechanisms involved in intracellular protein sorting, with reference primarily to eukaryotes but also to prokaryotes since both these systems share common features. More recent reviews on this subject are given by Davis & Tai(1980), Kreil(1981), Lodish *et al*(1981), Sabatini *et al*(1982), Strauss & Boime(1982), Silhavy *et al*(1983) and Wickner & Lodish(1985). As this thesis concerns a particular class of proteins, those secreted by the cell, the emphasis will be on the processes which result in the segregation of these proteins and also their transport to the cell surface for export.

I.A.2. Translocation into the Endoplasmic Reticulum

It was the classical work of Palade and his colleagues (summarised in Palade,1975) which established the intracellular route taken by secretory proteins. They found that secretory proteins were synthesized on membrane bound ribosomes and segregated immediately into the endoplasmic reticulum (ER) without appearing in the cytoplasm; from the ER the proteins were transported to the Golgi complex and then to secretory vesicles prior to discharge from the cell. Thus it appeared that proteins destined for export were distinguished at synthesis and were sequestered into the ER as the initial step along the secretory pathway.

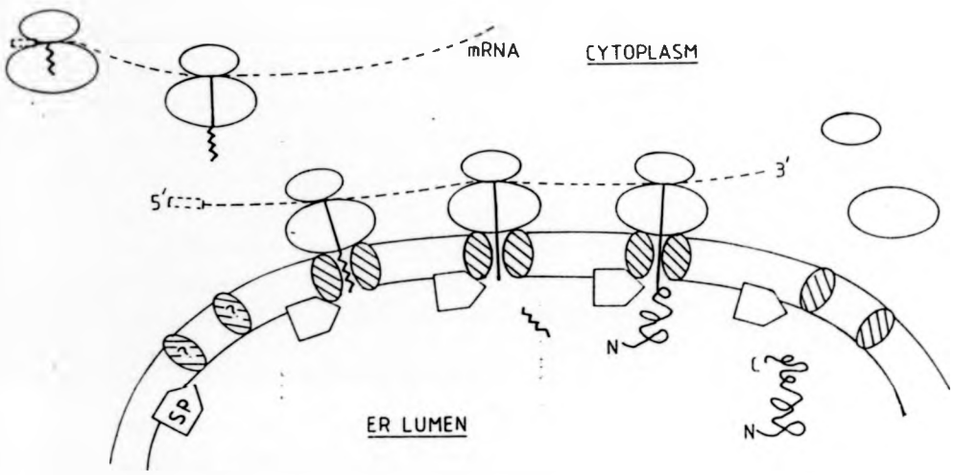


Figure I.1 The original 'Signal Hypothesis'
Representation of the translocation across the ER membrane of nascent secretory proteins, according to the original postulates of the Signal Hypothesis (Blobel & Sabatini,1971; Blobel & Dobberstein,1975a).
Key:

- mRNA
- RNA encoding the signal sequence
- nascent polypeptide
- signal peptide
- signal peptidase
- ribosome
- unknown membrane proteins

The signal hypothesis

In 1966 Redman & Sabatini demonstrated that proteins being synthesized on membrane-bound ribosomes were vectorially discharged into the ER. Studies by Milstein *et al*(1972) on the translation of immunoglobulin light chain messenger RNA (mRNA) in cell free systems found that the primary translation product of this secretory protein was a larger precursor. Subsequent *in vitro* translation experiments established that this precursor could be taken up and processed to the authentic mature product by dog pancreatic microsomes added at the start of translation (Blobel & Dobberstein,1975a and b). However addition of microsomes after completion of the polypeptide chain resulted in neither translocation or processing, which indicated that translation and translocation were obligatorily coupled. Such early observations led to the proposal of the 'Signal Hypothesis' for the mechanism by which secretory proteins are distinguished and sequestered for export (Blobel & Sabatini,1971; Blobel & Dobberstein,1975a and b). In its original form a number of postulates were made, these are represented in Figure 1:

1) mRNAs of proteins to be translocated across a membrane encode after the initiation codon a specific sequence of amino acids, termed the signal sequence.

2) Translation of mRNAs containing a signal sequence initiates on free ribosomes in the cytoplasm, but when the signal sequence emerges from the large subunit of the ribosome it causes attachment of the ribosome to the membrane.

3) Translation then continues on the membrane bound ribosome and is coupled to translocation as specific interactions with membrane proteins form a transient proteinaceous pore through which the nascent polypeptide is vectorially discharged across the membrane, in a thread-like manner.

4) Processing of the signal sequence occurs on the extracytoplasmic face of the membrane and the complete, mature polypeptide is released into the membrane compartment.

In their original paper Blobel & Dobberstein (1975a) suggest that the signal hypothesis of co-translational translocation could apply not only to secretory proteins but also to other proteins which transfer across the ER, like those destined for lysosomes, and to those proteins which cross other membranes such as cytoplasmically synthesized mitochondrial proteins. Indeed the Signal Hypothesis has remained a working model for protein translocation but has been extended and revised in the light of subsequent findings.

Structure and position of signal sequences

Work by Schechter *et al*(1975) with immunoglobulin light chains provided the first demonstration that a secretory protein precursor did contain an amino-terminal peptide extension. Examples were also soon found of transmembrane (Lingappa *et al*,1978a) and lysosomal proteins (Erickson *et al*,1981) synthesized with transient N-terminal signal peptides. Since the nascent transmembrane protein competed for uptake with a nascent secretory protein this suggested the two shared a common translocation pathway. It is now thought that all polypeptides which translocate the ER use the same transport machinery.

It was initially expected that signal sequences would show considerable structural homology since they were proposed to interact specifically with other proteins. However no uniformity of amino acid sequence or length of signal peptides has been observed; yet there is a common theme of a total of 15 to 30 amino acids with one or more basic, charged amino acid at the N-terminus, followed by a longer domain of hydrophobic amino acid residues, then towards the COOH end the hydrophilicity increases, with small side chain amino acids occurring around the cleavage site. This implies that constraints are imposed on the secondary structure of the signal peptide; and these observed structural features of signal sequences have been used as the basis of alternative hypotheses of protein segregation, which are discussed later. A more detailed consideration of the structural requirements of a signal sequence for a co-translational translocation is given in section B.3.

To date all eukaryotic secretory proteins studied have been found to have transient N-terminal signal sequences, with the notable exception of ovalbumin (Palmiter *et al*,1978). However it has been found that a number of integral membrane proteins do not possess a signal sequence which is cleaved on translocation; for example the erythrocyte anion transport protein - Band III (Braell & Lodish,1982a), cytochrome P-450 (Bar-Nun *et al*,1980; Sakaguchi *et al*,1984), and the viral proteins influenza neuraminidase (Markoff *et al*,1984) and the corona virus protein E1 (Rottier *et al*,1984). For some of these uncleaved translocated proteins it has been demonstrated that the NH_2-terminal region does act as the signal for translocation; for example a chimaeric influenza cDNA was constructed which encoded the first 40 amino acids of neuraminidase fused to haemagglutinin minus its normal cleaved signal peptide, when this was expressed in cells the chimaeric haemagglutinin was translocated into the rough ER and glycosylated (Bos *et al*,1984). The case of Band III showed a number of deviations from the classical Signal Hypothesis; in this

protein the N-terminal half of the protein faces the cytoplasm while the COOH half spans the membrane several times. Synchronised *in vitro* studies showed the protein was cotranslationally inserted into dog pancreatic microsomes without cleavage, but that insertion still occurred if microsomes were added as late as the time when about 50% (450 residues) of the polypeptide had been synthesized (Braell & Lodish,1982a). This suggested that in Band III the signal for translocation into the ER is internal, near the middle of the protein, and not at the NH_2 terminus. Thus signal sequences of secretory and integral membrane proteins can be N-terminal or internal, transient or permanent. It has been found that internal, uncleaved signal sequences of membrane proteins form regions which span the membrane (Markoff *et al*,1984). These findings led to a major modification of the Signal Hypothesis in which a loop insertion of the N-terminal or internal signal sequence into the ER membrane was envisaged, and the idea of multiple signal sequences was invoked to account for proteins which cross the membrane more than once (Blobel,1980). This sequence of events is depicted in Figure I.2.

Components of the Translocation Machinery

A key proposal of the signal hypothesis was the involvement of specific membrane proteins in the translocation process; and indeed it is this which distinguishes the hypothesis from others that have been proposed (discussed later). The development of *in vitro* systems enabled a dissection and biochemical analysis of the process of translocation. Early work by Kreibich *et al*(1978a and b) identified two integral membrane proteins, found only in microsomes derived from rough ER, which appeared to be components of the ribosome binding site and could be cross-linked to the ribosome large subunit; these were designated ribophorin I and II. Evidence of a saturable membrane receptor for translocated proteins then came from the work of Majzoub *et al*(1980). In the late 1970s and early 1980s several lines of investigation found that translocation could be inhibited *in vitro* by the treatment of microsomes with agents which affect proteins, such as proteases, alkylating agents and high ionic strength buffers (Warren & Dobberstein,1978; Walter *et al*,1979; Prehn *et al*,1980; Jackson *et al*,1980; Meyer & Dobberstein,1980a and b; Prehn *et al*,1981). These led to the identification of two components from canine pancreatic microsomes which are required for translocation of nascent polypeptides across membranes.

Dobberstein(1978) and later Walter & Blobel(1980) purified a protein component from the high salt extract of microsomal membranes which

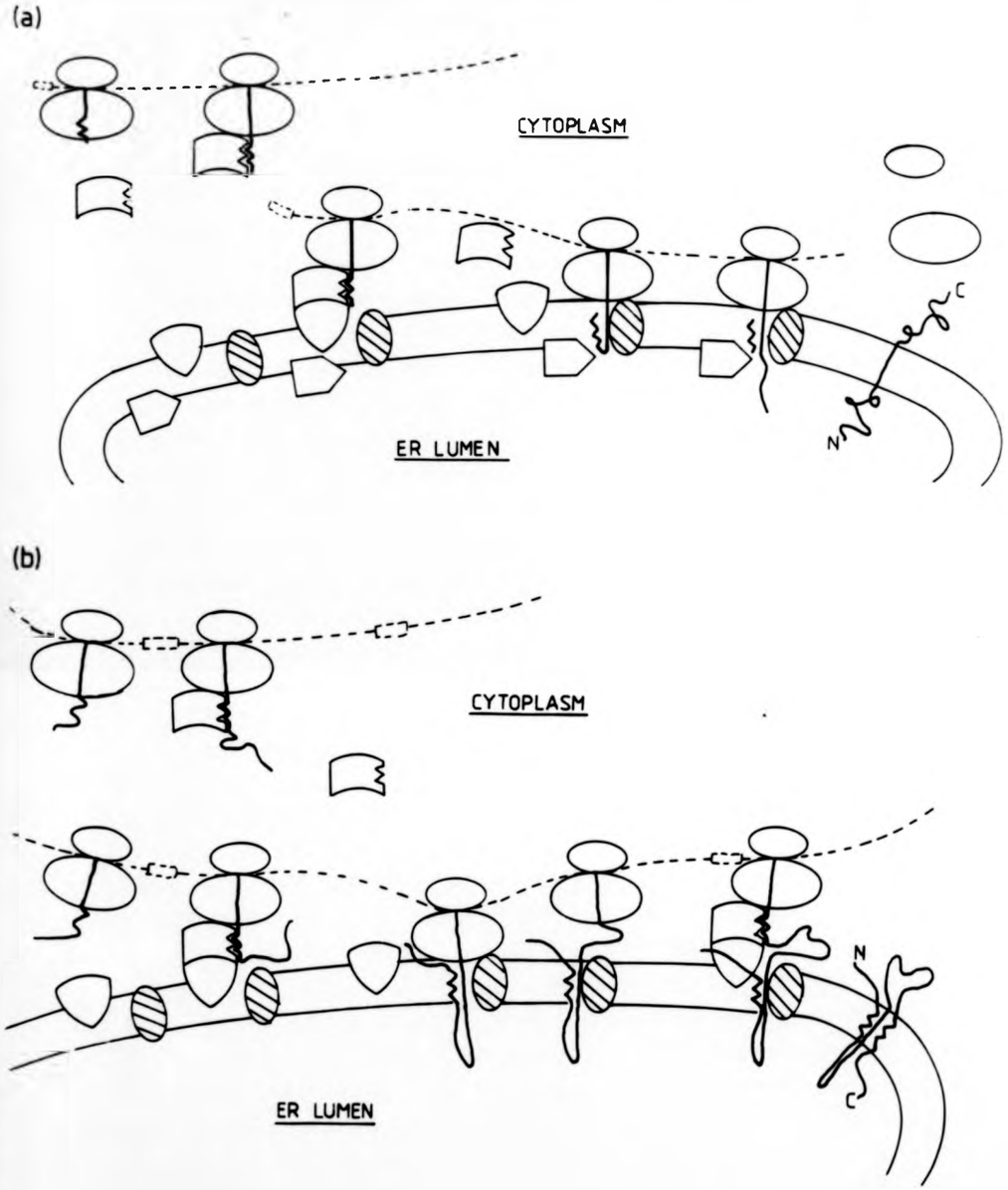


Figure I.2 The updated 'Signal Hypothesis'
Representation of the revised 'Signal Hypothesis', depicting co-translational translocation into the ER mediated by SRP and SRP Receptor (Meyer,1982; Walter *et al*,1984).
(a) A transmembrane protein with a cleaved N-terminal signal peptide.
(b) An integral membrane protein with two internal signal sequences.
Symbols include those used in Figure I.1 plus:

SRP

SRP Receptor

reconstituted translocation activity when added back to extracted membranes. This 11S (M_r=250,000) protein, termed the Signal Recognition Protein, is comprised of six polypeptide subunits M_r 72,000, 68,000, 54,000, 19,000, 14,000, and 9,000 in equal stoichometry; and at least one sulphydryl group is required for its activity. Later it was discovered that the signal recognition protein also contains a 7S RNA molecule which is essential to its structure and function (Walter & Blobel,1982; Walter & Blobel,1983a), hence the ribonucleoprotein complex was renamed Signal Recognition Particle (SRP). The 7S RNA was found to be the cytoplasmic 7SL RNA studied by Ullu *et al*(1982) and sequenced by Li *et al*(1982); its organisation in the SRP has been elucidated by Gundelfinger *et al*(1983). Extensive *in vitro* studies using the wheat germ cell-free system were carried out and from these a picture was built up of the role of SRP in translocation and of its relationship to the other microsomal component that has been characterised. This is described below and in the reviews of Meyer(1982) and Walter *et al*(1984), and is depicted in Figure I.2.

When the signal sequence of a secretory protein emerges from the large subunit of the ribosome it interacts specifically with SRP, which binds selectively to wheat germ ribosomes synthesizing secretory proteins, but not cytoplasmic proteins (Walter *et al*,1981). Cell fractionation studies have shown that SRP is associated with the cytoplasm, and both membrane-bound and free ribosomes (Walter & Blobel,1983b). The interaction of SRP with the signal peptide results in a site-specific arrest of translation which is released on the selective and SRP-mediated binding of the elongation-arrested ribosome to microsomal membranes (Walter & Blobel,1981a and b). Completion of the nascent polypeptide chain coupled to translocation into the microsomal vesicle then takes place.

A second ER specific membrane protein was also found to be required for vectorial transfer of secretory proteins; this 72kd protein is inactivated by sulphydryl group blocking agents and has a cytoplasmic domain of 60kd which is released from the membrane by certain proteases, resulting in abolition of translocation activity (Meyer & Dobberstein,1980a and b; Meyer *et al*,1982a). Meyer *et al*(1982b) subsequently reported that this 72kd protein interacted with SRP and relieves the SRP-induced arrest of translation seen in the wheat germ system, allowing translation to proceed; accordingly they called this protein Docking Protein. Independent work by Gilmore *et al*(1982a and b) also identified this 72kd integral membrane protein as responsible for releasing the SRP-mediated elongation arrest of nascent secretory

polypeptides, but they refer to the protein as the Signal Recognition Particle Receptor. Recently the function of different proteolytic fragments of the Docking Protein was analysed and a 13kd part of the cytoplasmic domain was identified as essential for the membrane association of the SRP-ribosome complex and release of translational arrest (Hortsch *et al*,1985). Further information on the topology and functional domains of the SRP Receptor was gained from studies by Lauffer *et al*(1985), which included the isolation and sequencing of a SRP Receptor cDNA clone. It appears that a transient interaction of the ribosome-bound SRP with the SRP Receptor in the membrane causes the displacement of SRP from the ribosome, but it does not result in binding of the SRP Receptor to the translating ribosome (Gilmore & Blobel,1983). Recent *in vitro* studies using protein denaturants indicate that other ER membrane proteins, besides SRP Receptor, are responsible for signal sequence binding to the membrane and the subsequent translocation of the nascent chain (Gilmore & Blobel,1985).

The discovery of SRP and its properties, of binding to the signal sequences of nascent secretory polypeptides and causing translation arrest until it mediated the interaction with the ER membrane via the SRP Receptor, could all be very neatly accommodated in an updated Signal Hypothesis (see Figure 2). Whilst the concept of a translational arrest of proteins destined for translocation seems sensible energetically, as it ensures synthesis does not occur in the absence of translocation, not all the *in vitro* experimental data was consistent with this idea and there is increasing evidence that the model of translation-arrested insertion may not be valid either for all *in vitro* cell-free systems or *in vivo* (reviewed and discussed in Hortsch & Meyer,1984). It has been found that SRP is also required for the integration into microsomes of integral membrane proteins, for example acetylcholine receptor δ-subunit (Anderson *et al*,1982), calcium ATPase, lens plasma membrane MP26 (Anderson *et al*,1983), cytochrome P-450 (Sakaguchi *et al*,1984) and corona virus E1 (Rottier *et al*,1985). However a SRP-induced elongation arrest in the wheat germ system was only demonstrated for the acetylcholine receptor δ-subunit, which contains a cleavable signal sequence, and for the uncleaved corona virus E1 protein; the translation of the other two proteins, which also have permanent signal sequences, was not inhibited by SRP in the wheat germ system. The phenomenom of translation arrest has in fact only been observed in the wheat germ cell-free system. It has been shown that the same mRNAs which display SRP-induced elongation arrest in the wheat germ are freely translated in both a rabbit

reticulocyte lysate and a He-La cell-free system, in the presence of endogenous and high exogenous levels of SRP; and in these systems the secretory proteins are translocated and processed by SRP-depleted microsomes (Meyer *et al*,1982b; Meyer,1985). It is suggested that the phenomenom of the SRP-induced arrest in wheat germ is an artefact due to the system being reconstituted from heterologous components - canine SRP and membranes, with plant ribosomes (Hortsch & Meyer,1984; Meyer,1985). Clearly caution should be taken in interpreting the results obtained from *in vitro* systems, and extending them to the translocation process *in vivo*. It also is evident that not all the components of the translocation machinery are identified or fully characterised.

Early work with several *in vitro* systems (Milstein *et al*,1972; Blobel & Dobberstein,1975b; Szczesna & Boime,1976; Walter *et al*,1979) established the location, on the lumenal side of the ER membrane, of a protease which removes the signal peptide during translocation; this enzyme was termed 'signal peptidase', and does not show species-specificity. Detergent solubilised signal peptidase from dog pancreas microsomes correctly cleaved *in vitro* synthesized full-length preprolactin, but less than 50% of these precursors reacted with the solubilised enzyme (Jackson & Blobel,1977). None of a range of inhibitors of proteolytic enzymes were found to inhibit the activity of the solubilised signal peptidase (Jackson *et al*,1980), but it was found that its activity depended on phospholipid (Jackson & White,1981). Using hen oviduct RER membranes Lively & Walsh(1983) developed a protocol for solubilising and partially purifying signal peptidase, which they found to be an integral membrane protein; the solubilised enzyme retained the activity and specificity of the membrane-bound form. Perlman & Halvorson(1983) analysed the distribution of amino acid residues around the signal sequence cleavage site of several proteins, and observed that a β-turn occurs in the polypeptide chain near the cleavage site and that Ala-X-Ala is the most frequent sequence preceeding the cleavage site. They proposed a signal peptidase recognition site of A-X-B↓, where ↓ is the site of cleavage, X is any amino acid, B is Ala, Gly or Ser, and A can be B-type residues or larger aliphatic amino acids - Leu, Val and Ile. Von Heijne(1983,1984a) has made similar comparisons and he proposed a '(-3,-1) rule' for a functional cleavage site: the last amino acid of the signal peptide (position -1) must be Ala, Ser, Gly, Cys, Thr or Gln, position -3 must not be a large polar(Asn, Gln), aromatic or charged residue, and no Pro residues are permitted in the region -3 to +1. Using hydrophilic analogues of leucine and threonine Hortin & Boime(1980,1981)

have shown that incorporation of these analogues into secretory proteins *in vitro* and *in vivo* may result in inhibition of segregation and processing, or miscleavage of the precursor; although it should be pointed out that amino acid substitutions also occurred in regions other than the signal sequence. Their data indicates that modification of the structure of the secretory precursor can alter or block the signal peptidase cleavage site, but that the complete and accurate removal of the signal peptide is not required for secretion.

The updated signal hypothesis and alternative models of protein translocation

It was mentioned earlier that other mechanisms, besides the Signal Hypothesis, have been proposed to account for the compartmentation of proteins into the ER. One aspect of the original Signal Hypothesis which has provided controversy is the obligatory coupling of translation to translocation. As described earlier several *in vitro* experiments showed the co-translational translocation of a number of secretory and membrane proteins; however it does not appear that the segregration of all membrane proteins is a strictly co-translational or SRP-mediated event. An example is the behaviour of cytochrome b_5 which is anomalous in several respects; not only is it synthesized on free ribosomes (Rachubinski *et al*,1980), but its integration into microsomes was found to be independent of SRP (Anderson *et al*,1983). It was also found that cytoplasmically synthesized mitochondrial and chloroplast proteins were post-translationally translocated *in vitro* into their respective organelles (see A.4.). Although many features of the Signal Hypothesis were found to apply to the export of proteins in prokaryotes, thus implying a universal translocation mechanism (discussed in A.5); it was observed that several exported prokaryotic proteins were segregated post-translationally. For example the precursor of the M13 phage coat protein was shown to be synthesized complete prior to its conversion into the processed, membrane-bound protein (Ito *et al*,1980). Consideration of the above experimental data and theoretical calculations based on the observed structure of the signal peptide has given rise to several alternative theories to the Signal Hypothesis.

The Membrane Trigger Hypothesis, pioneered by Wickner(1979,1980), proposes the role of the signal peptide is to allow the growing polypeptide chain to fold in a manner compatible with the aqueous environment of the cytoplasm and then, on binding to the appropriate membrane, a conformational change occurs exposing hydrophobic residues.

thus facilitating transfer of the nascent polypeptide across the lipid bilayer. This model is distinct from the Signal Hypothesis since neither protein transport components nor a specific ribosome-membrane interaction is required; and it is envisaged that the signal peptide interacts with sequences present in the mature protein. The Helical Hairpin Hypothesis (Engelman & Steitz,1981) and the Direct Transfer Model (Von Heijne & Blomberg,1979; Von Heijne,1981a) were based on theoretical considerations of the energetics involved in the transfer of a polypeptide chain from a polar to a non-polar environment; it is proposed that the hydrophobic signal sequence can directly and spontaneously penetrate the lipid bilayer to initiate translocation. Like the Membrane Trigger Hypothesis these models do not envisage the specific interaction of the signal sequence with membrane proteins which is a key point of the Signal Hypothesis, and these theories have largely been outdated by the discovery of SRP and SRP receptor. However, as yet, no membrane proteins have been characterised which are directly involved in binding the signal sequence in the context of the ribosome to the membrane, or in the transfer of the nascent chain across the ER membrane.

The concept, contained in the signal hypothesis, that a discrete region of a polypeptide, the signal sequence, is responsible for targetting a protein to the ER was subsequently extended in the theory of topogenic sequences of Blobel(1980), which proposes the subcellular location and orientation of all compartmented proteins is directed by a limited number of discrete and autonomous elements of the polypeptide; this is discussed fully in section B.

It should be noted that the evidence for co-translational translocation of eukaryotic proteins is purely derived *in vitro*, and the evolving concept of membrane translocation overall is no longer of a universal mechanism which covers all systems, but of 'variations on a theme' which range from a co-translational to an entirely post-translational process (Wickner & Lodish,1985). Whilst the initial steps of the translocation of secretory and membrane proteins into the ER have been partially elucidated, mainly from *in vitro* work, it is not understood how the nascent chain is subsequently transferred across the membrane bilayer, although it is known that this is an energy independent process. It is possible that facets of the other models outlined above may ultimately be incorporated into the Signal Hypothesis to provide a complete picture of the translocation process.

I.A.3. Transport of Proteins Beyond the Endoplasmic Reticulum

The primary concern of this thesis is the initial step in the secretory pathway - the translocation of proteins into the ER, which has been dealt with in the previous section. Here I will consider in less detail the subsequent steps in the localisation of secretory and membrane proteins; more information can be found in the reviews cited in A.1., and those of Farquar & Palade(1981), Tartakoff(1982), Dunphy & Rothman(1985) and Rothman(1985) which deal with the Golgi apparatus, and Kelly(1985) which discusses constitutive and regulated pathways of protein secretion.

Endoplasmic reticulum to Golgi complex

In addition to the proteolytic cleavage of transient signal peptides two other types of modifications can occur to translocated proteins in the rough ER (RER); these are the formation of intra- and intermolecular disulphide bridges, and the addition, plus first processing stages, of asparagine-linked glycosylation. It is possible that fatty acylation also occurs in the ER.

Glycosylation is a characteristic, although not a universal, feature of secretory and membrane proteins, and experiments carried out by Rothman and his coworkers (Rothman & Lodish,1977; Rothman *et al*,1978) showed that glycosylation of the Vesicular Stomatitis Virus transmembrane glycoprotein (VSV-G) occurred on the nascent, growing polypeptide chain, but that this co-translational glycosylation was not essential for membrane insertion. The initial step in the glycosylation of proteins involves the transfer of an oligosaccharide precursor to asparagine residues in the nascent chain which must be part of the sequence -Asn-X-Ser- or -Asn-X-Thr-, where X is any amino acid except Pro; this co-translational N-glycosylation takes place in the RER (see Hanover & Lennarz,1981). The Asn-linked oligosaccharides are subsequently post-translationally processed and modified to form complex oligosaccharide side chains, in an ordered assembly process which is initiated in the RER but is largely carried out in the Golgi complex (reviewed in Kornfeld & Kornfeld,1985). The removal of three glucose and one mannose residues from the primary side chain (which consists of three glucose, nine mannose and two N-acetylglucosamine residues) occurs in the RER. It was found that the maturation of certain secretory proteins from the RER to the Golgi body depended on early steps of the Asn-linked oligosaccharide processing taking place; but the inhibition of the trimming of N-linked oligosaccharides did not affect the transfer of other secretory proteins, or interfere with the surface expression of some integral membrane

proteins (Lodish & Kong,1984; Burke *et al*,1984). The passage of proteins from the RER to the Golgi complex often seems to be the rate limiting step in intracellular transport but it does not appear to be a synchronous process; instead several secretory proteins (glycosylated and non-glycosylated) have been found to migrate at different, characteristic rates, which show up to ten-fold variation (Lodish *et al*,1983; Fries *et al*,1984; Scheele & Tartakoff,1985). This suggests that the transfer of proteins from the RER to the Golgi is not a bulk-phase movement but is selectively mediated by membrane-bound receptors, and transport vesicles are thought to be involved. Studies in yeast (reviewed in Scheckman,1982) have shown that the transport of proteins from the RER to the Golgi complex requires energy and the involvement of at least 9 gene products.

The Golgi complex

The Golgi apparatus plays a crucial role in the processing and sorting of polypeptides (reviewed in Farquar & Palade,1981; Tartakoff,1982; Dunphy & Rothman,1985 and Rothman,1985). The Golgi complex typically consists, in mammals, of an asymmetric stack of flattened, smooth membrane-bound cisternae, and recent work shows it is organised into three functionally and compositionally distinct compartments (Griffiths *et al*,1983; and see also reviews of Dunphy & Rothman,1985; Rothman,1985).

The step wise processing of N-linked oligosaccharides has been elucidated using a variety of biochemical and immunocytochemical techniques (see Dunphy & Rothman,1985; Rothman,1985; Kornfeld & Kornfeld,1985); it occurs in a strict sequence of events as proteins enter the *cis* face of the Golgi apparatus from the RER, bearing the same oligosaccharide chains, and then pass through the *cis*, *medial*, and *trans* compartments where processing and construction of oligosaccharides is carried out. Lysosomal proteins are distinguished from secretory and membrane proteins by the specific phosphorylation of mannose residues on at least one oligosaccharide side chain, which probably takes place in the *cis* cisternae (Reitman & Kornfeld,1981; Varki & Kornfeld,1981; Waheed *et al*,1981a and b). The processing of other protein side chains involves the removal of mannose residues and the addition of N-acetylglucosamine in the *medial* compartment, followed by addition of galactose and sialic acid residues in the *trans* cisternae. O-linked glycosylation of serine, threonine or tyrosine residues also takes place in the Golgi complex (Hanover & Lennarz,1981).

Individual proteins can display a heterogeneity in their glycosylation patterns and it is not clear to what extent the accurate and specific glycosylation of membrane and secretory proteins is necessary for their correct localisation; the experimental data indicates that each protein needs to be considered individually. It has been suggested that oligosaccharide side chains may form part of the recognition site for transport receptors (Olden *et al*,1982), although experimental evidence for such receptors has yet to be published. As mentioned earlier, it was found that inhibition of trimming of N-linked oligosaccharides did not interfere with the surface membrane expression of human class 1 histocompatibilty antigens, influenza haemagglutinin and VSV-G (Burke *et al*,1984), yet inhibition of the RER processing step prevented the secretion of α1-antitrypsin and α1-antichymotrypsin, but not of transferrin, C3 and albumin - since albumin is unglycosylated this last result is not surprising (Lodish & Kong,1984). It was also found that inhibition of N-glycosylation of influenza neuraminidase greatly reduced its appearance at the cell surface when neuraminidase cDNA cloned into a SV40 expression vector was infected into cells (Markoff *et al*,1984). Experiments with α1-acid glycoprotein showed that non- or partially glycosylated molecules were secreted more slowly than the fully glycosylated native form; when a threonine analogue, which interferes with Asn-linked glycosylation, was incorporated into the protein both glycosylation and secretion were inhibited, but it was not possible to distinguish what effects on the intracellular transport were due to changes in the peptide backbone or to the extent of glycosylation (Docherty & Aronson,1985). The transmembrane protein VSV-G has been widely studied as a model for glycoprotein biosynthesis, it was found non-glycosylated VSV-G failed to reach the plasma membrane when tunicamycin was used to block N-linked glycosylation (Gibson *et al*,1978; Morrison *et al*,1978). However it has recently been shown that N-glycosylation at either of the two normal sites is sufficient to transport VSV-G to the cell surface nearly as efficiently as wild-type doubly glycosylated VSV-G; but that nonglycosylated mutants were trapped in the Golgi (Machamer *et al*,1985).

The specific phosphorylation of mannose residues in lysosomal enzymes is recognised by a mannose-6-phosphate receptor and is important for their correct routing within the cell (see Sly & Fischer,1982). The role of glycosylation as a protein sorting signal is discussed further in section B.

Sorting of proteins after the endoplasmic reticulum

It has been proposed, as a simple model, that in the absence of other specific signals proteins sequestered in the ER ultimately get secreted at the cell surface, whilst retention at some location along the secretory pathway requires ancillary signals. Experimental evidence which supports this hypothesis, ie. Wiedman *et al*(1984), Poruchynsky *et al*(1985), is discussed later (I.B.5). It is expected, however, that certain properties of the secretory polypeptide would be necessary to enable its transit through the cell, such as solubility in the milieu of the ER and Golgi complex. Two alternative secretion pathways exist - constitutive and regulated, the latter occurs in specialised secretory cells but both pathways can be found in the same cell (reviewed by Kelly,1985). Whilst the above model may be true for constitutive secretion, in cells which display regulated secretion secretory proteins destined for storage must contain information to target them to the storage secretory vesicles until secretion is triggered. It is envisaged that membrane proteins are treated like secretory proteins but are anchored in the membrane by a segment(s) of the polypeptide which cannot translocate the ER membrane and, unless retained along the secretory pathway, membrane proteins will appear in the plasma membrane. At some stage in polarised cells proteins destined for the basolateral and apical domains of the plasma membrane must be segregated and correctly routed. ER membrane proteins must contain information which specifies that they remain as components of the ER; and a parallel situation must occur for proteins which are components of the Golgi apparatus. Lysosomal enzymes are also diverted from the route to the cell surface.

The mechanisms by which proteins travel through the Golgi complex and to the cell surface are not yet understood, nor is it clear at what stage proteins destined for different subcellular compartments are routed to their correct location. Using temperature sensitive mutants, which enabled the synchrony of transport, and immunoelectron microscopy Bergmann & Singer(1983) showed the transfer of VSV-G from the ER to the Golgi (possibly in small, 50-70nm, vesicles), and then to the plasma membrane via vesicles of 200nm diameter. Rothman *et al*(1984) concluded from their experiments that transport through the Golgi was a vectorial process in which successive intercisternal transfer occurs by the budding and fusing of transport vesicles. Immunoelectron microscopy was also used by Strous *et al*(1983) in experiments which demonstrated that VSV-G, a protein which appears at the cell surface, is present in the same RER cisternae, Golgi compartments and secretory granules as both a secretory

glycoprotein (transferrin) and a non-glycosylated secretory protein (albumin). Experiments by Gumbiner & Kelly(1982) with specialised secretory cells, which have both constitutive and regulated pathways of protein secretion, found that viral envelope proteins reached the surface in vesicles different from the secretory granules containing mature adrenocorticotrophic hormone. In another study, by Green & Shields(1984), which used growth-hormone secreting cells the results suggested that in these cells the membrane and regulated secretory proteins are sorted into distinct compartments either late in, or after exit from, the Golgi; these workers found that exogenously added somatostatin (which inhibits the secretion of various peptide hormones) selectively inhibited the secretion of growth hormone without reducing the appearance at the cell surface of VSV-G in the same cells. Work in polarised Madine Darby Canine Kidney cells has shown that VSV-G, which is directed to the basolateral membrane, is found in the same *trans* Golgi compartment as influenza haemagglutinin or neuraminidase, which appear in the apical domain; these findings indicate that the divergence of the transport pathway for apical and basolateral membrane proteins occurs at a later intracellular stage (Fuller *et al*,1985; Rindler *et al*,1985; Pfeiffer *et al*,1985). The Mannose-6-Phosphate Receptor which binds to the specifically phosphorylated lysosomal enzymes is concentrated in the *cis* Golgi cisternae when phosphorylation is taking place (Brown & Farquar,1984), but as some lysosomal enzymes are sialated this indicates they pass through the *trans* Golgi. Recent immunoelectron microscopy experiments showed coated vesicles budding off the *trans* Golgi were involved in the transport of lysosomal enzymes from the Golgi to the lysosomes (Geuze *et al*,1985).
There are also some reports that coated vesicles are involved in the transport of newly synthesized membrane proteins to the cell surface (eg. Bursztain & Fischbach,1984).

I.A.4. Post-translational Transport into Organelles

In contrast to the co-translational translocation of proteins across the ER, the compartmentation of proteins to other organelles in eukaryotic cells - nucleus, mitochondria, chloroplasts, peroxisomes and glyoxisomes - is a post-translational process. As such these sorting mechanisms are distinct from those which are adopted by proteins following (all or part of) the secretory pathway. Since different organelles exist within the same cell mechanisms must distinguish proteins specific to each type of organelle, and their correct suborganellar location. Details of each system can be found in the reviews cited below and in Lodish *et al*(1981)

and Strauss & Boime(1982). Here I will only briefly summarise what is understood of the localisation of polypeptides to the mitochondrion and chloroplast, to provide a comparison of the features of these post-translational translocation mechanisms with the process of translocation across the ER (described in I.A.2), and to protein translocation in prokaryotes which is described in the next section (A.5). The transport of proteins to the nucleus is quite different to the translocation of proteins into other organelles and to translocation across the ER, as nuclear proteins do not translocate the membrane to enter the organelle but pass instead through pores in the nuclear envelope. A discussion of the theories and experimental information on transport of proteins to the nucleus can be found in the reviews of Bonner(1978), De Robertis(1983) and Dingwall(1985).

Import into mitochondria

The import of proteins into mitochondria has been the most widely studied post-translational sorting process; the experimental evidence and theories concerning the localisation of mitochondrial proteins are discussed fully in the reviews of Chua & Schmidt(1979), Neupert & Schatz(1981), Schatz & Butow(1983) and Hay *et al*(1984). Proteins are found in four compartments within the mitochondrion - the outer membrane, the intermembrane space, the inner membrane and the matrix. Although the mitochondrial genome encodes some mitochondrial proteins most are translated from nuclear-encoded transcripts on polysomes in the cytoplasm (Suissa & Schatz,1982). The newly synthesized polypeptides are subsequently transported to and recognised by the mitochondria prior to translocation across one or both mitochondrial membranes, with processing and assembly to form functional proteins. It is evident now that different mitochondrial proteins use disparate mechanisms of import; but in general the transport of proteins to the intermembrane space, inner membrane and matrix share common features which are distinct from the integration of outer membrane proteins. Details of the translocation mechanisms are not known, neither is it clear how the sorting process discriminates proteins destined for different compartments.

Like secretory proteins the primary translation product of most mitochondrial proteins is a precursor with an N-terminal extension relative to the mature polypetide (Maccheccchini *et al*,1979). Since these precursors are detected *in vivo* this demonstrates that vectorial translation is not obligatory to import. In contrast with the signal peptide of secretory and membrane proteins the N-terminal extension is

quite long (40-70 amino acids) and polar (Viebrock *et al*,1982; Watson,1984). Not all mitochondrial proteins are synthesized as precursors, an exception is the apoprotein of cytochrome c (Zimmerman *et al*,1979), but there is evidence that these polypeptides and those with N-terminal extensions adopt a configuration in the cytoplasm which is different to their mature form, and this may be pertinent to the import process. Precursors, but not the mature forms, of several mitochondrial proteins have been shown to specifically associate with the outer membrane and then internalise into mitochondria; translocation, but not binding, required an energised membrane (Riezman *et al* ,1983). Apocytochrome c, however, binds to a different receptor and does not require an electrochemical potential for import to the inner membrane (Zimmerman *et al*,1981; Hennig *et al*,1983). Whilst apocytochrome c does not require energy for its import to the inner membrane, the transport of other proteins to the same location, or to the intermembrane space and matrix, is energy dependent (Gasser *et al*,1982). It appears that the transfer of precursors into the inner membrane or matrix space occurs through 'translocation contact sites' where the precursor can span both inner and outer membranes (Schleyer & Neupert,1985). Proteolytic processing of precursors also occurs for polypeptides with different destinations. In the case of cytochrome b_2 (an intermembrane space enzyme) and cytochrome c_1 (inner membrane protein) the cleavage process is in two steps; the first requires an energised inner membrane and results in a membrane-bound intermediate, this is cleaved by a second enzyme yielding the mature form correctly localised (Daum *et al*,1982; Ohashi *et al*,1982).

Distinctive features are common to the transport of proteins to the outer mitochondrial membrane. These proteins are not synthesized as larger precursors, nor does insertion require energy or any covalent modification of the protein, and there is little evidence for the involvement of receptors. It appears that binding and integration are separate events with a conformational difference between the bound and integrated forms (Gasser & Schatz,1983).

Recent use of recombinant DNA techniques to create chimaeric proteins has led to the identification of specific regions involved in the targetting and cleavage of certain mitochondrial proteins; studies include proteins normally located in the outer membrane (Hase *et al*,1984) and matrix (Hurt *et al*,1984). This evidence for the involvement of topogenic

sequences in the localisation of mitochondrial proteins will be discussed in section B.4.

Transport of proteins into chloroplasts

Less is known about the transport of proteins into chloroplasts; the published work is discussed in the reviews cited at the beginning of this section and in Chua & Schmidt(1979) and Robinson(1984). The chloroplast structure is similar to the mitochondrion, both being enclosed by two delimiting membranes; but in chloroplasts the thylakoids constitute a third suborganellar membrane with its enclosed thylakoid space. Most chloroplast proteins are synthesized in the cytoplasm and many features of protein transport into this organelle show parallels to the import of mitochondrial proteins. Much of the work done has concentrated on the most abundant chloroplast protein the nuclear-encoded small subunit of ribulose bisphosphate carboxylase (ssRuBPC). The primary translation product of this polypeptide was found to be a precursor, about 5000 dalton larger than the mature form (Dobberstein *et al*,1977). It became clear that the signal hypothesis could not account for the uptake of ssRuBPC into the chloroplast since *in vitro* reconstitution experiments demonstrated the completed precursor was taken up and processed by intact, isolated chloroplasts in the absence of protein synthesis (Highfield & Ellis,1978). Instead a post-translational transport model has been proposed in which the precursors released from cytoplasmic ribosomes then bind to specific chloroplast outer membrane receptors prior to uptake and processing. Specific binding of chloroplast protein precursors to isolated chloroplast envelopes has been demonstrated, and this did not need energy (Pfisterer *et al*,1982). The uptake process does appear to be energy-dependent, but in contrast to the import of mitochondrial proteins the requirement was for ATP and not for a membrane potential. To date all the nuclear-encoded chloroplast polypeptides studied have been found to be synthesized as a larger precursor with an N-terminal extension, of 2000-5000 daltons, (see Robinson,1984) which is termed the transit sequence or peptide. A few transit peptides have been sequenced (see Watson,1984) and these are over 30 residues long and rich in hydrophilic amino acids. Removal of the transit peptide was found to prevent the uptake and processing of ssRuBPC by isolated, intact chloroplasts (Mishkind *et al*,1985). It appears that the transit peptides are processed in two steps by a stromal enzyme during import of the precursor into the chloroplast, yielding the mature polypeptide (Smith & Ellis,1979; Robinson & Ellis,1984; Mishkind

et al,1985). Recent experiments indicate that the transit sequence does contain all the information necessary for correct localisation of chloroplast proteins, since foreign proteins can be targetted to the chloroplast by fusion to the transit peptide of ssRUBPC (Van den Broeck *et al*,1985; Schreier *et al*,1985); these results are discussed later (section B.4.).

I.A.5. Sorting of Proteins in Prokaryotes

Whilst in prokaryotes there is not the range of subcellular organelle compartments which exist in the eukaryotic cell, transport mechanisms are still required for traffic of proteins to the plasma (inner) membrane, periplasm, and outer membrane of the Gram negative bacterial envelope, and for secretion. It is now apparent that protein sorting in prokaryotes does share common themes with the transport of eukaryotic proteins, but that a spectrum from co-translational to entirely post-translational translocation occurs through the same membrane. The aim here is to compare and contrast the two systems with a view to discerning to what extent results obtained in prokaryotic systems may provide information about protein transport in eukaryotes, in particular the translocation of proteins across the ER. Not surprisingly most work has been done with the Gram negative *Escherichia coli*. Reviews concerned with the sorting of proteins in bacteria can be found in Emr *et al*(1980); Michaelis & Beckwith(1982); Silhavy *et al*(1983); Wickner(1983); Benson *et al*(1985); Oliver(1985a) and Wickner & Lodish(1985).

Support for co-translational translocation in bacteria

Analysis of nascent polypeptides associated with polysomes in *E.coli* showed that, as in eukaryotes, membrane-bound polysomes were synthesizing exported proteins, including periplasmic, outer membrane and secreted proteins (Cancedda & Schlesinger,1974; Randall & Hardy,1977). In 1977 Smith *et al* provided direct evidence of vectorial co-translational translocation *in vivo* in prokaryotes by demonstrating that growing nascent polypeptide chains could be labelled by an extracellular reagent which did not cross the membrane of the *E.coli* spheroplasts; one of the major labelled products was the periplasmic protein alkaline phosphatase previously found to be synthesized on membrane-bound ribosomes. In subsequent *in vitro* translation studies, which paralleled those of Blobel & Dobberstein(1975a and b) with a eukaryotic system, Smith(1980) found that the precursors of two exported prokaryotic proteins, diptheria toxin and alkaline phosphatase, were segregated within and processed by

inverted *E.coli* inner membrane vesicles if these were added early in translation. If the cytoplasmic face of these vesicles was treated with pronase they were no longer translocationally competent, indicating the involvement of a cytoplasmically exposed inner membrane protein in the transmembrane transfer; outer membrane vesicles could not mediate translocation. At that time such close agreement with the postulates of the signal hypothesis led to the idea that the mechanism of membrane translocation was universal to prokaryotes and eukaryotes. The finding that cleaved N-terminal leader peptides of secreted, periplasmic and outer membrane proteins of prokaryotes were similar in structure to signal sequences of eukaryotic secretory and membrane proteins added weight to this idea (see Watson,1984).

Although it is now clear that there is not a single mechanism that applies to the transfer of all proteins across membranes, there is experimental evidence that the processes of translocation across the plasma membrane of prokaryotes and across the RER in eukaryotes are closely related, since components from one system will recognise and functionally interact with components of the other. Fraser & Bruce(1978) reported that chick ovalbumin synthesized in *E.coli*, under the transcriptional and translational control regions of the lac gene, was secreted through the cell membrane into the periplasmic space. Talmadge *et al*(1980,1981) studied the localisation in bacteria of a series of constructs containing fusions of different amounts of the periplasmic protein β-lactamase and the eukaryotic secretory protein preproinsulin. Not only was proinsulin, lacking its signal peptide, secreted by *E.coli* if fused to the signal sequence of pre-β-lactamase, but it was also found that if all the leader peptide of the bacterial protein was replaced by the signal sequence of preproinsulin the hybrid protein was still secreted. Consistent with results obtained with eukaryotic proteins (see A.2.) Müller *et al*(1982) translated *in vitro* synthesized pre-β-lactamase RNA in a wheat germ cell-free system and demonstrated both a SRP-mediated translation arrest and SRP-dependent translocation into canine pancreatic microsomes, where the bacterial precursor was correctly cleaved. Likewise Kronenberg *et al*(1983) found pre-β-lactamase was sequestered and accurately processed by pancreatic microsomes in a rabbit reticulocyte lysate *in vitro* translation system. In addition Wiedman *et al*(1984) have reported the secretion of pre-β-lactamase from *Xenopus* oocytes, following microinjection of *in vitro* synthesized capped RNA. These and other examples (ie. Lingappa *et al*,1984; Gray *et al*,1985)

suggest a conserved mechanism of protein transport in eukaryotes and prokaryotes.

Evidence against ER-like translocation in prokaryotes

Studies of protein localisation in prokaryotes have tended to focus on a limited number of proteins, and one of these is the coat protein of the bacteriophage M13. This protein has been the subject of much work by Wickner's group which has demonstrated that its mechanism of membrane insertion is quite distinct from an ER-like co-translational process. The M13 coat protein is integrated into the inner membrane of infected *E.coli* prior to forming the viral coat. Although the primary translation product of the coat protein, termed procoat, contains a 23 amino acid N-terminal leader peptide which resembles that of eukaryotic secretory proteins (Sugimoto *et al*,1977), it was found by Ito *et al*(1979 and 1980) that this precursor was synthesized on free ribosomes in the cytoplasm and the soluble M13 procoat was post-translationally assembled into the cytoplasmic membrane and processed to the mature coat protein. In contrast also to integration into the RER this membrane insertion requires an electrochemical membrane potential, but the binding of procoat to the inner face of the plasma membrane is not prevented by uncoupling agents (Date *et al*,1980a and b). *In vitro* experiments showed that phospholipid vesicles containing the bacterial processing enzyme leader peptidase (discussed later) were able to post-translationally bind and cleave M13 procoat, and some of the processed coat protein was inserted into the liposomes and spanned the bilayer; this suggests these are the only components required for binding and processing of procoat (Watts *et al*,1981). Furthermore M13 procoat is integrated as transmembrane coat protein into canine pancreatic microsomes *in vitro* under conditions when eukaryotic secretory protein precursors can not segregate (Watts *et al*,1983). So it is clear that the membrane insertion of this prokaryotic protein is not in accordance with the Signal Hypothesis, and the M13 procoat data provided the basis for the Membrane Trigger Hypothesis proposed by Wickner(1979,1980).

It has been found that an energised membrane is required for export of periplasmic and outer membrane proteins (Enquist *et al*,1981) and for a typical inner membrane protein which does not have a transient leader peptide (Wolfe & Wickner,1984). So it appears that, as with the import of proteins to the mitochondria, a membrane potential is essential for translocation of proteins in prokaryotes; but this is in contrast to protein translocation across the ER.

It became clear that for exported prokaryotic proteins there is no strict coupling of translocation to elongation of nascent polypeptide chains *in vivo*. By looking at the timing of signal sequence processing it appeared that, in *E.coli*, outer membrane and periplasmic precursors were processed at characteristically different stages during synthesis (Josefsson & Randall,1981). Cleavage may either occur whilst translation is still in progress or after the precursor is complete, with some proteins showing both the co- and post-translational modes of processing. Where cleavage occured co-translationally the proteolytic removal of the leader peptide was initiated as late as when 80% of the precursor had been synthesized. These findings were supported by further work by Randall(1983) who studied the timing of translocation using the criterion of the accessibilty of nascent chains of periplasmic proteins to externally added proteases. She showed that for one protein, ribose binding protein, translocation occurs entirely post-translationally but for another, maltose binding protein, it is a late event initiated only after 80% of the protein has been synthesized. She proposed that entire domains of polypeptide are transferred across the membrane after their synthesis, instead of a continuous process of translocation proceeding with elongation of the nascent polypeptide as envisaged in the signal hypothesis. Whatever the precise mechanism of transfer *in vivo* it is clear that the different exported prokaryotic proteins form a spectrum across the range from co- to post-translational translocation.

Compartmentation of proteins in bacteria

The process of protein localisation in prokaryotes has been much studied using mutants generated by bacterial genetics and recombinant DNA techniques. However, the categorisation of proteins as being localised in the periplasmic space, inner or outer membranes has generally been on the basis of cell fractionation data which is not always definitive. From a number of experiments in which the leader peptide of prokaryotic proteins has been altered it has been established that the presence of a functional signal sequence is crucial for export from the cytosol; these results are discussed in section B.2. There is, however, evidence that the signal peptide alone is not suficient to achieve export of prokaryotic proteins. Moreno *et al*(1980) found the signal sequence of the outer membrane protein LamB fused to the LacZ gene for the cytosolic protein β-galactosidase resulted in a hybrid protein which was not exported from the cytoplasm, although it is possible this is a consequence of the structure of β-galactosidase being non-permissive for membrane

translocation. Other fusions of various N-terminal regions of LamB onto LacZ did appear to be exported and some were localised in the outer membrane; but in addition to the leader peptide part of the mature LamB protein was required (Benson & Silhavy,1983; Benson *et al*,1984). Work by Koshland & Botstein(1980, 1982) showed that in *Salmonella typhimurium* the secretion of β-lactamase into the periplasmic space needed the COOH-end of the protein, but chain terminator mutants lacking this region still translocate the inner membrane.

Mechanisms are required to distinguish proteins destined for the inner membrane, periplasm, outer membrane and extracellular medium, but it is not yet clear how this differential routing is achieved. It is thought translocation of proteins in the bacterial cell may occur via sites of adhesion between the inner and outer membranes. To date all outer membrane and periplasmic proteins have been found to have a cleaved N-terminal signal sequence and it is thought that these proteins share at least initial steps in the export process. Inner membrane proteins are a more diverse group, many are not synthesized as precursors and some are very hydrophobic and may spontaneously integrate into the lipid bilayer (see reviews cited at beginning of this section for examples). However one inner membrane protein, leader peptidase, which is not synthesized as a precursor does appear to use the same export pathway as the precursors of outer membrane and periplasmic proteins (Wolfe & Wickner,1984). It was recently shown (Jackson *et al*,1985) that if the cleaved N-terminal pre-sequence (which is of relatively low hydrophobicity) of the inner membrane penecillin binding protein 5 replaces the signal sequence plus the first 11 amino acids of an outer membrane protein (OmpF) the fusion protein is still directed to the outer membrane. This sugests that this inner membrane protein uses the same export machinery as outer membrane proteins, but that information for sorting proteins to the outer and inner membranes does not reside in the signal sequence or the extreme N-terminus of the mature protein. Further evidence implying the function of the signal peptide is to initiate translocation, and that other signals are required to retain proteins in the inner or outer membranes and periplasm comes from work by Nagahari *et al*(1985). They found that if the region of OmpF encoding the leader peptide and first 8 amino acids was fused to human β-endorphin the resulting hybrid protein was not only exported across the cytoplasmic membrane in *E.coli* but also selectively secreted into the culture medium. Further discussion of prokaryotic protein sorting signals is given in section B.

Components of the protein export machinery in bacteria

From genetic studies several genes have been identified which appear to encode components of the *E.coli* protein export machinery. One of the first to be identified was the secA gene (Oliver & Beckwith,1981) and recently five new genes were identified which were involved in the synthesis of secreted proteins in *E.coli* (Oliver,1985b); so it appears that the list may not yet be complete, neither are the components fully characterised functionally. The experimental evidence concerning these genes involved in protein export is reviewed by Michaelis & Beckwith(1982), Oliver(1985a) and Benson *et al*(1985).

The proteolytic enzymes responsible for cleaving prokaryotic signal peptides have been better characterised than the eukaryotic counterpart. *E.coli* leader peptidase (also called signal peptidase I) appears to be localised equally in the inner and outer membranes; the purified enzyme from either membrane will correctly cleave the N-terminal leader peptide from M13 procoat without requiring other factors, and when reconstituted into liposomes (Zwizinski & Wickner,1980; Zwizinski *et al*,1981; Watts *et al*,1981; Wolfe *et al*,1983). Its specificity in processing appears to be the same as eukaryotic signal peptidase since, as noted earlier, eukaryotic secretory precursors are correctly cleaved in bacteria, and prokaryotic signals are accurately processed by *in vitro* and *in vivo* eukaryotic systems, even in the case of prokaryotic proteins which are not exported by an ER-like mechanism (see Watts *et al*,1983). Thus the same theoretical considerations of a peptidase recognition and cleavage site (Perlman & Halvorson,1983; Von Heijne,1983 and 1984a), described in A.2. apply. Although this enzyme was found to process M13 procoat and several other precursors of periplasmic and outer membrane proteins, it would not cleave prolipoprotein, the precursor of an outer membrane protein (Tokunaga *et al*,1982). There is a second signal processing enzyme, termed signal peptidase II, responsible for processing a group of lipoprotein precursors in which glyceride modification is essential to cleavage (Yamada *et al*,1983). This protein, unlike signal peptidase I, is sensitive to globomycin and resides in the cytoplasmic membrane of *E.coli*, it is encoded by the lspA gene which is now cloned and sequenced (Innis *et al*,1984).

I.B. PROTEIN SORTING SIGNALS

I.B.1. The Theory of Topogenic Sequences

It has been pointed out several times in section A that the signals which direct a protein to its correct subcellular location reside in the protein structure itself, as either a permanent or transient component. The accumulating experimental evidence towards this conclusion resulted, in 1980, in Blobel proposing the theory of topogenic sequences; the concept that a limited number of discrete, positively acting, autonomous regions of the polypeptide chain are responsible for sorting proteins to the range of intracellular compartments and for achieving, where appropriate, their particular orientation in the membrane. Four types of topogenic sequences are predicted - 'signal sequences', 'stop-transfer sequences', 'insertion sequences' and 'sorting sequences', the information in each type being decoded and processed by specific effectors. 'Signal sequences' initiate the translocation of proteins across membranes, in a process mediated by specific protein receptors in the membrane. These must be subdivided into signal sequences involved in the unidirectional translocation of proteins across the different translocationally-competent cellular membranes, *ie.* of the ER, mitochondria, chloroplast and the prokaryotic plasma membrane. 'Stop-transfer' sequences are regions of the polypeptide which halt the translocation process previously initiated by a signal sequence; this results in the asymmetric integration of proteins into translocationally-competent membranes so that part of the protein spans the membrane. In contrast 'insertion sequences' are proposed to effect the unilateral integration of proteins into a lipid bilayer but without involving protein effectors, and these do not result in the protein spanning the membrane. Lastly 'sorting signals' constitute a diverse, 'and the rest' category comprising those determinants which direct proteins after translocation to other membranes and compartments which can not translocate proteins.

Experimental work which tests this hypothesis has been steadily growing. Two complementary experimental criteria can be used to identify such a topogenic sequence as a discrete, positively acting unit; the alteration of a protein's location by the deletion or mutation of the proposed region, and addition of the same domain resulting in redirection of a foreign protein to the appropriate compartment specified by the topogenic sequence. The involvement of specific effectors for decoding and processing the information has been less easy to test experimentally. Initially most work was done with prokaryotes since it was possible to

carry out the appropriate genetics in bacteria (reviewed in Shuman,1981; Emr & Silhavy,1982; Benson *et al*,1985). However more recently the advent of recombinant DNA techniques has enabled eukaryotic systems to be studied (see Colman,1982; Ellis,1985 and particularly the review of Garoff,1985). This section will consider the experimental data and ideas concerning the structure of protein sorting signals. I will concentrate on the signal sequence for translocation across the ER and its equivalent in prokaryotes, and to avoid confusion the term signal sequence will only be used to refer to this translocation signal. I will also summarise the experimental approaches used to study and characterise other topogenic sequences.

I.B.2. Prokaryotic Signal Sequences

Genetic studies in bacteria have provided confirmation that the signal sequence is essential for initiation of protein export and have begun to define the structural features of the signal peptide which are required for its function. Early experimental work involved the selection or construction of gene fusions encoding hybrid proteins with amino-terminal sequences from an exported protein and part of the cytoplasmic protein β-galactosidase at the carboxyl end. Export-defective strains in which the hybrid protein was not transported from the cytoplasm were isolated and analysed, and these were found to have alterations in the signal sequence region. *In vitro* mutagenesis has also been used to construct signal sequence mutants of prokaryotic proteins. An outline of the experimental methods used to isolate gene fusions and export defective mutants and a full discussion of the results obtained can be found in the reviews of Emr *et al*(1980); Michaelis & Beckwith(1982); Silhavy *et al*(1983) and Benson *et al*(1985).

Experiments on the localisation of various hybrid proteins containing parts of exported proteins - such as the outer membrane λ receptor protein (Hall *et al*,1982) and maltose binding protein or alkaline phosphatase which are periplasmic proteins (Bassford & Beckwith,1979; Michaelis & Beckwith,1982) - established that the information necessary to export proteins from the cytoplasm resides at the N-terminal end of exported proteins. Export-defective fusions, recombinants and mutants have been isolated as well as revertants in which export is restored, and these have enabled a more detailed analysis of the information specifying the initiation of protein translocation in bacteria.

The extensive work of Benson, Emr, Silhavy and their colleagues has concentrated on the bacteriophage λ receptor, *lamB*, which is found in the

outer membrane of *E.coli*. Their data showed that in all the export-defective *lamB-lacZ* hybrids they characterised there was a mutation in the signal sequence of the λ receptor, however not all these mutations when recombined onto an otherwise wild-type *lamB* gene resulted in a block of *lamB* export. It seemed that in the 25 amino acid leader peptide the hydrophobic core was important for initiating protein export, since introducing charged residues in this region interfered with export. Furthermore certain residues within the hydrophobic segment were more sensitive to mutation in terms of maintenance of signal sequence function (Emr & Silhavy,1982). An analysis of the structure of export-defective mutants showed that if the predicted α-helical conformation of the hydrophobic region was disrupted, by amino acid subsitution or deletion, export was inhibited. Double mutant pseudorevertants were isolated from a deletion mutant; in these the secondary point mutation extended the portion of the shortened signal peptide which is predicted to form an α-helix (Emr & Silhavy,1983). It is proposed that the central region of the signal sequence must be able to maintain an α-helical configuration to effect translocation, and a critical subset of four amino acids comprises a recognition site that interacts directly with a component of the export machinery.

Similar conclusions were drawn from work with the *malE* gene product, maltose binding protein (MBP) (Bassford & Beckwith,1979; Bedouelle *et al*,1980; Bassford *et al*,1981). Bankaitis *et al*(1984) then isolated intragenic suppressor mutations of an export-defective *malE* mutant with a 7 amino acid deletion in the signal peptide, these restored to varying degrees the export of MBP to the periplasm. Most of these suppressor mutants contained further mutations in the signal sequence which were predicted to physically lengthen the truncated hydrophobic region; either through the insertion of additional hydrophobic residues, or by substituting one of the charged residues at the N-terminal end with an uncharged amino acid, or lastly by amino acid substitution such that the hydrophobic segment could adopt a more extended conformation. However one suppressor mutation, which was least efficient in exporting MBP, contained an amino acid substitution not in the signal peptide but at residue 19 of mature MBP. This result provided further support for the idea that information for initiating protein secretion is contained within the mature protein (see also I.A.5). However it is important to distinguish effects due to mutations which make the mature polypeptide chain no longer permissive for translocation across membranes, from mutations which disrupt the protein sorting signals; for example it may

be found that the signal peptide of MBP when fused to a translocation-permissive cytoplasmic protein, such as α-globin, will direct the export of the hybrid protein, indicating the signal sequence does act as an autonomous unit in initiating protein translocation in prokaryotes.

Inouye and coworkers have focussed their attention on the major outer membrane lipoprotein precursor, termed prolipoprotein; which, as explained in section A.5, requires glyceride modification for translocation and processing. They carried out studies to look at the role of the positive charge on the amino-terminal of the signal peptide, which in wild-type prolipoprotein is +2; and used oligonucleotide directed mutagenesis to generate mutants with a net signal sequence charge of +1, 0, -1, and -2 (Inouye *et al*,1982; Vlasuk *et al*,1983). Their results showed that, although there is not an absolute requirement for a positively charged N-terminal for translocation across the plasma membrane, as the positive charge of the signal sequence decreased in the mutants there was a reduction in synthesis of lipoprotein and a progressive decrease in the rate of assembly of the precursor into the membrane. Those mutants carrying a net negative charge on the signal peptide accumulated as unmodified prolipoprotein in the cytoplasm and showed slow and post-translational translocation. This demonstrated the importance of the N-terminal positive charge on the signal peptide in initiating translocation early in synthesis. The loop model for protein secretion in prokaryotes, put forward earlier by Inouye & Halegoua(1980), proposes that the basic N-terminal region of the signal peptide facilitates interaction with the negatively charged inner surface of the cytoplasmic membrane. Inouye *et al*(1984) have also looked at the effect of mutations in the hydrophobic region of the prolipoprotein leader peptide and these results indicated that there is considerable flexibility in the primary structure of this region in terms of functioning in protein export. The localisation and orientation of lipoprotein variants which contain two signal peptides in tandem array was examined recently by Coleman *et al*(1985). The internalised signal sequence seemed to function either as the usual translocation signal for the downstream polypeptide chain, or as a 'stop-transfer sequence' anchoring the protein in the membrane as a transmembrane protein. The role adopted by the second signal peptide depended upon the separation of the two signal sequences and whether the first leader peptide acted as a co- or post-translational translocation signal.

I.B.3. The Signal Sequence for Translocation across the Endoplasmic Reticulum

It is proposed that the function of the signal sequence of eukaryotic secretory and membrane proteins is to initiate the translocation of these proteins across the ER membrane. The importance of the role played by the signal sequence in the process of translocation is now well established, but precisely how this function is achieved is still not clear and the structural features of the signal sequence essential for its interaction with the export machinery are still being elucidated. As Von Heijne points out in a recent paper concerning the limits of variation of the signal sequence structure "one of the outstanding features of the signal sequences as a group is their extraordinary variability in terms of overall length and amino acid sequence" (Von Heijne,1985). According to the Signal Hypothesis the signal sequence will interact specifically with several proteins, including SRP; such specific receptor protein binding usually involves a ligand of defined and limited structure. Protein translocation according to models which envisage a direct integration of the signal sequence into the membrane lipid bilayer would probably impose less rigorous limitations on the primary and secondary structure of the signal sequence, since maintenance of overall hydrophobicity would be the key feature (Von Heijne & Blomberg,1979; Wickner,1980; Engelman & Steitz,1981).

Role of the signal peptide in insertion into the ER

Several workers have reported that translocation of secretory proteins is inhibited in the presence of synthetically synthesized signal peptides (Majzoub *et al*,1980; Prehn et al,1980; Koren *et al*,1983; Austen *et al*,1984). This effect was seen with an 'unnatural' signal peptide which represents a consensus of naturally occurring signal sequences (Austen *et al*,1984), but not with a hydrophobic hexapeptide (Prehn *et al*,1980) or a peptide, glucagon, of similar size (26 amino acids) to the synthetic signal peptide (Koren *et al*,1983). The observations of Koren *et al*(1983) on the effect of microinjection of a synthetic signal peptide into *Xenopus* oocytes indicated that the signal peptide was not only involved in the translocation process but also in later steps of the secretory pathway. However these results should be interpreted with caution since such hydrophobic synthetic peptides are difficult to solubilise and can act as detergents, furthermore they are often used at unphysiological concentrations.

As described in B.2. there is much evidence from work with prokaryotes, using fusion and mutant proteins, for the importance of the signal peptide in translocation. In the case of eukaryotes two groups have shown that the deletion of the N-terminal signal sequence from influenza haemagglutinin resulted in the accumulation of truncated, unglycosylated haemagglutinin in the cytosol (Gething & Sambrook,1982; Sekikawa & Lai,1983). The recent work of Lingappa *et al*(1984) established unequivocally that the signal peptide of a prokaryotic secretory protein contains all the information for translocation across the ER: they constructed a fusion protein encoding the signal sequence plus first five amino acids of β-lactamase and 141 codons of the cytoplasmic protein α-globin at the C-terminus; this was translocated into pancreatic microsomal vesicles *in vitro* and the signal peptide was processed. A concept of the current signal hypothesis and theory of topogenic sequences is that membrane proteins which span the membrane more than once will contain more than one signal sequence (see A.2); additional signal sequences are envisaged to reinitiate translocation of another domain of the nascent chain following the action of the previous stop-transfer signal. The recent work of Friedlander & Blobel(1985) supports this concept of multiple signal sequences; these workers localised 2 of the 4 theoretical signal sequences required for the integration of bovine opsin which crosses the ER membrane 7 times.

It was noted earlier that the work of Hortin & Boime(1980,1981) indicated that the correct processing of cleaved signal peptides is not required for translocation into the ER. More recently Schauer *et al*(1985) reported the isolation of yeast mutants with mutations in the secretory protein invertase such that the mutant invertase protein still maintained enzymic activity but showed delayed intracellular transport. In one of these mutants the ultimate amino acid of the signal sequence was changed from Ala to Val, this mutation was found to result in defective signal processing causing precursor molecules to translocate into the ER but then remain associated with it; some of the mutant precursors were processed at the downstream adjacent peptide bond and these were secreted but showed a 50 fold slower transport to the Golgi than wild type invertase.

Structure and conformation of signal sequences

The sequences of around 200 eukaryotic and 50 prokaryotic signal peptides are now known, many of which are included in the compilation by Watson(1984). A recent comparison by Von Heijne(1985) of 118 eukaryotic

and 32 prokaryotic cleaved signal peptides reiterated the common theme described in A.2, and reported the following general features. Eukaryotic signal sequences range in length from a lower limit of 15 residues with a mode of 18-20 residues. Three structural regions are recognised, a positively charged N-terminal region, a central hydrophobic region and a more polar C-terminal region, these were termed the n, h, and c-regions respectively. Irrespective of the overall length of the signal peptide the c-region showed a consensus length of 5 or 6 amino acids, putting the start of the h-region at residue -6 or -7 (calling the final residue of the signal peptide -1, and the first amino acid after the cleavage site +1). No strong sequence constraints were found in the h-region beyond an observed enrichment for hydrophobic residues (Phe, Ile, Leu, Met, Val, Trp). The length of the n-region does vary strongly with the overall length of the signal peptide and accounts for half the total length variations, but this does not cause a change in net charge - a mean of +1.7 is maintained. The size of the h-region also varies but no change in amino acid composition was seen with increasing length. These observations formed a picture of the 'minimum' eukaryotic signal peptide structure; a 13 amino acid sequence composed of one positively charged amino acid as a n-region followed by a seven residue h-region, with no more than one Ser, Gly, Thr or Pro, and a five residue c-region which obeys the '(-3,-1) rule' defining the cleavage site (Von Heijne,1984a). In prokaryotic signal peptides the n-, h- and c-regions are all on average one residue longer; making the minimum functional sequence 16 amino acids long.

The maximal limits for a functional signal sequence are harder to define. It is suggested that if the h-region of a signal sequence becomes longer than about 20 residues it may then permanently anchor the protein in the membrane (Von Heijne,1981b; Bos *et al*,1984). The n-region containing charged residues prior to the hydrophobic core does appear to be able to be extended considerably. For example, the uncommonly long 51 amino acid N-terminal cleaved signal peptide of human insulin-like growth factor has at its N-terminus a stretch of 26 amino acids containing 6 charged (4+, 2-) residues (Jansen *et al*,1983). It has also been shown in mutant or fusion secretory proteins that additional amino acids can be added N-terminally to the signal peptide without abolishing translocation. It was found that if the signal peptide of preproinsulin was displaced 18 amino acids from the N-terminus, making a n-region of 21 amino acids, this did not affect translocation or cleavage; but preproinsulin with a 68 amino acid N-terminal extension was not translocated (Talmadge

et al,1981; Kozak,1983a and 1984).The uncleaved signal sequence of ovalbumin, which appears to be located towards but not at the N-terminus (within residues 22-41), also functions in a variant ovalbumin with a 21 amino acid N-terminal extension; however a 50 amino acid N-terminally extended ovalbumin was not secreted but remained cytoplasmic (Krieg *et al*,1984; Tabe *et al*,1984). Interestingly when Perara & Lingappa(1985) inserted the preprolactin gene into the chimpanzee α-globin gene such that the normally N-terminal signal peptide of preprolactin was located downstream of the first 109 amino acids of α-globin; they found the entire fusion protein could be translocated, glycosylated and cleaved, yielding a protein identical to mature prolactin and also globin with the preprolactin signal sequence attatched to its COOH-terminus. These results indicate that the amino acid sequence of such N-terminal extensions to secretory proteins must be permissive for translocation if the signal sequence is still to function, and constraints are imposed on the structure of the n-region. In a different analysis of 134 eukaryotic and 39 prokaryotic proteins Von Heijne(1984b) noted a consistency between the two samples in terms of the net N-terminal charge on the signal peptide and in the distribution of the charged residues. He suggested that the N-formyl group on the initiation methionine residue is not removed from prokaryotic signal sequences, and it is compensation for the absence of a free NH_3^+ group on the first amino acid that lies behind the observation that the balance of the charges on the N-terminal charged residues of prokaryotic signal sequences is most commonly +2 whereas in eukaryotic proteins it is +1. The modal charge carried by the eukaryotic signal peptides is also +2 if the positive charge on the free NH_2 group of the terminal amino acid is taken into consideration, but the range is -1 to +5.

Conformational studies of a synthetic signal peptide region of preproparathyroid hormone were carried out by Rosenblatt *et al*(1980); these showed the synthetic peptide had distinct and different conformations in an aqueous and non polar environment, which correlated with the two highest-probability structures predicted from the amino acid sequence. The same rules devised by Chou & Fasman(1978) to predict the secondary structure of proteins have been used in several instances to assess the likely conformation of wild type and mutant signal sequences (for example Emr & Silhavy,1983); from such predictions it seems that to function in translocation the 8-20 residues of the hydrophobic core region must retain a conformation which is compatible with it spanning the thickness of the membrane (25-30Å). A β-turn in the polypeptide

structure is observed to occur near the signal peptide cleavage site consistent with the proposed peptidase recognition site.

Incorporation of β-hydroxyleucine, a leucine analogue, into preprolactin synthesized in a tumour ascites *in vitro* system abolished translocation (Hortin & Boime,1980). The same analogue used in a wheat germ *in vitro* system resulted in the inhibition of high affinity binding of SRP to polysomes synthesizing secretory protein, and the binding of these polysomes to microsomal membranes (Walter *et al*,1981; Walter & Blobel,1981a). This was interpreted as evidence for the direct interaction of the signal peptide with SRP. Although it is also argued that the observed variation in terms of length and amino acid sequence of the n and h-regions of signal peptides is too great to allow specific interaction with protein receptors; instead a non-specific binding is envisaged of the n-region to the surface of the membrane and the h-region to the membrane interior (Von Heijne,1985). Further experimental data is clearly needed to characterise the structure function relationship of the signal sequence for translocation across the ER.

I.B.4. Sorting Signals for Transport of Proteins into Mitochondria and Chloroplasts

Targetting of proteins to the mitochondria

The transient leader sequence of a few mitochondrial proteins has been determined (for examples see Watson,1984) and the structure of these is quite distinct from signal sequences for translocation across the ER, being 30-70 residues long with basic charged amino acids distributed along their length. These basic residues appear to be essential for function since when arginine residues in ornithine transcarbamylase (OTC) were replaced with an acidic analogue, canavanine, import and processing were inhibited (Horwich *et al*,1985a).

It has been established from a number of investigations that the transient leader peptide alone contains the information for the correct localisation of proteins, at least to the mitochondrial matrix. A series of experiments carried out by Hurt *et al*(1984a,1984b,1985a) have shown that a hybrid protein containing as little as the first 12 of the 25 amino acid leader sequence of a mitochondrial matrix protein (yeast cytochrome c oxidase subunit IV) fused to a cytosolic protein (mouse dihydrofolate reductase - DHFR) was correctly targetted to the mitochondrial matrix *in vitro* and *in vivo*. Although the natural cleavage

site had been deleted the pre-piece was cleaved in a hybrid which contained 22 residues of the N-terminal leader peptide, but when only the first 12 amino acids were present no proteolytic processing occurred. Similar fusion experiments were carried out by Horwich *et al*(1985b) using the leader peptide of another matrix protein (OTC) to target DHFR to mitochondria *in vitro* and *in vivo*; the fusion protein was again abnormally cleaved within the DHFR region in both systems, although a cleavage product consistent with normal processing was also seen *in vivo*.

Extensive deletion and fusion manipulations of the cloned yeast mitochondrial 70kd outer membrane protein found that all the information for targetting and anchoring this protein in the outer membrane was contained in the amino-terminal 41 residues (Hase *et al*,1984). Like all outer membrane proteins the 70kd protein does not have a transient leader sequence, but the N-terminus consists of a stretch of 28 uncharged amino acids flanked on both sides by four basic residues. Certain deletions in this region resulted in a small fraction of the protein being misrouted to the matrix, although most remained cytosolic. The analysis indicated that this N-terminal segment could be divided into two domains; the first 11 residues, which are hydrophilic and basic like cleaved mitochondrial leader peptides, mediate targetting to the mitochondria, whereas the following uncharged region acts as a 'stop-transfer sequence' and anchors the protein in the outer membrane. This hypothesis was born out by further work (Hurt *et al*,1985b) which demonstrated that the first 12 amino acids of the 70kd outer membrane protein directed DHFR to the matrix and also could replace the leader peptide of cytochrome c oxidase subunit IV in targetting this mitochondrial protein to the matrix. It is not clear, however, if integration of proteins into mitochondrial membranes always involves membrane protein effectors (see Hurt *et al*,1985b). The anchoring of a foreign protein in the outer mitochondrial membrane by the putative stop-transfer sequence has yet to be demonstrated.

The transit peptide of chloroplast proteins

The transit peptide of the few nuclear-encoded chloroplast proteins which have been sequenced (see Watson,1984) are, like mitochondrial leader peptides, typically 30-60 residues long and positively charged. A comparison of the available sequences of the transit peptide of ssRuBPC from various species revealed a conserved 9 amino acid region which encompassed one of the processing sites (Mishkind *et al*,1985). Van den Broeck *et al*(1985) first reported the targetting of a foreign protein to

chloroplasts, using the 57 amino acid transit peptide of pea ssRuBPC fused to the coding region of bacterial neomycin phosphotransferase. When this construct was introduced into the genome of plant cells the fusion protein expressed was translocated to the chloroplast stroma and processed in a manner similar to the ssRuBPC precursor. Likewise the fusion protein was taken up and processed by intact chloroplasts *in vitro*. Similar results were obtained by Schreier *et al*(1985) using the transit peptide and first 22 amino acids of pea ssRuBPC fused to the same bacterial gene. The signals involved in sorting of proteins to the different compartments and membranes of the chloroplast have not yet been characterised.

I.B.5. Other Topogenic Sequences

Stop-transfer sequences

Several investigations over the past few years have provided evidence for the existence of 'stop-transfer' sequences as proposed by Blobel(1980); discrete, autonomously acting segments of membrane proteins which serve to anchor the protein in the membrane by stopping translocation of the nascent polypeptide chain. Work has mainly been carried out with proteins which span the membrane once with their amino terminal in the lumen of the ER and the carboxyl terminal exposed in the cytoplasm; and it is envisaged that this disposition is achieved by the action of a signal sequence followed by a stop-transfer sequence. Groups have been able to convert such transmembrane proteins to secretory proteins by deletion of their transmembrane segments, ie. influenza haemagglutinin (Gething & Sambrook,1982), VSV-G (Rose & Bergmann,1982). Yost *et al*(1983) inserted the region encoding the transmembrane domain of the membrane form of IgM heavy chain between the coding regions of β-lactamase and globin in the fusion of Lingappa *et al*(1984), described in B.3, which had been shown to translocate completely *in vitro* into microsomal vesicles. The presence of the IgM domain resulted in the fusion protein being anchored as an integral membrane protein with the N-terminal lactamase region within the lumen and the globin domain on the cytoplasmic face of the vesicles. Although the transmembrane region acted as a stop-transfer sequence after initiation of translocation by the signal sequence, it could not itself initiate translocation. A similar result was demonstrated *in vivo* by Guan & Rose(1984) who constructed a hybrid gene encoding the secretory protein growth hormone fused to the membrane-spanning and cytoplasmic domains of VSV-G. From transfection,

immunofluorescence, cell fractionation, immunoprecipitation and proteolysis studies they found that this hybrid protein was anchored in microsomal membranes and transported to the Golgi, but it was not transported to the cell surface. Using oligonucleotide site-specific mutagenesis Adams & Rose(1985a) created DNAs encoding a series of VSV-G variants in which the transmembrane domain was shortened. In wild-type VSV-G the membrane spanning region consists of 20 uncharged, mainly hydrophobic amino acids, and it was found that this region could be reduced to 14 amino acids without affecting either the transmembrane orientation of VSV-G or its appearance at the cell surface. Surprisingly mutants with only 8 or 12 amino acids also spanned intracellular membranes but the proteins were not transported beyond the Golgi. These results argued against the idea that membrane-spanning regions must consist of an entirely α-helical structure, since this conformation requires at least 20 amino acids to span the lipid bilayer; however it is suggested that parts of the polypeptide chain flanking the hydrophobic region are drawn into the membrane-spanning domain in these mutants. It does appear that the length and structure of the transmembrane domain affects not only its function as a membrane anchor, but also transport to the cell surface. If a charged residue is introduced into this region the domain can still span the membrane but cell surface transport of the protein is blocked (Adams & Rose,1985b).

Similar results and conclusions have also been obtained from studies on the membrane anchoring domain of prokaryotic proteins (Abrahmsén *et al*,1985; Boeke & Model,1982; Davis *et al*,1985).

Insertion sequences

It is proposed that insertion sequences are regions of the polypeptide which adopt a conformation which enables the protein to interact directly with membranes, without involving any protein effectors. This concept is used to explain the membrane association of proteins found at the cytoplasmic face of the ER membrane which do not span the membrane and are translated on free ribosomes, such as NADH-cytochrome b_5 reductase and cytochrome b_5 (Borgese & Gaetani,1980; Rachubinski *et al*,1980). Cytochrome b_5 is known to be integrated post-translationally into membranes independent of SRP (Anderson *et al*,1983), and the protein shows spontaneous binding to natural and synthetic membranes (Bendzko *et al*,1982). The properties of the membrane-binding region of cytochrome

b_5 is different to those of signal sequences, and it is thought the former adopts a compact structure with a highly hydrophobic surface. Transmembrane proteins which have a complex topology in the membrane may contain a combination of signal sequence(s), stop-transfer sequences and insertion sequences. It is also envisaged that insertion sequences are responsible for the integration of certain prokaryotic inner membrane proteins which insert post-translationally and do not span the membrane (see Michaelis & Beckwith,1982). A detailed analysis of the structural requirements for a functional insertion sequence has yet to be made.

Sorting Sequences

Blobel proposed that sorting sequences would be discrete positively acting signals which, following translocation, would direct proteins to their different subcellular locations ie. lysosomes, Golgi complex, plasma membrane. However it has also been suggested that after translocation into the ER membrane there is a nonselective routing of all proteins to the cell surface and only those which are retained 'en route' in the ER and Golgi, or diverted, ie. to the lysosomes, need positive sorting signals. Certain observations have suggested that after the action of the signal sequence no further topogenic information is required for export of a protein; for example the prokaryotic periplasmic protein β-lactamase is secreted when expressed in *Xenopus* oocytes (Wiedman *et al*,1984).

It is not yet resolved if there are positive signals for transport along the constitutive pathway to the cell surface. In the absence of a suitable 'marker' protein which can be tagged with the proposed sorting sequence to test its activity, studies have been limited to looking at the effect of mutations or deletions on the transport of various plasma membrane proteins, for example influenza haemagglutinin (Doyle *et al*,1985). In these cases it is difficult to distinguish whether an observed effect on intracellular transport is due to alteration of a positive plasma membrane targetting signal, or to denaturation of the protein. It was noted earlier, in the section on stop-transfer sequences, that whilst secretory proteins could be anchored in the membrane by a membrane-spanning domain from a transmembrane protein these hybrid proteins were not necessarily transported to the cell surface when expressed in cells. However, in experiments which followed from the work by Guan & Rose(1984) on the growth hormone/VSV-G hybrid which was not transported beyond the Golgi, it was found that this hybrid was transported to the cell surface if it was glycosylated through the introduction of a N-linked glycosylation consensus sequence,

Asn-X-Ser/Thr, into the translocated growth hormone domain (Guan *et al*,1985). This result indicates that for this hybrid protein, as in full-length VSV-G (Machamer *et al*,1985), N-linked glycosylation does provide a positive signal for protein transport to the cell surface. However, as mentioned in I.A.3, different proteins show different requirements of glycosylation for intracellular transport, and the observed effects on transport when the normal glycosylation is blocked could be due to altered physical properties of the mutant protein, ie. in terms of solubility or conformation, rather than specifically due to the inactivation of a signal region. These and other problems have meant that limited progress has been made towards defining, or even proving the existence of, discrete signals which interact with specific protein effectors and target proteins along a constitutive pathway to the cell surface; likewise there is little experimental evidence yet for signals for differential routing in polarized cells, or signals responsible for directing proteins to secretory granules in a regulated pathway of secretion (see Garoff,1985).

Recently there was a report of work with deletion mutants of a rotavirus glycoprotein which normally resides in the ER; this showed that deletion of the second of the two N-terminal hydrophobic domains of this protein resulted in products which were transported from the ER, N-linked glycosylated and secreted (Poruchynsky *et al*,1985). These results suggests that this hydrophobic region contributes to a positive signal for localisation in the ER and acts as a membrane anchor, and this sorting sequence overides constitutive transport to the cell surface.

The specific phosphorylation of mannose residues of the oligosaccharide side chains of lysosomal hydrolases (described in A.3) was recognised early on as providing a potential signal for routing these enzymes to the lysosomes. Several studies have tended to support the idea that the mannose-6-phosphate residues act as a component of a recognition marker for sorting of lysosomal enzymes, although this is not necessarily the only lysosome sorting signal. In patients with I-cell disease it was found that active lysosomal hydrolases were secreted, leading to the idea that they lacked the signal for correct localisation in the lysosomes. In these lysosomal enzymes from I-cell disease patients the characteristic phosphomannosyl residues were absent (Hasilik & Neufeld,1980), suggesting that it was this which caused their miscompartmentation - the absence of a sorting signal resulting in their secretion at the cell surface. In mutant cell lines which have defective mannose-6-phosphate receptors the phosphomannosyl signal is not

recognised and these cells secrete lysosomal precursors which contain functional mannose-6-phosphate groups (Robbins & Myerowitz,1981). Recently a study was carried out on the phosphorylation of lysosomal enzymes *in vitro* in the presence of deglycosylated lysosomal enzymes, proteolytic fragments or denatured forms of hydrolases (Lang *et al*,1984). The results provided evidence for a peptide signal or determinant which is recognised by the phosphorylating enzyme, but this determinant is only active in the native enzyme indicating it may consist of discontinuous stretches of amino acids in the polypeptide chain unlike a 'classical' topogenic sequence.

I.C. THE *XENOPUS* OOCYTE AS A SYSTEM FOR STUDYING PROTEIN SEGREGATION

I.C.1. Transcription and Translation of Foreign DNA and RNA by *Xenopus* Oocytes

The *Xenopus* oocyte is now well established as a system in which to study the fate of DNA, RNA, proteins and organelles which can be introduced by microinjection into these large cells, which are the immature, unfertilised eggs of these South African frogs.

In 1971 Gurdon *et al* first reported that eukaryotic mRNA microinjected into the cytoplasm of *Xenopus* oocytes was translated efficiently, and this led to an expanding use of the amphibian oocyte as a translational assay system. The scope of *Xenopus* oocytes in the study of translation of microinjected mRNA and the fate of the expressed foreign proteins has been reviewed recently in Lane(1983) and Soreq(1985), whilst Colman(1984a) also gives practical details concerning the microinjection and subsequent analysis of *Xenopus* oocytes. A major attraction of the oocyte is that, as a complete cell, it provides a system in which post-translational events can be analysed, such as secondary modification and intracellular sorting of proteins. Since the foreign proteins expressed from injected mRNA can be assembled into their biologically active form, oocyte microinjection can also be employed as a bioassay.

It was found later, by Mertz & Gurdon(1977), that DNA microinjected into the germinal vesicle (nucleus) of *Xenopus* oocytes could be transcribed. Since genes transcribed by all three eukaryotic RNA polymerases were found to be active this extended the potential of the oocyte to a system for studying transcription and RNA processing, as well as providing a coupled transcription-translation assay for cloned genes and constructs derived from them. A full discussion of studies of foreign genes in oocytes is given by Gurdon & Melton(1981), Wickens & Laskey(1981) and Lane(1983); once again Colman(1984b) provides a practical description of the techniques involved. It should be noted, however, that not all eukaryotic promoters are active in *Xenopus* oocytes. For example, the promoters of SV40 (Wickens & Gurdon,1983) and the *Xenopus* histone genes (Old *et al*,1982) do function in oocytes, whilst that of chick ovalbumin is inactive (Wickens *et al*,1980).

In addition to making use of the transcriptional and translational capacities of the oocyte, the work in this thesis exploits the *Xenopus* oocyte as a means of determining the post-translational processing and

segregation of foreign proteins. These aspects of the oocyte system are discussed next.

I.C.2. Compartmentation of Foreign Proteins by *Xenopus* Oocytes

It became clear that *Xenopus* oocytes would not only translate injected mRNA, but that the foreign protein products are sorted to their normal intracellular location. In 1977 Zehavi-Wilner & Lane reported that secretory proteins, encoded by mRNA from *Xenopus* liver or guinea-pig mammary gland, fractionated with membrane vesicles where they were resistant to added proteases; whilst globin translated from injected mRNA was found in the cytosolic fraction of the oocyte. Colman & Morser(1979), using coinjected mRNA species, confirmed that oocytes distinguished cytosolic and secretory proteins; they also showed that the latter were not only sequestered by the ER but ultimately exported from the oocyte. The extensive studies of Lane *et al*(1980) demonstrated that cytosolic, membrane, secretory and nuclear proteins were all compartmented correctly by oocytes injected with mRNA from a wide range of sources - mammals, birds, insects, plants and viruses. The polypeptide sequence encoded by the mRNA appears to be sufficient to achieve export, as other factors from the mRNA donor species are not required (Cutler *et al*,1981; Colman *et al*,1983). Hence the oocyte can act as a surrogate segregation and secretory system.

Several groups have studied the segregation of a variety of proteins in *Xenopus* oocytes and their results are summarised in Colman *et al*(1983), Lane(1983), Soreq(1985) and Colman(1984a). These studies have confirmed that proteins synthesized from injected eukaryotic or viral RNA and DNA are generally correctly localised in oocytes. The only foreign protein found to be incorrectly sorted by oocytes is promellitin which is secreted by the venom gland of the honey-bee yet is not exported by oocytes injected with promellitin mRNA; instead the protein is associated with a vesicle fraction (Lane *et al*,1981). It is interesting that not only eukaryotic proteins can be correctly segregated by oocytes; when an *in vitro* synthesized, capped RNA encoding bacterial pre-β-lactamase was injected into *Xenopus* oocytes the prokaryotic secretory protein was exported from the oocyte (Wiedman *et al*,1984).

The oocyte has been employed as a surrogate system to study the compartmentation of mutant proteins. A point mutation in an immunoglobulin light chain which results in its accumulation within myeloma cells was also found to give a non-secretory phenotype in *Xenopus* oocytes, although the protein did gain access to the ER (Valle

et al,1983). The microinjection into oocytes of mRNA obtained from a patient suffering from genetic α1-antitrypsin deficiency showed that the mutant Z-variant protein, unlike the normal M-variant, was not secreted and remained intracellular (Foreman *et al*,1984). The presence of glycosylation on the mutant Z-protein indicated it was, however, able to translocate into the ER. Recently the oocyte has successfully been used in the identification of a topogenic sequence of a polypeptide. With a view to locating the karyophilic signal of the influenza nucleoprotein Davey *et al*(1985) made a series of constructs from the nucleoprotein cDNA and determined if, following microinjection into oocytes, the expressed protein was able to accumulate in the nucleus. Using this approach these workers were able to define a region of 18 amino acids responsible for targetting this protein to the nucleus.

The experiments described by Mishina *et al*(1985) elegantly and fully exploit the scope of the oocyte; firstly to translate injected RNAs encoding the four subunits of the *Torpedo* acetylcholine receptor, then to carry out appropriate post-translational processing (including disulphide bridge formation and N-glycosylation) which enable the assembly in the surface membrane of a functional acetylcholine receptor containing the correct stoichiometry of subunits. These workers injected oocytes with *in vitro* synthesized RNA encoding wild type and specifically mutated acetylcholine receptor subunits, in order to locate specific regions of the α-subunit which were involved in forming either the transmembrane ionic channel or the acetylcholine binding site.

I.C.3. Post-translational Processing of Foreign Proteins by *Xenopus* Oocytes

The cotranslational removal of the N-terminal signal peptide of secretory proteins is one of the early events in the secretory pathway of the cell, and it has been found that oocytes will faithfully cleave these transient signal sequences from foreign proteins. Processing of the signal peptides of chick lysozyme (Colman *et al*,1981) and mouse immunoglobulin (Valle *et al*,1981) in oocytes injected with mRNA was implied from the observation that on gel electrophoresis the protein secreted by oocytes showed an appropriate increase in mobilty relative to the precursor product of *in vitro* translation. Usually the cleavage of signal sequences has been inferred from such a comparison of the molecular weight of the oocyte and *in vitro* product, but in some cases such comparisons are complicated by post-translational modification to the oocyte product. However, Lane *et al*(1981) demonstrated by N-terminal

sequence analysis that oocytes correctly processed the signal peptide of promellitin. Since the transient signal peptides of all foreign polypeptides appear to be recognised and cleaved by the oocyte machinery it seems that this processing event is neither species nor tissue specific, in agreement with other data (see I.A.2).

The cleavage of 'pro' peptides from translocated proteins does seem to be a tissue specific process as the oocyte fails to remove the 'pro' sequence from proinsulin (Rapoport,1981) and promellitin (Lane *et al*,1981). Whilst the absence of further cleavage does not prevent secretion of proinsulin from the oocyte, promellitin remains intracellular as mentioned previously.

The oocyte will carry out other types of post-translational processing on foreign proteins, but these do not necessarily reflect the precise modification which occurs in the 'parent' cell of the foreign protein. In addition to disulphide bridge formation and secondary hydroxylation, acetylation, phosphorylation and glycosylation, injected oocytes have also been shown to carry out the covalent and non-covalent assembly of protein subunits. Examples of these modifications, and the other processing events discussed, are given in the reviews of Asselberg(1979), Lane(1983), Colman(1984a) and Soreq(1985).

I.D. PROCHYMOSIN

Much of the work described in this thesis concerns the segregation of the secretory protein preprochymosin and constructs derived from this protein. A brief outline is included here of the biochemistry of prochymosin and its gene structure, together with a summary of published work concerning the expression of preprochymosin *in vitro* and in *E.coli* and yeast.

I.D.1. Chymosin and Prochymosin Proteins

Chymosin (EC 3.4.23.4) is the major proteolytic enzyme in the fourth stomach, or abomasum, of unweaned calves where it functions in the partial digestion of casein in the dietary milk. The milk clotting activity of this enzyme has given it considerable importance in the cheese making industry as the main component of calf rennet. In common with other proteolytic enzymes chymosin is secreted as an inactive zymogen, prochymosin. In the acidic environment of the stomach an

autocatalytic process removes the N-terminal 42 amino acid residues of prochymosin yielding enzymatically active chymosin (Pederson & Foltmann,1975). Originally this enzyme was known as rennin and its zymogen as prorennin, however to avoid confusion resulting from the similarity of this name to that of another protein, renin, the nomenclature of chymosin and prochymosin was subsequently adopted.

Chymosin has two aspartate residues which are essential for its activity, and this has led to the enzyme being classed in a family of 'aspartate proteases' which includes pepsin, penecillinopepsin, cathepsin D and renin. Calf prochymosin shows extensive homology (55%) with pig pepsin, and both these gastric proteases share homology with penecillinopepsin (Foltmann *et al*,1977 and 1979), leading to the suggestion that the aspartate proteases evolved from a common ancestral gene. The biochemistry and evolution of prochymosin and chymosin has been reviewed by Foltmann(1981).

Prochymosin and chymosin proteins isolated from calves stomachs display a heterogeneity when analysed by chromatographic and electrophoretic techniques. Foltmann(1970) found that, by DEAE-cellulose chromatography, crystalline chymosin could be resolved into three components which he designated chymosin A, B and C. He was able to isolate zymogen precursors of A and B chymosin but not for chymosin C which he thought to be a mixture including degredation products. Further work by Foltmann *et al*(1977 and 1979) determined the complete amino acid sequence of prochymosin B (365 residues) and calculated the expected molecular weight to be 40,777. The protein has three disulphide bridges between the cysteine residues 89 and 94, 249 and 253, 292 and 325. Processing of the zymogen to chymosin leaves 323 residues with an expected molecular weight of 35,652. Partial sequencing of 156 residues of chymosin A (Foltmann *et al*,1979) detected only one amino acid difference in comparison with chymosin B; residue 286 is aspartate in prochymosin A whereas in prochymosin B it is glycine. Using different isolation and purification procedures Asato & Rand(1971 and 1972) distinguished up to four protein species for both chymosin and prochymosin, two of which displayed the same chromatographic properties as authentic A and B chymosin. However when Donnelly *et al*(1984) analysed the proteins of individual calf stomachs, in each of four cases they detected only two enzymatically active forms of chymosin and their corresponding prochymosin precursors. One of these had the same mobility on urea-polyacrylamide gel electrophoresis as chymosin B but the other corresponded neither to chymosin A nor B.

Several possibilities could account for the observed heterogeneity of chymosin proteins *in vivo*; different species may represent post-translational modifications of the same gene product, alternatively some proteins may be encoded by different genes which could be either allelic or non-allelic. To date only chymosin A and B, as designated by Foltmann, have been characterised by protein sequencing; the observed difference in their amino acid sequence indicates these are products of different genes.

I.D.2. Expression of Prochymosin mRNA and Isolation of Prochymosin cDNA Clones

The commercial importance of chymosin made the enzyme an attractive target for gene cloning, whilst the high levels of chymosin synthesized by the calf abomasum indicated that this tissue would be a rich source of prochymosin mRNA for the purposes of making cDNA.

With this end in mind several groups have extracted poly-A RNA from the mucosal layer of the calf abomasum and studied its translational activity in cell-free systems. The translation of prochymosin mRNA in a wheat germ cell-free system appears to be poor and abberant. Early work by Jones & Nicholson (1979) found that whilst free polyribosomes from the calf gastric mucosa stimulated protein synthesis in a wheat germ system, they were unable to obtain translation from poly-A RNA isolated from this tissue. Uchiyama *et al* (1980) did obtain a translation product from abomasum poly-A RNA using the wheat germ system but this had an apparent molecular weight of only M_r=37,000, which corresponds more closely to the expected molecular weight of chymosin than of prochymosin and it is likely that this was an artefact. The rabbit reticulocyte lysate cell-free translation system has also been used and with greater success (Nishimori *et al*,1981; Harris *et al*,1982; Moir *et al*,1982; Nicholson & Jones,1984; McConnell *et al*, 1984). In general translation of calf abomasum poly-A RNA in this system was found to give a major polypeptide product which was specifically immunoprecipitated by antibodies raised against calf prochymosin. The apparent molecular weight of this polypeptide, from its migration on SDS-polyacrylamide gels, was variously estimated to be between M_r=40,000 and 43,000. However the experimental data presented in this thesis are not in complete agreement with these published results as *in vitro* translation of preprochymosin mRNA was found to give two electrophoretically distinct proteins precipitated with prochymosin antibodies; this is discussed in Chapter IV in the light of the observed heterogeneity of chymosin proteins noted above and the information about the cDNA species described below.

Table I.1:
A comparison of the sequence of prochymosin proteins and prochymosin cDNAs.

	Amino Acid Residue in Prochymosin										
	-2	128	158	202	214	236	257	286	318	320
Prochymosin A - Foltmann *et al*,1979	-	?	?	Asp	Asp	?	Leu	Asp	Ser	Gln
Prochymosin B - Foltmann *et al*,1979	-	Leu	Thr	Asp	Asp	Tyr	Leu	Gly	Ser	Gln
Prochymosin cDNA - Nishimori *et al*,1982a	End TAG	Leu CTA	Tyr TAT	Asn AAT	Asp GAC	Thr ACT	Leu CTG	Asp GAT	Gly GGC	Gln CAG
Prochymosin cDNA - Moir *et al*,1982	Gln CAG	Leu CTG	Tyr TAT	Asn AAT	Asp GAC	Thr ACT	Leu CTG	Asp GAT	Ser AGC	Gln CAG
Prochymosin cDNA - Harris *et al*,1982	Gln CAA	Leu CTA	Tyr TAT	Asn AAT	Asn AAC	Thr ACT	Leu TTG	Gly GGT	Ser AGC	Gln CAA

Three groups have isolated and characterised prochymosin cDNA clones (Nishimori *et al*,1981 and 1982a; Harris *et al*,1982; Moir *et al*,1982), however only those obtained by Harris *et al* and Moir *et al* were full length clones encompassing the entire coding region. The DNA sequence of these full length clones showed that, as with most secretory proteins, the prochymosin cDNA gene encoded a precursor form of prochymosin which has an additional 16 amino acids N-terminal to the known protein sequence of prochymosin; this was thought to be the signal sequence of prochymosin necessary for its translocation across the ER membrane. The prochymosin signal peptide contains a high proportion of hydrophobic amino acid residues and conforms to the structure of the signal sequence of other eukaryotic secretory proteins (see I.A.2, I.B.3 and Watson,1984). By convention this precursor primary translation product of the cDNA is called preprochymosin. The cDNA gene characterised by Nishimori *et al*(1982a) has an incomplete signal sequence and contains only ten codons upstream from the prochymosin coding region, one of these is a termination codon; in addition it lacks a methionine initiation codon prior to the sequence encoding prochymosin. These workers suggest that the termination codon, TAG, may have arisen from an error in the *in vitro* synthesis of the cDNA using reverse transcriptase. Moir *et al*(1982) found this codon, the penultimate of the signal peptide, to be CAG whilst Harris *et al*(1982) sequenced this position as CAA - both are codons for glutamine. Nishimori *et al* suggest, alternatively, that the cDNA clone is derived from the RNA of a pseudogene.

The amino acid sequence encoded by the cDNA clones shows minor discrepancies in comparison with the sequences of prochymosin proteins determined by Foltmann *et al*(1979), these differences are summarised in Table I.1 together with the differences noted between the cDNAs. It should be noted that DNA sequence data presented in this thesis showed an error in the published preprochymosin cDNA sequence of Harris *et al*; this error was notably in the signal sequence region of preprochymosin, which was of particular importance in terms of the work of this thesis (see Chapter VI). Spontaneous deamidation during protein sequencing can account for the misidentification of residue 202 as aspartate instead of the asparagine encoded by the cDNAs at this position. Foltmann found residues 158 and 236 of prochymosin to be threonine and tyrosine respectively, whereas all three cDNAs encode position 158 as tyrosine and 236 as threonine. Apart from the above differences and another aspartate/asparagine discrepancy at residue 214 (discussed below), the prochymosin encoded by the cDNA cloned by Harris' group is the same as

the prochymosin B sequence determined by Foltmann *et al*(1977). This indicates that Harris *et al*(1982) isolated the cDNA for preprochymosin B. Nishimori *et al* found a glycine codon (GGC) for amino acid 318, whilst the cDNA clones of Harris *et al* and Moir *et al* contained a serine codon (AGC) at this position in agreement with the protein sequence data. Except for this difference at residue 318 the prochymosin encoded by the cDNA clones of Moir *et al* and Nishimori *et al* are identical, but they differ at two positions from the product of the cDNA cloned by Harris and his workers. The presence of an aspartate codon at position 286 in the clones of Moir *et al* and Nishimori *et al* corresponds to the known difference between prochymosin A and B. The other difference at residue 214 (see Table 1) may either also be a real difference between the two prochymosin proteins or is due to an error in DNA sequencing by Harris *et al*. The sequence data strongly suggest that the complete cDNA isolated by Moir *et al*(1982) is that of preprochymosin A.

An analysis of bovine genomal DNA using cloned chymosin cDNA as hybridization probes was carried out by Moir *et al*(1982), this indicated there is only a single chymosin locus which has at least two introns, and that preprochymosin A and B are products of different alleles of this gene. If there is only one chymosin locus in the bovine genome then the incomplete clone characterised by Nishimori *et al*(1982a) can not be a pseudogene but may be part of the preprochymosin A cDNA gene (with an error in DNA sequencing accounting for the difference in codon 318). Alternatively it may be the product of a third preprochymosin gene, allelic to both preprochymosin A and B.

I.D.3 Expression of Prochymosin cDNA in Bacteria

A gene coding for a methionyl-prochymosin protein was constructed by Emtage *et al*(1983) using synthetic oligonucleotides and the preprochymosin cDNA cloned by Harris *et al*(1982). This methionyl-prochymosin gene was used to obtain the expression of prochymosin in *E.coli* to avoid any problems resulting from incorrect processing of the preprochymosin signal peptide in the bacteria. Calf prochymosin was expressed in *E.coli* when the methionyl-prochymosin cDNA was inserted in a bacterial expression plasmid containing the *E.coli* tryptophan promoter. After centrifugation of lysed cells prochymosin was found to be pelleted with the cell debris which indicated that the protein accumulates in an insoluble or aggregated form in the bacteria. This characteristic aided the purification of prochymosin produced in *E.coli*, and the partially

purified protein could be processed by acidification to give active chymosin.

A different approach was adopted by Nishimori *et al*(1982b) who replaced the first four codons of prochymosin in the cDNA they had isolated (Nishimori *et al*,1982a) with sequences coding for the N-terminal ten amino acids of β-galactosidase. This fusion gene was placed under the control of the lac UV5 promoter, but gave poor expression. However Nishimori *et al*(1984) were able to increase the yield of prochymosin tenfold in *E.coli* by fusing the N-terminal part of the trp E gene (anthranilate synthetase), preceded by the trp promoter and attenuator, to cDNA which encoded all but the first four codons of prochymosin. The prochymosin fusion product, like the complete prochymosin expressed in *E.coli* by Emtage *et al*(1983), appeared to be localised in the bacterial cell membranes and was not secreted.

I.D.4. Expression of Chymosin Clones in Yeast

The expression of chymosin cDNA genes has also been tested in yeast since it was hoped this would provide a more suitable system from which to purify a protein for use in the food processing industry. Mellor *et al*(1983) made three derivatives of preprochymosin for expression in yeast, using the cDNA cloned by Harris *et al*(1982). Synthetic oligonucleotides were used to construct fragments which coded for methionyl-prochymosin and methionyl-chymosin polypeptides. These fragments and an insert encoding the complete preprochymosin protein were each cloned into a yeast expression vector. Chymosin-specific polypeptides, immunoprecipitated by prochymosin antisera, were expressed from all three inserts. Most immunospecific protein was expressed from the methionyl-prochymosin insert (5% total cell protein), but the methionyl-chymosin was very poorly expressed (<0.1% total cell protein) in comparison with both the methionyl-prochymosin product and that from the preprochymosin insert (1% total cell protein). The observed difference in levels of expression from the three plasmids was not a consequence of plasmid copy number or levels of chymosin-specific RNA. A polypeptide, which comigrated with chymosin (M_r=36,000) was seen in the yeast harbouring the plasmid containing the methionyl-chymosin cDNA. The polypeptide produced from the methionyl-prochymosin insert comigrated with authentic prochymosin with an apparent molecular weight M_r=41,000. The protein synthesized from the preprochymosin insert also showed the same migration as prochymosin on SDS-polyacrylamide gels, instead of having the mobility expected from a precursor containing the 16 amino

acid signal peptide. This comigration of the preprochymosin insert product with prochymosin indicated the signal sequence of preprochymosin had been processed in yeast. Enzymatically active chymosin could be produced by acidification of the products of the preprochymosin and prochymosin inserts.

It was anticipated that the yeast secretory system might recognise the chymosin polypeptides as a secretory proteins. It appeared that the signal sequence of preprochymosin was processed by yeast, implying a compatibility of the bovine secretory precursor with the initial processes of the yeast secretory pathway, including translocation across the ER and cleavage by signal peptidase. However no secretion of chymosin proteins was seen either from yeast spheroplasts or whole cells. Therefore the full length preprochymosin polypeptide does not contain all the information necessary to achieve secretion in yeast. Yet it was observed that none of the chymosin-specific polypeptides, produced from the three different cDNA constructs, was detected intracellularly, but the proteins were all associated with the yeast cell wall. The precise subcellular localization of the chymosin polypeptides expressed in yeast was not clear. The association of the methionyl-chymosin and methionyl-prochymosin with cell membranes led Mellor *et al* to suggest that these may contain a topogenic sequence which directs their subcellular location in yeast. This topogenic sequence must, however, be distinct from the signal sequence which is only contained in the preprochymosin insert.

Goff *et al*(1984) also constructed a methionyl-prochymosin for expression in yeast, they used the cDNA isolated by Moir *et al*(1982). These workers placed this gene under the control of the GAL1 gene promoter which enabled regulation of prochymosin expression by varying the yeast carbon source. The prochymosin purified from yeast could be activated by incubation at pH2 to yield a protein with the same milk-clotting activity and migration on gels as chymosin. In agreement with the work of Mellor *et al*(1983) Goff and his workers found that about 80% of prochymosin was not freely soluble in the yeast cytoplasm, but they did not determine its subcellular location.

I.E. THE WORK OF THIS THESIS

The work of this thesis addresses questions concerning the function and structure of the signal sequence of eukaryotic secretory proteins, using both *in vitro* and the *Xenopus* oocyte *in vivo* assay systems. The experimental work it includes was carried out from July 1982 to March 1985.

The ultimate aim of the project was to use site-specific mutagenesis to modify the cleaved, N-terminal signal peptide of a eukaryotic secretory protein; then to look at the segregation of these mutants *in vitro* and their interaction with components of the translocation machinery, with a view to understanding the molecular basis for mutant nonsegregated phenotypes. Initially the *Xenopus* oocyte would be developed as an *in vivo* assay system to study the segregation of proteins expressed from cDNAs encoding wild-type and mutant secretory proteins; since, as described in section I.C., these oocytes can efficiently translate RNA transcribed from microinjected DNA, and correctly localise the foreign proteins. Ultimately, however, mutant cDNA constructs would be transcribed *in vitro* using the SP6 system (Melton *et al*,1984; Krieg & Melton,1984) which was being developed at the time the project was started but did not become available in our laboratory until mid 1984. The SP6 synthetic RNAs could then be translated *in vitro* in cell-free systems in the presence of dog pancreatic vesicles and specific microsomal extracts, containing for example SRP or SRP receptor. It was hoped to be able to sub-group nonsegregating proteins according to which stage in the translocation process specific mutant signal peptides were defective. For example does SRP recognise the mutant signal sequence and cause an arrest in translation, or can the SRP/ribosome/nascent secretory protein complex interact correctly with the SRP receptor to release the elongation arrest?

Before constructing and characterising signal sequence mutants it was important to establish whether eukaryotic signal sequences function autonomously in their role of acting as the signal for translocation across the ER membrane. Would mutations in the signal sequence cause a nonsegregating phenotype just as a consequence of altering the conformation of the signal peptide region independent of the rest of the protein, or were interactions between the signal sequence and the mature protein also to be considered in the interpretation of experimental data? At the stage this work was initiated it had not been demonstrated that eukaryotic signal sequences function as discrete, autonomous units. Most of the published data related to experiments on the signal sequence

of prokaryotic secretory proteins, but it was not clear to what extent the situation in prokaryotes reflected the process of translocation across the ER membrane in eukaryotes. As discussed earlier (I.A.5 and I.B.2) there was growing evidence that, although similar in some respects, protein secretion in bacteria has only a limited resemblance to the process in eukaryotes. Furthermore it was also not known whether the obligate coupling of translocation to translation seen in eukaryotic *in vitro* systems provided an accurate model for the *in vivo* process. A major part of the work described in this thesis was carried out to look at this question of whether a eukaryotic signal peptide was a self-contained unit which could function when transferred to another eukaryotic protein, both *in vivo* and *in vitro* (see below).

The first chapter of results (III) describes the development of vectors which enable the expression of cDNAs in *Xenopus* oocytes, thus facilitating a study of the segregation of proteins encoded by specific cDNA constructs in this *in vivo* system. Using these vectors with cDNA inserts encoding prelysozyme and preprochymosin the localisation of these two eukaryotic secretory proteins was examined. In the context of the aim of generating signal sequence mutants which would have a nonsegregating phenotype, the expression and localisation in the oocyte of two truncated polypeptides which lacked a signal peptide region was also characterised. These signal-minus cDNA constructs were both derived from the preprochymosin cDNA, one variant encoded methionyl-prochymosin and the other methionyl-chymosin.

In the course of *in vitro* translation experiments it was found unexpectedly that preprochymosin mRNA expressed two electrophoretically distinct proteins, both precipitated by antibodies raised against prochymosin; this was in contrast to published data (see I.D.2). Chapter IV describes studies on the translocation and processing of these polypeptides encoded by preprochymosin mRNA, *in vitro* and also *in vivo* using *Xenopus* oocytes.

Chapter V introduces the SP6 transcription system, which first became available at this stage and was subsequently used to generate synthetic RNAs for translation in the oocyte and in cell-free systems. This chapter covers experiments on the expression *in vivo* and *in vitro* of SP6 RNAs transcribed from preprochymosin cDNA and the methionyl-prochymosin and methionyl-chymosin derivatives (described earlier).

In order to establish whether eukaryotic signal sequences function autonomously, two hybrid genes were constructed in which the signal sequence of prelysozyme was replaced either by the 'pre-' sequence only

of preprochymosin, or by the signal peptide plus part of the mature prochymosin sequence. The results on the translocation of these hybrid proteins in the oocyte and *in vitro* are discussed in Chapter VI; these experiments included the expression of synthetic SP6 RNAs encoding these fusion proteins and prelysozyme, both in oocytes and *in vitro*.

The long term objective of isolating signal sequence mutants was not achieved within the time period of my PhD studentship, but Chapter VII includes a discussion of *in vitro* mutagenesis techniques which were considered with a view to generating specific mutations within the signal peptide of preprochymosin for subsequent study as outlined above.

The same vectors, described earlier, which enable the expression of cDNAs in *Xenopus* oocytes were also used in another series of experiments, which were started in parallel with those outlined in this section. The objective of this second investigation was to identify the region of the polypeptide chain of ovalbumin which functions as the signal sequence for translocation across the ER. Ovalbumin has long attracted attention as the only example of a eukaryotic secretory protein which does not have a cleaved signal sequence (Palmiter *et al*,1978; Lingappa *et al*,1978b). From work by Lingappa *et al*(1979) it was proposed that an internal region, residues 229-276, acted as the signal sequence. Later this was found to be incorrect and it was shown that the signal sequence function lies at the N-terminal, within the first 150 amino acids (Braell & Lodish,1982b). After work was started at Warwick a further report was published by Meek *et al*(1982) which narrowed down the location of the ovalbumin signal sequence to within the first 60 residues. The approach used by Dr. Colman's group, here at Warwick, was firstly to make specific deletions of ovalbumin cDNA and insert these into expression vectors, then inject these constructs into the nuclei *Xenopus* oocytes in order to study the segregation of the mutant ovalbumin proteins *in vivo*. Secondly, once identified, the putative signal sequence region of ovalbumin was to be fused to a nonsegregating polypeptide to determine whether the resulting hybrid protein was translocated *in vivo*. The success of these fusion experiments would depend on the signal sequence region alone being able to effect translocation, without interaction between the signal sequence and other regions of the protein being required. These ovalbumin signal sequence fusions would therefore provide an opportunity of testing the autonomy of a eukaryotic signal sequence which is not cleaved on translocation. This work, to which I contributed, will not be described in this thesis but it is published (Krieg *et al*,1984; Tabe *et al*,1984) and reprints of the papers are included in the Appendix.

II.A. MATERIALS

Source of mRNAs, recombinant DNAs and antibodies

PolyA preprochymosin mRNA was prepared from unweaned calves as described in Harris *et al*(1982), and was a kind gift from T. Harris (Celltech Ltd., UK). Oviduct mRNA from Rhode Island Red hens was prepared in this laboratory by D. Drummond and L. Tabe as described in Cutler *et al*(1981). Human placental lactogen mRNA was a kind gift from Amersham International.

Recombinants pSV2 SFV(d-1) and pSV S-SFV(d-1) (Garoff *et al*,1983; Kondor-Koch *et al*,1983) were kind gifts from H. Garoff (EMBL, Heidelberg, W.Germany). The plasmids pTKMoLTR$_1$, which contains the Moloney murine sarcoma virus long terminal repeat (Dhar *et al*,1980), and pTK1, which contains the *Herpes simplex* thymidine kinase gene (Wilkie *et al*,1980), were both gifts from N. Wilkie (Beatson Institute, Glasgow, UK). The plasmid pls184, which contains all the chicken lysozyme coding sequence (Land *et al*,1981), was a gift from A. Sippel (Cologne, W.Germany). *BclI* DNA fragments containing the coding sequences for calf preprochymosin, prochymosin and chymosin were gifts from T. Harris (Celltech); these are described in Emtage *et al*(1983) and Mellor *et al*(1983) as fragments '82', '70' and '86' respectively.

Rabbit anti-chick lysozyme was a kind gift from D. Cutler (University of Warwick). One batch of rabbit anti-prochymosin sera (αpC_{CT}) was a gift from P. Lowe (Celltech), but more prochymosin antibodies (αpC_{RS}) were raised at Warwick (see II.G)

Chemicals

Generally 'AnalaR' grade chemicals were obtained from BDH Chemicals Ltd., Dorset.

The following compounds were purchased from Sigma Chemical Co. Ltd., Poole, Dorset: Ampicillin, Chloramphenicol, Dithiothreitol (DTT), deoxynucleotide 5' triphosphates (dNTPs), dideoxynucleotide 5' triphosphates (ddNTPs), Ethidium bromide, Phenylmethylsulphonylfluoride (PMSF), Trizma base, 5-bromo-4-chloro-3-indolyl-β-galactoside (BCIG), Isopropyl-β-D-thiogalactoside (IPTG).

Nucleoside 5' triphosphates (ATP, CTP, UTP, GTP) and the capping dinucleotides m^{7}G(5')ppp(5')G and m^{7}G(5')ppp(5')Gm were obtained from Pharmacia, Hounslow, Middlesex. A *HindIII/AluI* oligonucleotide linker [d5'GCAAGCTTGC3'] and nuclease free bovine serum albumin (BSA) were obtained from GIBCO BRL, Uxbridge, Middlesex.

The '15mer' M13 sequencing primer [dTCCCAGTCACACGT] was purchased from Boehringer Mannheim, BCL, Lewes, E.Sussex.

Electrophoresis reagents were purchased as follows: Agarose Type II (Sigma). Acrylamide, specially purified for electrophoresis (Fisons, Loughborough, Leics.). N,N'-methylenebisacrylamide (bisacrylamide) (Eastman Kodak Co., Kirkby, Liverpool). N,N,N',N'-tetramethylethylenediamine (TEMED) (Biorad Laboratories, Watford, Herts.).

Materials for bacterial culture were obtained as follows: Oxoid yeast extract (Oxoid Ltd., Basingstoke, Hants). Bacto-agar, -tryptone, -casamino acids (Difco Laboratories, E.Molesey, Surrey.

Radiochemicals

The following radiochemicals were purchased from Amersham International plc, Amersham, Bucks.

[^{14}C]methylated protein mixture, molecular weights 14,300-200,000; 5μCi/ml.

L-[^{35}S]methionine, in aqueous KCH_3COO solution; ≈1060Ci/mmol (10-15mCi/ml)

Deoxyguanosine 5'-[α-^{32}P]triphosphate, triethylammonium salt in aqueous solution; ≈3000Ci/mmol (10mCi/ml).

Deoxycytidine 5'-[α-^{32}P]triphosphate, triethylammonium salt in aqueous solution; ≈3000Ci/mmol (10mCi/ml).

Adenosine 5'-[γ-^{32}P]triphosphate, triethylammonium salt in aqueous solution; 5,000Ci/mmol (10mCi/ml).

Guanosine 5'-[α-^{32}P]triphosphate, triethylammonium salt in aqueous solution; 410Ci/mmol (10mCi/ml).

Enzymes

DNA restriction enzymes were obtained from Bethseda Research Laboratories (GIBCO BRL). Calf intestinal alkaline phosphatase (CIP), Klenow DNA polymerase, T_4 ligase, T_4 kinase and micrococcal nuclease were obtained from Boehringer Mannheim. SP6 RNA polymerase was from NEN Research Products, through Dupont (UK) Ltd., Stevenage Herts. RNasin (ribonuclease inhibitor) was purchased from P & S Biochemicals, Liverpool. Lysozyme and RNase A were obtained from Sigma.

Photographic Materials

Fuji RX X-ray film (Fujimex Ltd., Swindon). DX80, FX40 (Kodak Chemicals Ltd., Kirkby, Liverpool). Polaroid Land Film type 665 (Polaroid (UK) Ltd., St. Albans, Herts).

Commonly Used Buffers

TE: Tris/EDTA (10mM-Tris.HCl, 1mM-EDTA, adjusted to required pH (between 7.4-8.0))

TEA: Tris/EDTA/Acetate (40mM-Tris.HCl, 20mM-$NaCH_3COO$, 2mM-EDTA, pH8.3)

TBE: Tris/Borate/EDTA (100mM-Tris.HCl, 100mM-boric acid, 2mM-EDTA, pH8.3)

II.B. MANIPULATION OF DNA

Many of the techniques described in this section are included in 'Molecular cloning: a laboratory manual' by Maniatis *et al*(1982) to which the reader is referred for further details and background information. It should be noted that the precise conditions I used here vary, in general only slightly, from those described in Maniatis *et al*.

II.B.1. Restriction Endonuclease Digestion

Restriction enzymes were used according to the suppliers instructions. The DNA to be digested was dissolved in the appropriate buffer to a concentration of ≤100μg/ml and then incubated with the required enzyme as described in Maniatis *et al*(1982). To digest RNA present heat treated RNase A (10μg/ml) was included in the incubation when necessary.

II.B.2. Phosphatase Treatment of DNA

Calf intestinal alkaline phosphatase (CIP) was used to remove the 5' phosphate groups from DNA. When a new batch of CIP was received it was dissolved in sterile water to 1unit/μl and aliquoted into 25unit lots which were lyophilised and stored at -20°C with desiccant. The lyophilised CIP was reconstituted as required in 10mM-Tris.HCl(pH7.6), 10mM-$MgCl_2$ to give a solution of 0.5unit/μl; this was kept at 4°C and used for up to six weeks.

Up to 5μg of DNA in 50μl 100mM-Tris.HCl(pH8.0) or restriction buffer were incubated with 2.5units of reconstituted CIP plus 0.16%ʷ/v SDS. The reaction was incubated at 37°C for ≥2h, after which the enzyme was removed by extracting the mix twice with phenol:chloroform (1:1). Traces of phenol were removed by extracting the final aqueous phase with diethyl ether, and the DNA was recovered by ethanol precipitation.

II.B.3. Filling in 3' Recessed Ends

The 3' recessed ends of restriction fragments were filled in using the Klenow fragment of DNA polymerase I according to the suppliers instructions. The reaction was terminated by extracting it with phenol:chloroform(1:1) and the DNA was recovered by ethanol precipitation.

II.B.4. Ligation of DNA Fragments

DNA fragments with cohesive or blunt ends were ligated together by incubation overnight (approx. 16h) at 4°C with T_4 ligase in 70mM-Tris.HCl(pH7.5), 7mM-$MgCl_2$, 1mM-ATP, 10mM-DTT. Generally 50ng of each DNA fragment was used in a 10μl ligation mix with 1unit of T_4 ligase. After incubation the enzyme was inactivated by heating the reaction to 65°C for 5min.

II.B.5. Addition of Linkers to DNA Fragments

Synthetic oligonucleotide linkers, containing specific restriction endonuclease cleavage sites, were obtained with hydroxyl groups at their 5' ends. To produce 5' phosphate groups suitable for ligation the linkers were treated with T_4 polynucleotide kinase and ATP in a reaction containing 50mM-Tris.HCl(pH7.6), 10mM-$MgCl_2$, 10mM-DTT, 1mM-ATP, 0.5units/µl T_4 kinase, 100ng/µl linkers. The reaction was incubated for 1h at 37°C. When radiolabelled linkers were required [γ-^{32}P]ATP was used instead of unlabelled ATP during the first 30min of the incubation period; 'cold' ATP(1mM) was then added and the reaction continued a further 30min. The efficiency of the phosphorylation reaction was assessed by separating an aliquot of the ^{32}P-labelled linkers on a 10% nondenaturing polyacrylamide gel (see B.9), together with some ligated phosphorylated linkers. Autoradiography of the gel showed the incorporation of ^{32}P-label into a fragment the size of the linker, and into a ladder of fragments with the ligated linkers. The ligation of linkers to DNA fragments was carried out as described in B.4 but using 200ng of phosphorylated linkers and 500ng of the DNA fragment.

II.B.6. Transformation of Plasmid DNA into *E.coli*

The *E.coli* strain MC1061 was used as a transformation host for plasmids. Whilst this strain is recA$^+$ no problems of recombination between plasmid and bacterial DNA were encountered, and the strain was found to give higher transformation efficiencies than recA$^-$ strains.

MC1061 was grown at 37°C by liquid culture in L broth (1%(w/v)-tryptone, 0.5%-yeast extract, 0.5%-NaCl) or on L agar plates (L broth plus 1.5%(w/v) agar).

To prepare competent cells for transformation an overnight culture of MC1061 was diluted 1:50 with L broth then incubated in an orbital shaker (200rpm) at 37°C until the culture reached a density of approx. A_{600}=0.6. After chilling the culture on ice for 30min the bacteria were harvested by centrifugation at 4,000g for 5min at 4°C. The cells were resuspended in ice-cold 0.1M-$MgCl_2$ to half the original culture volume, and the suspension was centrifuged as before. The bacterial pellet was then resuspended to 1/20 the original volume in ice-cold 0.1M-$CaCl_2$ (made fresh from solid $CaCl_2.6H_2O$). Competent cells were left at 4°C for 4 to 24h before being used for transformation.

For transformation plasmid DNA (1-3µl of a ligation mix, see B.4) was added to 0.2ml competent cells and incubated for 30min on ice. The cells were then heat-shocked by transfer to 42°C for 2min, and afterwards

incubated on ice for a further 30min. Following addition of 0.5ml L broth the cells recovered at 37°C for 30min, after which they were mixed with 3ml L top-agar (L broth containing 0.7% agar) kept molten at 45°C, and plated onto L agar containing 100μg/ml ampicillin, in 15cm petri dishes. All plasmids used contained the ampicillin resistance gene of pBR322.

II.B.7. Isolation of Plasmid DNA

Plasmid DNA was prepared from transformed *E.coli* using the alkaline lysis method of Birnboim & Doly(1979). For analysis of recombinants plasmid DNA was isolated from a 1.5ml overnight culture of the transformant in L broth containing ampicillin(30μg/ml). For larger scale preparation a fresh overnight culture was diluted 1:50 with L broth+ampicillin(30μg/ml); when this culture reached A_{600} of 1.0-1.5 chloramphenicol was added to 200μg/ml and incubation continued overnight. An outline is given of the isolation of plasmid DNA from a 1l chloramphenicol treated culture, while the figures inside the square brackets are for the small scale isolation of plasmid DNA from 1.5ml cultures.

E.coli were harvested by centrifugation at 4,000g for 5min at 4°C [5min in a microfuge]. The pellet was resuspended in 8ml [40μl] of 25mM-Tris.HCl(pH8.0), 10mM-Na_2EDTA, 15%(w/v)-sucrose, 2mg/ml-lysozyme, and incubated on ice for 30min [5min]. Two volumes of 0.2M-NaOH, 0.1%-SDS were added and the mixture left on ice for 10min [5min]. After addition of 10ml [50μl] 3M-$NaCH_3COO$(pH4.6) the mixture was incubated on ice for a further 40min [5min] before pelleting the denatured chromosomal DNA by centrifugation at 15,000rpm for 15min at 4°C in a MSE HS18 8×50 rotor [10min in a microfuge]. The supernatant was removed and 2.5 volumes of absolute ethanol added to it. After incubation at -70°C for 15min the precipitated material was recovered by centrifugation at 10,000rpm for 10min at 4°C in a MSE HS18 8×50 rotor [10min in a microfuge]. The pellet was washed with 70% ethanol and then dissolved in 3ml [100μl] water. The large scale preparations were RNase treated at this stage by incubation with 100μg/ml heat treated RNase A at 37°C for 30min. The solution was then extracted twice [once] with phenol:chloroform(1:1) and once with water-saturated diethyl ether prior to carrying out a second ethanol precipitation. The final dried pellet was dissolved in sterile H_2O and stored at -20°C.

Where necessary the plasmid DNA was further purified using sucrose gradients (see below, B.8) or by banding on CsCl gradients. CsCl (208g) was dissolved in 10mM-Tris.HCl, 1mM-Na_2EDTA (160ml) and 25ml of this

solution was mixed with 5ml DNA solution and 1.5ml ethidium bromide(10mg/ml). The samples were then centrifuged in a V.Ti.50 vertical rotor in a Beckmann L8 at 50,000rpm and 23°C overnight (approx. 16h). The banded DNA was collected and the DNA recovered as described in Maniatis *et al*(1982).

II.B.8. Purification of DNA on Sucrose Gradients

Sucrose gradients (10-40%) for fractionation and purification of plasmids and DNA fragments were prepared as follows in polycarbonate tubes to fit a 6×14 swing out rotor of a MSE 65 centrifuge. 5.5ml 10%(ᵂ/v)-sucrose in TE Buffer (pH8.0) was layered on top of 5.5ml 40%(ᵂ/v)-sucrose in TE Buffer (pH8.0), and the tubes were sealed with 'Nesco' film. The tubes were gently turned horizontal and placed at 37°C for 90min to allow the gradient to form, after this they were carefully turned upright again and cooled to 4°C. After balancing the tubes in the centrifuge buckets the DNA samples (≤50µg) were layered onto the gradients. The gradients were then centrifuged in a precooled rotor at 35,000rpm and 4°C overnight (approx. 16h).

The gradients were fractionated using an ISCO Density Gradient Fractionator fitted with a filter absorbing at 254nm. A 'pusher solution' of 50%(ᵂ/v)-sucrose was pumped into the bottom of the tube at a flow rate of 1ml/min, and fractions were collected from the top of the gradient. When uncut plasmid DNA was fractionated under these conditions the supercoiled form was collected in the first DNA peak detected (around 7ml), migrating slower on the gradient than nicked open circles; this is the reverse of the relative mobility of these two forms on agarose gel electrophoresis. Any RNA present is found at the top of the gradient. DNA was recovered from gradient fractions by a short ethanol precipitation (10min at -70°C).

II.B.9. Gel Electrophoresis of DNA

Agarose gels

For analytical and preparative purposes DNA was electrophoresed on horizontal agarose gels. Agarose (TypeII, Sigma) was dissolved in TEA Buffer (see II.A) to give the required concentration of agarose, this ranged from 0.8-3.0%(ᵂ/v). Samples of DNA were loaded into the gel wells in a buffer of 6%(ᵂ/v)-sucrose, 0.1%-bromophenol blue, 1mM-Na_2EDTA. The gels were electrophoresed at a constant voltage with TEA Buffer, using wicks made from 4 layers of Whatman 3MM paper to conduct current from the buffer compartment to the gel. DNA was visualised by immersing the gels after electrophoresis in 5µg/ml-ethidium bromide in TEA for 15-30min, then viewing the stained gel on a UV light box. When required a

photographic record was made of the UV fluorescent DNA fragments using positive-negative Polaroid 665 film.

Polyacrylamide gels

Vertical nondenaturing polyacrylamide slab gels were used to resolve small DNA fragments, <100 base pairs(bp). 10%(w/v)acrylamide gels (30%acrylamide:1%bisacrylamide) were prepared in TBE Buffer (see II.A) using ammonium persulphate and TEMED for polymerization. The gels were pre-electrophoresed at 300V for at least 1h then the samples were loaded and electrophoresis carried out at 600V with TBE as running buffer.

II.B.10. Extraction of DNA From Agarose Gels

Initially in order to isolate a specific DNA fragment use was made of low gelling temperature agarose, but later DNA was routinely extracted from 'normal' agarose gels by electroelution onto filter paper.

Low gelling temperature agarose

Preparative gels were cast using a low gelling temperature agarose (Sea Plaque) and electrophoresed at <40mA. The part of the ethidium stained gel containing the required fragment was cut out and TE Buffer(pH8.0) was added to 1-2× its estimated volume. The gel slice was then melted by heating at 67°C and an equal volume of saturated phenol at 37°C added. Following brief vortexing (30s) the mixture was microfuged for 3min. The aqueous phase was reextracted successively with warm phenol as before until no interface was seen after centrifugation. The final aqueous phase was washed with diethyl ether and the DNA recovered by ethanol precipitation.

Electroelution

Ethidium stained gels were placed on a glass plate on the UV light box and a clean scalpel blade was used to make a slit just below the DNA band to be eluted. A strip of Whatman No.1 filter paper the width of the band backed by a single thickness of dialysis membrane was placed in the slit, with the paper adjacent to the DNA fragment. Electrophoresis was then carried out for a further 10-15min at 150V. After checking under UV that the band had run onto the paper, the paper and dialysis membrane were removed and put into a 0.4ml microfuge tube with a hole pierced in the bottom. This tube was placed inside a 1.5ml microfuge tube and both were microfuged for 20s, spinning the eluted DNA off the paper into the large tube. The paper and membrane were washed twice by adding 100µl TE Buffer(pH8.0) and microfuging as before. The combined eluate was extracted once with phenol/chloroform and once with diethyl ether and the DNA precipitated with ethanol.

II.C. M13 DNA SEQUENCING

Sequencing of DNA was carried out using the chain termination method of Sanger *et al*(1977) and the M13 vectors developed by Messing and his coworkers (Messing *et al*,1977; Gronenborn & Messing,1978; Messing,1983). The techniques involved are now well established and full practical details and background information are given in the texts on DNA sequencing by Davis(1982) and Hindley(1983).

II.C.1. Cloning DNA Fragments into M13 Vectors

The *E.coli* host strain JM103 was used; this was grown at 37°C by liquid culture in 2×YT broth (1.6%(ᵂ/v)-tryptone, 1.0%(ᵂ/v)-yeast extract, 0.5%-NaCl) or on Minimal agar (this contained per litre: 15g agar, 10.5g K_2HPO_4, 4.5g KH_2PO_4, 1g $(NH_4)_2SO_4$, 0.5g sodium citrate dihydrate, 0.2g $MgSO_4.7H_2O$, 5μg thiamine hydrochloride, 2g glucose).

The double stranded replicative form of M13mp10 was prepared as described in Davis(1982) by D. Drummond; this was digested with the appropriate restriction endonucleases and treated with CIP (see B.1 and B.2) to provide vectors for the cloning of specific DNA fragments. Ligation of vector and fragment was carried out as described in B.4.

Competent cells for transformation were made from a fresh culture of JM103 at $A_{600}\approx 0.2$, harvested by centrifugation at 5,000g at 4°C for 5min. The pelleted bacteria were resuspended in freshly made ice-cold 50mM-$CaCl_2$ to ½ the original volume of the culture and left on ice for 20min, then pelleted again and resuspended in 50mM-$CaCl_2$ to $^1/_{10}$ the culture volume. For transformation 300μl competent cells were mixed with 1μl of ligation mix and left on ice for 40min, they were then heat-shocked for 2min at 42°C and returned to ice. The heat-shocked cells were added to 3ml molten H agar (1%-tryptone, 0.5%-NaCl, 0.8% Difco agar) at 45°C containing 25μl BCIG (2% (ᵂ/v) in dimethylformamide), 25μl IPTG (25mg/ml), 20μl of exponentially growing JM103; after mixing this was plated onto Minimal agar plates (15cm diameter). Transformants with M13 containing inserts produced clear ('white') plaques whilst the M13 vector alone gave blue plaques in the bacterial lawn.

II.C.2. M13 Sequencing

Preparation of single stranded template DNA

Single stranded M13 DNA was isolated from the supernatant of infected *E.coli*, it was found that careful preparation of the template was crucial for good results in sequencing. A fresh overnight culture of JM103 was diluted 1:40 with 2×YT and divided into 1ml aliquots. Using a toothpick, virus particles from M13 plaques were transferred to the

II.C. M13 DNA SEQUENCING

Sequencing of DNA was carried out using the chain termination method of Sanger *et al*(1977) and the M13 vectors developed by Messing and his coworkers (Messing *et al*,1977; Gronenborn & Messing,1978; Messing,1983). The techniques involved are now well established and full practical details and background information are given in the texts on DNA sequencing by Davis(1982) and Hindley(1983).

II.C.1. Cloning DNA Fragments into M13 Vectors

The *E.coli* host strain JM103 was used; this was grown at 37°C by liquid culture in 2×YT broth (1.6%(ᵂ/v)-tryptone, 1.0%(ᵂ/v)-yeast extract, 0.5%-NaCl) or on Minimal agar (this contained per litre: 15g agar, 10.5g K_2HPO_4, 4.5g KH_2PO_4, 1g $(NH_4)_2SO_4$, 0.5g sodium citrate dihydrate, 0.2g $MgSO_4.7H_2O$, 5μg thiamine hydrochloride, 2g glucose).

The double stranded replicative form of M13mp10 was prepared as described in Davis(1982) by D. Drummond; this was digested with the appropriate restriction endonucleases and treated with CIP (see B.1 and B.2) to provide vectors for the cloning of specific DNA fragments. Ligation of vector and fragment was carried out as described in B.4.

Competent cells for transformation were made from a fresh culture of JM103 at A_{600}≈0.2, harvested by centrifugation at 5,000g at 4°C for 5min. The pelleted bacteria were resuspended in freshly made ice-cold 50mM-$CaCl_2$ to ½ the original volume of the culture and left on ice for 20min, then pelleted again and resuspended in 50mM-$CaCl_2$ to $^1/_{10}$ the culture volume. For transformation 300μl competent cells were mixed with 1μl of ligation mix and left on ice for 40min, they were then heat-shocked for 2min at 42°C and returned to ice. The heat-shocked cells were added to 3ml molten H agar (1%-tryptone, 0.5%-NaCl, 0.8% Difco agar) at 45°C containing 25μl BCIG (2% (ᵂ/v) in dimethylformamide), 25μl IPTG (25mg/ml), 20μl of exponentially growing JM103; after mixing this was plated onto Minimal agar plates (15cm diameter). Transformants with M13 containing inserts produced clear ('white') plaques whilst the M13 vector alone gave blue plaques in the bacterial lawn.

II.C.2. M13 Sequencing

Preparation of single stranded template DNA

Single stranded M13 DNA was isolated from the supernatant of infected *E.coli*, it was found that careful preparation of the template was crucial for good results in sequencing. A fresh overnight culture of JM103 was diluted 1:40 with 2×YT and divided into 1ml aliquots. Using a toothpick, virus particles from M13 plaques were transferred to the

diluted JM103 aliquots and the infected cultures were incubated for 5h at 37°C in an orbital shaker (200rpm). The M13 single stranded '+ strand' template DNA was prepared as described in Davis(1982); particular care being taken firstly to exclude any bacteria when pipetting off the initial supernatant material, and secondly to avoid taking any of the interface or organic layer with the aqueous phase after phenol extraction.

The sequencing reaction

Sequencing of the M13 single stranded template was carried out using [α-^{32}P]dGTP as the radiolabelled nucleotide and a 15 base sequencing primer (see II.A). The procedure used in general is given below, but exact conditions were varied as necessary; once again practical details can be found in Davis(1982).

Single stranded DNA equivalent to a fifth of a 1ml culture was annealed to 5ng of sequencing primer in a buffer of 10mM-Tris.HCl(pH7.4), 10mM-$MgCl_2$, 50mM-NaCl, 10mM-DTT, in a volume of 10μl in a 0.4ml microfuge tube. The microfuge tube and its contents were placed into a preheated test tube of water standing in a boiling water bath, and left for 5min; the test tube was then removed and, with the microfuge tube containing the annealing mix still inside, it was allowed to cool to room temperature. 5μCi [α-^{32}P]dGTP was added to the annealed primed DNA followed by 1unit Klenow enzyme. 2μl of this mix was then dispensed into four 1.5ml microfuge tubes each containing 2μl of one of the different dideoxynucleotide/deoxynucleotide mixes (ddNTP/N°) listed below.

N° Mixes	volume in μl			
	T°	C°	A°	G°
50mM-Tris.HCl(pH8.0),1mM-Na_2EDTA	5	5	5	5
0.5mM-dTTP	1	20	20	20
0.5mM-dCTP	20	1	20	20
0.5mM-dATP	20	20	1	20
66μM-dGTP	1	1	1	1
ddNTP Concentrations [μM]	ddTTP	ddCTP	ddATP	ddGTP
	500	200	125	200

ddNTP/N° Mixes:

Equal volumes of the relevant pairs of N° and ddNTP solutions were mixed together (ie. ddCTP + C°), and 2μl of the resulting ddNTP/N° mix used in sequencing reactions as described above.

The reactions were incubated for 15min at room temperature then 1μl 0.5mM-dGTP added and a further 15min 'chase' period allowed. At the end

of the chase incubation 5µl Sequencing Load (95%(v/v) deionised formamide, 10mM-Na_2EDTA, 0.1%-xylene cyanol FF, 0.1%-bromophenol blue) was added, except in the case of samples which were not to be electrophoresed immediately; these reactions were terminated by freezing in dry ice, then kept at -20°C and Sequencing Load added when the samples were thawed for electrophoresis.

Polyacrylamide sequencing gels

The products of the sequencing reactions were separated on 6%(w/v)-polyacrylamide (38%(w/v)acrylamide:2%(w/v)bisacrylamide), 8M-urea, TBE gels. These were cast with a thickness of 0.35mm using plates to fit the Raven Scientific model RCA 505 apparatus. Gels were pre-electrophoresed with TBE as buffer at 25mA for at least 1h using a LKB Bromma model 2197 power pack. Just prior to electrophoresis the reactions in Sequencing Load were put in a heating block at 100°C for 5min, then 1-2µl samples were loaded into the sample wells which had been thoroughly flushed out first with buffer. The gels were electrophoresed at 25-30mA with an aluminium plate clamped to the glass plates to evenly distribute the heat. Following electrophoresis the gels were first fixed with 10%(v/v)-glacial acetic acid, 10%(v/v)-methanol and then carefully transferred onto a sheet of damp Whatman 3MM paper. The gel was dried down under vacuum at 60°C and then exposed to X-ray sensitive film at room temperature.

II.D. *IN VITRO* TRANSCRIPTION USING THE SP6 SYSTEM

Synthetic RNAs, for translation *in vitro* (see section E) or *in vivo* following injection into *Xenopus* oocytes (section H), were made using the SP6 system developed by Melton and his coworkers (Melton *et al*,1984; Krieg & Melton,1984). Specific cDNA constructs were cloned into the HindIII site of pSP64 (Melton *et al*,1984). The DNA was linearised to provide a template for transcription by cutting at a unique restriction site in the polylinker region downstream of the insert, and the linearised DNA was purified by electroelution from agarose gels (B.10).

II.D.1. Transcription Reaction

Uncapped transcripts

Transcription was carried out using a modification of the method described in Melton *et al*(1984), taking care to ensure all reagents and equipment were free from nucleases. In brief the linearised DNA template (20-50µg/ml) was incubated for 1h at 40°C in a volume of 50µl in a reaction containing 40mM-Tris.HCl(pH7.5), 6mM-$MgCl_2$, 2mM-spermidine,

10mM-DTT, 2units/µl RNasin, 100µg/ml BSA, 500µM each of the ribonucleotides UTP, CTP, ATP and GTP, 200µCi/ml [α-^{32}P]GTP and 200-300units/ml SP6 RNA polymerase. At the end of the incubation aliquots were taken for quantification of the RNA yield, described below, and the remaining reaction was made up to 100µl and EDTA(pH7.6) added to 100µM. The mix was then extracted once with phenol:chloroform (1:1) and twice with chloroform. Ammonium acetate was added to 0.7M followed by 2.5 volumes of absolute ethanol and the RNA was precipitated overnight at -20°C. The RNA was then dissolved in 0.7M ammonium acetate and precipitated with ethanol a second time as before. The final RNA pellet was washed with 80% ethanol, dried under vacuum then dissolved in autoclaved double distilled water at 100ng/µl and stored at -70°C.

Capped transcripts

To produce capped transcripts the transcription reaction included 500µM-m^{7}G(5')ppp(5')Gm or 500µM-m^{7}G(5')ppp(5')G and only 50µM-GTP.

II.D.2. Estimation of RNA Yields

Agarose gel electrophoresis

A qualitative assessment was made of the amount and size of the transcripts synthesized by separating an aliquot (5µl) of the reaction mix by electrophoresis on a 2% agarose/TEA gel as described in B.9. It was found that the RNA remained as a tight band if electrophoresis was carried out for ≤20min at a constant voltage of 150V.

Incorporation of ^{32}P-labelled ribonucleotide into RNA

The maximum theoretical yield of RNA in an uncapped reaction containing 500µM-GTP is 35µg, whilst in a capped reaction in the presence of 50µM-GTP it is 3.5µg. The amount of RNA synthesized in a reaction was quantitated by determining the percentage of the total [α-^{32}P]GTP which had been incorporated into RNA. This was calculated by determining the radioactivity either in nucleic acids precipitated by trichloroacetic acid (TCA) or material binding to DE81 paper, the latter method being preferred.

For the first method four aliquots (1-2µl) of the reaction were spotted onto Whatman No1 paper (1cm^{2}) and allowed to dry. Duplicate sample were counted directly to give an estimate of the total radioactivity. The other pair were incubated with ice-cold 10%(w/v)-TCA for 10min, washed with more cold 10%-TCA, then ethanol, followed by acetone, and then dried. The radioactivity of both sample squares in Triton toluene scintillant (6ml) was determined using a scintillation counter with the window set for ^{32}P.

In the alternative method duplicate samples were spotted onto DE81 paper (1cm²). One was used directly to determine the total ^{32}P; whilst the other was washed for 5min each with 5 lots of 5%(w/v)-$Na_2HPO_4.12H_2O$, once with H_2O, once with methanol, then allowed to dry. Both samples were counted as above.

II.E. *IN VITRO* TRANSLATION OF RNA

Natural messenger RNA and synthetic SP6 RNAs were translated *in vitro* in rabbit reticulocyte lysate (Pelham & Jackson,1976) and wheat germ (Roberts & Paterson,1973) cell-free systems; practical aspects of both these systems are discussed in Clemens(1984).

II.E.1. Translation of RNA in a Wheat Germ Cell-Free System

The wheat germ extract was prepared by D. Jackson using the procedure of Roberts & Paterson(1973). Translation of exogenous RNAs in the wheat germ extract was carried out using a modification of the procedure described in Clemens(1984). In brief RNA (4-25ng/µl) was incubated for 1h at 27°C with 50µM each amino acid except methionine, 1mM-ATP, 0.1mM-GTP, 10mM-creatine phosphate, 0.7µg/µl creatine phosphokinase, 50µM-spermine, 250µM-spermidine, 80mM-KCH_3COO, 2mM-DTT, 10mM-HEPES (pH7.6 with KOH), 1unit/µl RNasin and approx. 1µCi/µl-[^{35}S]methionine. Prior to incubation an aliquot (1-2µl) of the mixture was removed, and further aliquots were taken at the end of the incubation period to determine the amino acid incorporation (see E.4.). The reaction was terminated by adding EDTA(pH7.5) to 100mM.

II.E.2. *In Vitro* Translation Using a Rabbit Reticulocyte Lysate

Two preparations of rabbit reticulocyte lysate were used for *in vitro* translations: a laboratory preparation made by Dr. G. Valle as described in Clemens(1984), and a commercial nuclease-treated batch (Amersham). Prior to use each 1ml aliquot of the laboratory batch was treated with haemin and micrococcal nuclease (see Clemens,1984). The thawed reticulocyte lysate was adjusted to 50µg/ml creatine phosphokinase and 40µM-haemin, and an aliquot (10µl) removed to determine the endogenous translation before nuclease treatment. The mix was then made to 2mM-$CaCl_2$ and 60units/ml micrococcal nuclease and incubated at 25°C for 6-15min (the optimum time was determined experimentally for each batch of nuclease). The reaction was terminated by the addition of EGTA(pH7.6) to 2.5mM and the lysate put on ice. Samples were taken to determine the endogenous and exogenous translation of the nuclease-treated lysate, and the remainder was divided into 50-100µl aliquots and

flash frozen then stored in liquid N_2. The nuclease treatment should reduce the endogenous translation by about 6%, and the nuclease-treated lysate shows a stimulation in translation of up to 7 fold in the presence of exogenous RNAs.

When the commercial reticulocyte lysate was used translation of RNAs was carried out according to the suppliers instructions, using 6ng/µl RNA, 65µM each amino acid except methionine, 1.5µCi/µl [^{35}S]methionine and 80%(v/v nuclease-treated reticulocyte lysate.

In vitro translations using the laboratory prepared lysate were carried out as described in Clemens(1984). RNA (5-25ng/µl) was incubated for 1h at 30°C in the presence of 50µM each amino acid except methionine, 1mM-ATP, 0.1mM-GTP, 10mM-creatine phosphate, 0.7µg/µl creatine phosphokinase, 80mM-KCH_3COO, 10mM-HEPES (pH7.6 with KOH), approx. 1µCi/µl [^{35}S]methionine and 50%(v/v) haemin and nuclease treated reticulocyte lysate. The reaction was terminated by the addition of EDTA(pH7.5) to 2.5mM and samples were taken to determine the amino acid incorporation (E.4).

II.E.3. Translation and Translocation Assays

Translation in the presence of microsomal membranes

Stripped dog pancreatic microsomal membranes (Walter *et al*,1981) were included in *in vitro* translations to determine whether the products of translation would translocate microsomal membranes. Translations were carried out as described in E.1. and E.2. except that dog pancreatic microsomes (1µl) were included in the assay; these microsomal membranes were a kind gift from G. Blobel.

Protease protection assays

After incubation *in vitro* translation-translocation assays were treated with proteases to determine which translocated products were protected within membranes from protease digestion. The translation mix was divided into 3 portions; to the the first trypsin and chymotrypsin (50µg/ml each) were added, the same proteases plus 1%(v/v) Triton-X-100 were added to the second portion , whilst no additions were made to the third. All three samples were incubated together on ice for 60min, then PMSF was added to 1mM. The protein products in each incubation were then analysed by immunoprecipitation and gel electrophoresis (see II.F).

Separation of membranes and translocated proteins

An alternative method to distinguish those proteins translocated within microsomes involved the separation of membranes from the translation-translocation assay by centrifugation in an airfuge. An

aliquot (10µl) from the assay was mixed with carrier membranes and layered onto a 100µl cushion of 250mM-sucrose, 100mM-KCH_3COO in an airfuge tube (200µl capacity) and centrifuged at 4°C for 20min at 20psi. The supernatant was carefully removed and saved, and the pellet was resuspended in 50µl 'Detergent Buffer' (see II.F.1). PMSF (1mM) was added to both the supernatant and pellet fractions which were then stored at -20°C prior to analysis by immunoprecipitation and gel electrophoresis (as described in section F).

II.E.4. Determination of Amino Acid Incorporation in *In Vitro* Translations

To assess the amino acid incorporation in an *in vitro* translation the radioactivity in proteins precipitated with hot TCA was compared with the total ^{35}S-isotope. The method described in Clemens(1984) was used; in brief duplicate samples (1-2µl) of the incubation mixture were pipetted, using an Oxford sampler, onto Whatman No.1 filter paper squares ($1cm^2$) and allowed to dry. One sample was used for determination of total radioactivity, the other squares were gently swirled successively in ice-cold 5%(w/v) TCA for 15min; 5% TCA plus 3%(w/v) Casamino acids at 90°C for 15min; then at room temperature with 5% TCA for 15min, absolute ethanol for 1min and acetone for 1min. The radioactivity in the dried samples was then determined in a toluene based scintillant using a scintillation counter with a window set for ^{35}S.

II.F. ANALYSIS OF PROTEINS

II.F.1. Immunoprecipitation

The immunoprecipitation of proteins from oocyte fractions (see II.H) or *in vitro* translations (E.1 and E.2) was carried out in a buffer containing a mixture of detergents (Detergent Buffer) (100mM-Tris.HCl(pH7.95), 1%v/v Triton-X-100, 0.5% SDS, 5mM-$MgCl_2$, 100mM-KCl, 0.05% sodium deoxycholate, 1mM-methionine, 1mM-PMSF, final pH 8.2). This was made fresh for each set of immunoprecipitations by mixing stock solutions of the components in the order listed. The whole immunoprecipitation procedure was carried out on ice or at 4°C. The protein sample was mixed with Detergent Buffer (350µl) and an excess of the required antibody, then incubated for 30min. Formalin-fixed Staphylococcal A envelopes (prepared by S. Bahmra and D. Jackson) were prewashed twice with Detergent Buffer and then added in excess to the immunoprecipitation to adsorb the antibody-antigen complexes (Kessler,1975). After incubating at 4°C overnight on a rotastat, the membranes with bound material were pelleted lightly in a microfuge (20s)

and the pellet was washed four times by resuspension in Detergent Buffer followed by pelleting as before. The washed pellet was either prepared immediately for electrophoresis on SDS-polyacrylamide gels or stored at -20°C until required.

II.F.2. SDS-Polyacrylamide Gel Electrophoresis of Proteins

For background information and practical details about the technique of polyacrylamide gel electrophoresis (PAGE) of proteins refer to Hames & Rickwood(1981), only a brief outline of the methods used will be given here.

Preparation of samples for electrophoresis

Immunoprecipitate pellets (see F.1) were resuspended in Sample Buffer (200mM-Tris.HCl(pH8.8), 1M-sucrose, 0.01% bromophenol blue, 5mM-EDTA, 8mM-DTT) and placed in a heating block at 100°C for 5min. After allowing the samples to cool to room temperature they were then alkylated by incubating in the presence of 70mM-iodoacetamide at room temperature for 15-30min. Finally the samples were spun in a microfuge for 2min prior to loading on the gel. When aliquots from *in vitro* translations were to be electrophoresed without immunoprecipitation, these were added directly to Sample Buffer and treated as above; except in the case of *in vitro* translations primed with ^{32}P-labelled 'SP6 RNA' when the sample was incubated with RNase A (300µg/ml, 37°C, 30min) prior to adding Sample Buffer and treating as described above. Portions of treated samples which were not loaded on the gel were stored at -20°C; when a further aliquot of the same sample was required these were thawed, vortexed, spun in a microfuge for 2min and then loaded on a gel.

Gel system

Protein products from *in vitro* translations or immunoprecipitated oocyte fractions were separated by electrophoresis on SDS-polyacrylamide slab gels using a discontinuous buffer system (Laemmli,1970). According to the size of proteins to be resolved the gels were cast with different proportions of acrylamide and bisacrylamide (see below) in a buffer of 375mM-Tris.HCl(pH8.8), 0.1% SDS.

% w/v acrylamide in gel	% acrylamide: % bisacrylamide (cross linking ratio)
15.0	40: 0.2
12.5	30: 0.825
9.0	30: 0.825
9.0	30: 1.6

A single composition of stacking gel was used with 5%(ᵂ/v) acrylamide (30: 0.825 cross linked) in 125mM-Tris.HCl(pH6.8), 0.1% SDS. Ammonium persulphate and TEMED were used as polymerization agents. The gels were electrophoresed with a Tris/glycine running buffer (50mM-Trizma base, 384mM-glycine, 0.1% SDS) at 20mA per gel until the bromophenol blue dye had entered the resolving gel, then electrophoresis was continued at 25-30mA per gel.

Markers

Cytochrome c (25μg, horse heart typeIII) gave a visual indication of the resolution of the gel during electrophoresis. A range of radiolabelled molecular weight markers was provided by running an aliquot (10nCi) of a mixture of methylated ^{14}C-labelled proteins; this contained lysozyme (MW 14,300), carbonic anhydrase (30,000), ovalbumin (46,000), bovine serum albumin (69,000), phosphatase B (92,500) and myosin (200,000).

II.F.3. Fluorography and Autoradiography of SDS-Polyacrylamide Gels

Following electrophoresis the SDS-polyacrylamide gels of ^{35}S-labelled proteins were fixed in 45%(ᵂ/v) methanol, 10%(ᵂ/v) glacial acetic acid and -then fluorographed using the method of Bonner & Laskey(1974) as described in Hames & Rickwood(1981). Occasionally a rapid, single step commercial fluorography agent, 'Amplify' (Amersham) or 'Enhance' (NEN), was used according to the manufacturers instructions. However it was found that these were less sensitive than the longer method of Bonner & Laskey and gave a high background with samples from cell-free translations which were electrophoresed without immunoprecipitation.

Autoradiographs were made of the dried, fluorographed gel by placing the gel in contact with X-ray sensitive film (see Materials) at -70°C. The film was sometimes preflashed to approximately A_{440}=1.0 to increase its sensitivity (Laskey & Mills,1975).

II.G. PREPARATION AND CHARACTERISATION OF PROCHYMOSIN ANTIBODIES

II.G.1. Raising Rabbit Anti-Calf Prochymosin Sera

Antibodies to calf prochymosin for use in immunoprecipitation were raised in a male Lop rabbit (Hylyne, Marston, Cheshire) with the help of Dr. G. Valle who held the required animal licence and performed the injections and bled the rabbit.

Problems were encountered in solubilising the lyophilised prochymosin (a kind gift from P. Lowe, Celltech) for injection; this resulted in modification of the injection protocol to make best use of the material available. The injection routine used is outlined below:

Day	Protein Preparation(*)	Volume	Approx. µg Protein Injected
0	A	500µl	50
14	B	1ml	100
28	B	400µl	50
42	B	400µl	50
50	Bleed 1		
75	C	500µl	150
82	Bleed 2		

(*)Protein Preparations:

A. A solution of ≈100µg prochymosin in 500µl water (obtained as a solution from Celltech) emulsified with 500µl complete Freund's adjuvant to give prochymosin preparation A @ ≈100µg/ml.

B. Lyophilised prochymosin, 500µg, (new batch from Celltech) suspended in 500µl water; as the prochymosin did not dissolve 2ml of Buffer P (10mM-Tris.HCl(pH8.0), 100mM-NaCl, 1mM-Na_2EDTA) was then added, this was recommended by Celltech to aid solution. However, still not all the protein dissolved and the resulting suspension was emulsified with 2.5ml incomplete Freund's adjuvant to give preparation B at ≈100µg/ml prochymosin. This suspension was stored at -20°C between injections and re-emulsified prior to use.

C. 2mg lyophilised prochymosin (another batch from Celltech) dissolved as far as possible in 2ml Buffer P then microfuged to remove undissolved protein. The protein concentration of the supernatant containing solubilised prochymosin was estimated by measuring A_{280} against a set of BSA standards; this gave a concentration of 570µg/ml protein. This solution was emulsified with an equal volume of incomplete Freunds' adjuvant for injection @≈300µg/ml protein.

Bleeds:

1. Blood (40ml) was taken from the major vein in the ear and left to clot for 2h at room temperature then overnight at 4°C. The clotted blood was centrifuged at low speed and the

serum decanted. This serum was badly discoloured by haemolysed cells. A 1ml aliquot was removed for testing and stored at -20°C together with the rest of the serum. This serum is referred to as αpC_{Rs1}.

2. 25ml Blood was taken from the other ear vein and treated as above except that the serum was decanted from the clot prior to clarification by centrifugation; this gave a straw coloured serum (αpC_{Rs2}).

II.G.2. Testing of Antisera by Immunodiffusion Assays and Immunoprecipitation & SDS-PAGE

Immunodiffusion assays

It was hoped to test the cross reaction between the rabbit antiprochymosin sera and calf prochymosin by immunodiffusion 'Ouchterlony' assays, using the following conditions. Glass microscope slides were covered with a layer molten agarose (1.5%^{w}/v Noble agarose, 0.8% NaCl, 0.01% sodium azide, approx. pH8). Wells (2mm diameter) were cut in the set agarose with one central well surrounded by 6 others (at a distance of 7mm), these were filled with the required antigen and antibodies and the slide was left overnight in a humid atmosphere at 37°C. Under these conditions a precipitin arc was seen in a control assay between ovalbumin protein and anti-ovalbumin sera. However, in 5 separate experiments using various preparations of prochymosin and antiprochymosin sera (see below for details) on no occasion was a visible precipitin arc obtained.

Prochymosin used as antigen in central well:

i) Injection solution C (see G.1.), before addition of adjuvant ie 570µg/ml.

ii) A 1:3 dilution of a solution of prochymosin obtained from Celltech (said to be at 1mg/ml in water), ie. diluted to 330µg/ml.

iii) Soluble extract of yeast harbouring a plasmid containing an insert encoding prochymosin (Mellor *et al*, 1983), this contained 1.5mg total protein/ml; it was a kind gift from Dr. A. Kingsman, Oxford University.

Antisera tested in antibody wells (neat and at various dilutions):

a) Rabbit sera raised against calf prochymosin by the group at Celltech and used in our laboratory for immunoprecipitation of prochymosin produced in microinjected oocytes (see III.E) - referred to as αpC_{CT}.

b) Sera from Bleed 1 - αpC_{RS1}, see G.1.

c) Sera from Bleed 2 - αpC_{RS2}, see G.1.

The reason for this failure to precipitate prochymosin in any of the combinations of antigen and antisera is unknown; the Celltech group report they obtain precipitin arcs with αpC_{CT} and prochymosin antigen solution ii), using similar conditions (personal communication from Dr. Peter Lowe). It was therefore decided to determine whether αpC_{RS1} and αpC_{RS2}, like αpC_{CT}, could immunoprecipitate radiolabelled prochymosin from homogenised oocytes following microinjection with cDNA encoding preprochymosin.

Immunoprecipitation and SDS-PAGE of Radiolabelled Prochymosin

Oocytes were microinjected with cDNA encoding preprochymosin ($pTK_2PPChy+$, see III.E.), labelled with ^{35}S-methionine and homogenised as described in II.H. Equal portions were immunoprecipitated (see II.F.1) with the following antisera: 3µl αpC_{CT}, 15µl αpC_{RS1}, 3µl and 15µl αpC_{RS2}. The immunoprecipitated proteins were separated by electrophoresis on a 12% SDS-polyacrylamide gel, as described in II.F.2, together with samples of unlabelled prochymosin used as the antigen in the Ouchterlony assays (i, ii and iii). The part of the gel containing radiolabelled proteins was fixed with acetic acid then fluorographed using 'Amplify' and the dried gel was exposed with preflashed X-ray sensitive film (see F.3); the remainder of the gel was fixed with 12% TCA and stained with Coomasie Blue (as described in Hames & Rickwood,1981), then dried.

On the portion of the gel stained with Coomasie Blue the prochymosin antigen solutions i and ii each gave a single band; a protein which also displayed the same mobility was the major component of the yeast extract (iii) which contained many proteins. The immunoprecipitated samples all gave the same pattern of bands on the autoradiograph of the dried fluorographed gel, with the main band migrating to the same position as the Coomasie Blue stained prochymosin band, this was slightly further than the 46,000 molecular weight marker. So it seemed that the prochymosin antisera I had raised (αpC_{RS1}, αpC_{RS2}) will specifically immunoprecipitate a protein which shows the mobility expected of prochymosin on SDS-PAGE. However, from the relative intensities of the bands on the autoradiograph, neither 15µl of αpC_{RS1} or αpC_{RS2} appeared to precipitate as much ^{35}S-labelled prochymosin as the 3µl aliquot of αpC_{CT}; so it was decided to concentrate the prochymosin antibodies from the sera by salt precipitation.

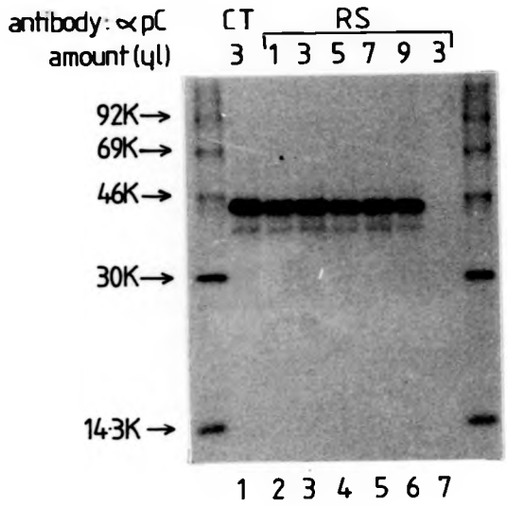


Figure II.1 Immunoprecipitation of prochymosin by prochymosin antibodies, αpC_{Ms}
Preprochymosin mRNA was translated *in vitro* in a reticulocyte lysate cell-free system, as described in II.E.2. Equal portions (8µl) of the *in vitro* translation were immunoprecipitated either with prochymosin antisera raised at Celltech Ltd. (αpC_{CT}, track 1) or with various amounts of salt precipitated prochymosin antibodies raised at Warwick (αpC_{Ms}, tracks 2-6). The immunoprecipitated products were separated by SDS-PAGE on a 12.5% polyacrylamide gel (cross linking ratio 30:0.825) and the radiolabelled proteins were detected by fluorography and autoradiography (see II.F). The dried fluorographed gel was exposed to preflashed X-ray sensitive film for 15h. Track 7 is of material precipitated by αpC_{Ms} from an *in vitro* translation carried out in the absence of exogenously added RNA. The size of the ^{14}C-labelled protein in the marker tracks are indicated.

II.G.3. Sodium Sulphate Precipitation of Prochymosin Antibodies

The immunoglobulin prochymosin antibodies from the rabbit antiprochymosin sera were precipitated with sodium sulphate as described by Heide & Schwick(1978). In brief αpC_{Ms2} was diluted 1:1 with 0.2M sodium phosphate(pH8.0), then this solution was precipitated once with 18%(w/v)-Na_2SO_4, twice with 15%(w/v)-Na_2SO_4, dialysed against phosphate buffered saline and concentrated in an Amicon B15 to 1/6 the original volume of αpC_{Ms2}. The purified antibody fraction was divided into 15µl aliquots which were frozen on dry ice and stored at -70°C. When required an aliquot was thawed and adjusted to 50%(v/v) glycerol; such a glycerol stock, which will be referred to as αpC_{Ms}, was then stored at -20°C and used as a working stock.

II.G.4. Specificity of Prochymosin Antibodies

Preprochymosin mRNA was translated *in vitro* (as described earlier, II.E.2.) in order to provide material to determine the specificity and activity of the antibody fraction (αpC_{Ms}) purified from the rabbit antiprochymosin sera, in comparison with the sera obtained from Celltech (αpC_{CT}). It was found that 1µl αpC_{Ms} was as efficient as a 3µl aliquot of αpC_{CT} (the amount recommended by Celltech for use in immunoprecipitations) in immunoprecipitating the products in an 8µl aliquot of a reticulocyte lysate translation primed with preprochymosin mRNA (12.5ng/µl) (Fig. II.1, tracks 1 & 2). The autoradiograph of the precipitated proteins separated by SDS-PAGE on a 12.5% polyacrylamide gel (cross linking ratio 30:0.825) also showed that, in addition to a major band at M_r=43,000, presumed to be preprochymosin, a minor component with an apparent molecular weight M_r=38,500 was also precipitated both by αpC_{Ms} and αpC_{CT} (see also Chapter IV). This is in agreement with the work of Moir *et al*(1982) who note that on translation of preprochymosin mRNA in a reticulocyte lysate cell-free system a minor species is immunoprecipitated by rabbit antiprochymosin sera, which is seen on SDS-PAGE of the products. They note that this species migrated slightly faster than the chymosin sample on the same gel (expected M_r=36,000) and it can be competed out specifically by excess unlabelled chymosin. It is likely that the minor species (estimated M_r=38,500) seen in Figure II.1 is the same as that observed by Moir *et al*, it is probable the apparent discrepancy in molecular weight is due to different electrophoresis conditions, but Moir *et al* do not give the polyacrylamide composition of their gels. Moir *et al* suggest that this minor species may either be a degredation product of preprochymosin, or the result of translation

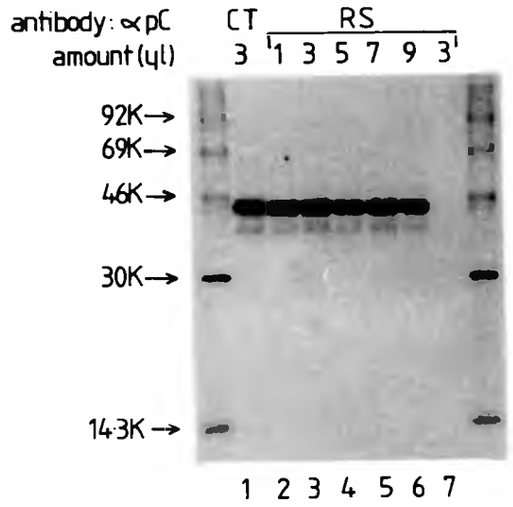


Figure II.1 Immunoprecipitation of prochymosin by prochymosin antibodies, αpC_{RS}
Preprochymosin mRNA was translated *in vitro* in a reticulocyte lysate cell-free system, as described in II.E.2. Equal portions (8µl) of the *in vitro* translation were immunoprecipitated either with prochymosin antisera raised at Celltech Ltd. (αpC_{CT}, track 1) or with various amounts of salt precipitated prochymosin antibodies raised at Warwick (αpC_{RS}, tracks 2-6). The immunoprecipitated products were separated by SDS-PAGE on a 12.5% polyacrylamide gel (cross linking ratio 30:0.825) and the radiolabelled proteins were detected by fluorography and autoradiography (see II.F). The dried fluorographed gel was exposed to preflashed X-ray sensitive film for 15h. Track 7 is of material precipitated by αpC_{RS} from an *in vitro* translation carried out in the absence of exogenously added RNA. The size of the ^{14}C-labelled protein in the marker tracks are indicated.

II.G.3. Sodium Sulphate Precipitation of Prochymosin Antibodies

The immunoglobulin prochymosin antibodies from the rabbit antiprochymosin sera were precipitated with sodium sulphate as described by Heide & Schwick(1978). In brief αpC_{RS2} was diluted 1:1 with 0.2M sodium phosphate(pH8.0), then this solution was precipitated once with 18%(ᵂ/v)-Na_2SO_4, twice with 15%(ᵂ/v)-Na_2SO_4, dialysed against phosphate buffered saline and concentrated in an Amicon B15 to $^1/_8$ the original volume of αpC_{RS2}. The purified antibody fraction was divided into 15µl aliquots which were frozen on dry ice and stored at -70°C. When required an aliquot was thawed and adjusted to 50%(ᵛ/v) glycerol; such a glycerol stock, which will be referred to as αpC_{RS}, was then stored at -20°C and used as a working stock.

II.G.4. Specificity of Prochymosin Antibodies

Preprochymosin mRNA was translated *in vitro* (as described earlier, II.E.2.) in order to provide material to determine the specificity and activity of the antibody fraction (αpC_{RS}) purified from the rabbit antiprochymosin sera, in comparison with the sera obtained from Celltech (αpC_{CT}). It was found that 1µl αpC_{RS} was as efficient as a 3µl aliquot of αpC_{CT} (the amount recommended by Celltech for use in immunoprecipitations) in immunoprecipitating the products in an 8µl aliquot of a reticulocyte lysate translation primed with preprochymosin mRNA (12.5ng/µl) (Fig. II.1, tracks 1 & 2). The autoradiograph of the precipitated proteins separated by SDS-PAGE on a 12.5% polyacrylamide gel (cross linking ratio 30:0.825) also showed that, in addition to a major band at M_r=43,000, presumed to be preprochymosin, a minor component with an apparent molecular weight M_r=38,500 was also precipitated both by αpC_{RS} and αpC_{CT} (see also Chapter IV). This is in agreement with the work of Moir *et al*(1982) who note that on translation of preprochymosin mRNA in a reticulocyte lysate cell-free system a minor species is immunoprecipitated by rabbit antiprochymosin sera, which is seen on SDS-PAGE of the products. They note that this species migrated slightly faster than the chymosin sample on the same gel (expected M_r=36,000) and it can be competed out specifically by excess unlabelled chymosin. It is likely that the minor species (estimated M_r=38,500) seen in Figure II.1 is the same as that observed by Moir *et al*, it is probable the apparent discrepancy in molecular weight is due to different electrophoresis conditions, but Moir *et al* do not give the polyacrylamide composition of their gels. Moir *et al* suggest that this minor species may either be a degredation product of preprochymosin, or the result of translation

starting after the normal ATG initiation codon, thus producing a smaller protein. From the sequence of the cloned preprochymosin cDNAs the next in phase ATG codon, after the initiation codon, corresponds to residue 140 (Harris *et al*,1982; Moir *et al*,1982), so the expected molecular weight of a polypeptide initiating translation from this codon is approx. 28,000. It is therefore unlikely that abberant initiation of translation is the correct explanation for the smaller protein precipitated by the prochymosin antisera, it is suggested instead that it is either a product of premature termination or a degredation product. It should be noted that this minor species was not always seen in immunoprecipitates of *in vitro* translation of preprochymosin mRNA or RNA from cloned chymosin genes (see Fig IV.2) and was never seen in immunoprecipitated oocyte fractions after microinjection of preprochymosin mRNA or cloned chymosin genes (Figs IV.1, III.6, V.3).

II.H. MICROINJECTION AND ANALYSIS OF *XENOPUS* OOCYTES

Natural mRNAs and synthetic SP6 RNAs were translated *in vivo* following microinjection into the cytoplasm of *Xenopus laevis* oocytes. *Xenopus* oocytes were also used as a coupled transcription-translation assay for cDNA constructs contained in eukaryotic expression vectors which were injected into the germinal vesicle. The Introduction section I.C. discusses the *Xenopus* oocyte as an assay system, and information concerning the practical details of the technique is given in the texts by Colman(1984a and b).

The ovarectomy of the frogs and all microinjections were kindly carried out by Dr. A. Colman. The technical assistance of Ms. E. Wallis is also acknowledged here, she frequently prepared the oocytes for injection and carried out the fractionation of injected oocytes and the subsequent immunoprecipitation and SDS-PAGE.

Microinjection of oocytes

Xenopus laevis frogs were maintained under the conditions given in Colman(1984a) by Mrs. C. Kwasinik. Oocytes for microinjection were obtained from anaesthetised animals by partial ovarectomy, then prepared and injected using the procedures described in Colman(1984a and b) throughout. The excised oocytes were firstly thoroughly washed in modified Barths' saline (MBS) (88mM-NaCl, 1mM-KCl, 2.4mM-$NaHCO_3$, 15mM-HEPES (pH7.6 with NaOH), 0.3mM-$CaNO_3.4H_2O$, 0.41mM-$CaCl_2.6H_2O$, 0.82mM-$MgSO_4.7H_2O$, 10µg/ml sodium penecillin, 10µg/ml sodium streptomycin sulphate); the oocytes were then manually separated under MBS using

watchmakers forceps and kept at 20°C, generally overnight, before microinjection.

In the cytoplasmic injections approx. 40nl of RNA solution (0.25-1mg/ml in distilled water) was injected into the vegetal region of each oocyte. Supercoiled plasmid DNA for nuclear injections was purified from sucrose gradients (B.8.) and dissolved at 150µg/ml in 88mM-NaCl, 10mM-HEPES (pH7.6) or in distilled H_2O. Approximately 40nl of DNA solution was introduced into the germinal vesicle of oocytes using the 'blind' injection method (Colman,1984b). In general 20-30 oocytes were injected with each type of RNA or DNA in an experiment.

Before incubating the injected oocytes with radiolabelled methionine, oocytes from both RNA and DNA injections were left overnight in MBS at 20°C, and any unhealthy oocytes were then discarded. The healthy oocytes were incubated at 20°C in MBS containing 1mCi/ml [^{35}S]methionine in the wells of microtitre plates, with 30µl radiolabelled medium per batch of 5 oocytes. After incubation for approx. 24h the medium from wells in which all oocytes were still healthy was removed and saved (to analyse for secreted proteins). Any unhealthy oocytes were discarded and the remaining oocytes were washed thoroughly in MBS, and then fractionated as outlined below.

Oocyte fractionation

Labelled oocytes were homogenized at 4°C, usually as groups of 15-25, in 500µl of T Buffer (50mM-NaCl, 10mM-$Mg(CH_3COO)_2$, 20mM-Tris.HCl (pH7.6)) supplemented with 10%(w/v) sucrose, 100mM-NaCl and 1mM-PMSF. Homogenates were layered onto 1ml of T Buffer containing 20%(w/v) sucrose, in 5ml polycarbonate tubes and centrifuged in a MSE HS18 in a 8×5ml swinging bucket rotor at 17,000g for 30min at 4°C. The supernatants, representing the oocyte cytosol, were removed and retained for immunoprecipitation (F.1.). The pellets, containing yolk and membranes were then extracted with 500µl of Resuspension Buffer (100mM-Tris.HCl (pH7.6), 5mM-$Mg(CH_3COO)_2$, 1%(v/v) Triton-X-100 and 1mM-PMSF) followed by centrifugation for 1min in a microfuge. The supernatant, containing solubilised membranes, was kept for immunoprecipitation (F.1.). Those fractions representing the cytosol and membranes of the oocyte are referred to as the C and M fractions respectively, whilst the medium containing secreted proteins is designated the S fraction.

III. EXPRESSION AND SEGREGATION OF LYSOZYME AND CHYMOSIN BY OOCYTES AFTER INJECTION OF cDNA

The work described in this chapter was carried out in collaboration with Dr. Paul Krieg and Dr. Alan Colman, and it is published as part of Krieg *et al*(1984) which is included in the Appendix.

III.A. Introduction

The objective of this project was to introduce site-specific mutations into the signal sequence region of cDNAs encoding eukaryotic secretory proteins and study their effect on translocation across the ER membrane (see I.E.). The published material to date concerning the structure and function of signal sequences in the translocation of secretory proteins had largely used prokaryotic systems. However it was decided in this project to use eukaryotic expression systems for studying the segregation of such mutants since (as discussed in the Introduction) there was growing evidence that, although similar in some respects, protein secretion in bacteria has only a limited resemblance to the process in eukaryotes.

The oocytes of *Xenopus laevis* was chosen as an *in vivo* expression system for this project, continuing work that had been carried out in this laboratory using these oocytes as a surrogate secretory system (eg. Colman & Morser,1979; Cutler *et al*,1981; Valle *et al*,1981). As noted in the Introduction (I.C.) it is known that oocytes will correctly post-translationally modify many foreign proteins expressed from mRNA and localise the proteins according to their behaviour in their 'parental' cell type, including the secretion of secretory proteins. Other advantages of this system are the ease of the well tested techniques for analysis of the compartmentation of foreign proteins, with a few oocytes providing sufficient material for analysis. These features made the oocyte system preferable to other eukaryotic *in vivo* assay systems such as the transient expression of cDNA following microinjection or transfection of cells in tissue culture. Whilst it was established that *Xenopus* oocytes will transcribe exogenous DNAs microinjected into the germinal vesicle, it had been found that many eukaryotic promoters were not functional in oocytes (ie Wickens *et al*,1980, and see I.C.). In order to study the expression of a variety of cDNA constructs it would be necessary to place the genes under the control of promoters known to function in oocytes. The development of vectors for the expression of cDNAs in oocytes is described in this chapter. The expression of specific cDNA constructs

contained in such vectors, following injection into the oocyte nucleus, would determine whether proteins with signal sequence mutations were translocated and secreted *in vivo*. It was hoped to be able to relate alterations in the structure of the signal peptide to its role in translocation by looking closer at the reason for a nonsegregating phenotype, to determine at which stage in the process of translocation the mutant signal sequence was defective in function. The *Xenopus* oocyte system, however, could not be used for such a biochemical dissection of translocation and *in vitro* assay systems would be needed for this part of the project.

When this work was started in 1982 eukaryotic *in vitro* expression systems had been used to study the process of translocation across the ER membrane. These studies involved the translation of polyA RNA for secretory proteins in cell-free systems supplemented with microsomal membranes and extracts of membranes containing components of the translocation machinery ie. SRP and SRP Receptor (see Introduction section A.2). It was known also that Melton and his laboratory were developing a eukaryotic *in vitro* transcription system based on vectors containing a bacteriophage SP6 promoter from which transcription could be efficiently and specifically initiated by SP6 RNA polymerase (personal communication from D. Melton to A. Colman). It was therefore hoped to use the SP6 transcription system, when it became available, to generate RNAs encoding wild-type and mutant secretory proteins from specific cDNA constructs. Translation of these synthetic RNAs in the established cell-free translation systems could then be used to characterise signal sequence mutants in terms of which stages in the translocation process the signal peptide could function correctly.

III.B. Construction of Vectors for the Expression of Eukaryotic cDNAs in *Xenopus* Oocytes

As noted above not all eukaryotic promoters function in oocytes, so to provide a general system for expression of cDNAs in *Xenopus* oocytes it was decided to use vectors which firstly placed a cDNA insert under the control of a promoter which functions efficiently in oocytes, and secondly could replicate in *E.coli* to facilitate the isolation of DNA for experiments. Vectors based on the pBR/SV40 shuttle vectors constructed by Mulligan & Berg(1980) had been developed by Garoff's group and used to study the expression of cDNA encoding viral membrane proteins on microinjection into cultured cells (Garoff *et al*,1983; Kondor-Koch *et al*,1983). These placed the cDNA insert under the control of the Simian

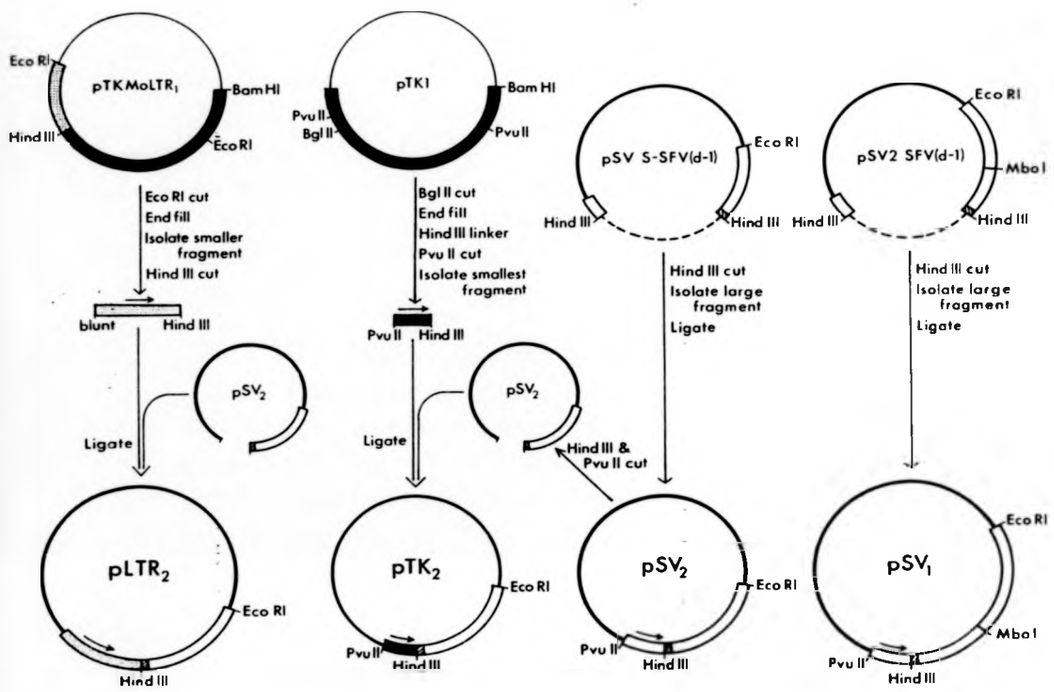


Figure III.1: Construction of the expression vectors pSV_1, pSV_2, pTK_2 and $pLTR_2$.
The plasmids pSV_2 and pSV_1 were derived from pSV S-SFV(d-1) and pSV2 SFV(d-1) respectively, by excision of the Semliki Forest Virus sequence (broken line) using the strategy indicated. These plasmids resembled the parental pSV2 described by Mulligan & Berg(1980) in that they contain 2300 bases of pBR (thick lines) which specify the origin for plasmid replication and the ampicillinase gene. They also both contain SV40 sequences (open boxes) including a *PvuII-HindIII* fragment containing the SV40 early promoter region. Downstream from the unique *HindIII* site, relative to this promoter, both vectors have a fragment carrying the early transcripts termination and polyadenylation signals. In pSV_2, this region lies almost adjacent to the *HindIII* site but in pSV_1 there is, between the *HindIII* site and the polyadenylation region, an additional *MboI* fragment containing the SV40 small t intron. The hatched box represents a 21 base synthetic oligonucleotide that contains translational stop codons in all three reading frames (see Garoff *et al*,1983). The vectors pTK_2 and $pLTR_2$ were were derived from pSV_2 by replacing the SV40 early promoter with either a region of the *Herpes simplex* thymidine kinase gene (shaded box) or the Moloney murine sarcoma virus long terminal repeat (stippled box); manipulations were carried out as indicated in the Figure using techniques described in the Materials & Methods (section II.B). For clarity the various regions of the vectors are not drawn to scale. The small arrows indicate the position and polarity of the promoter in the vectors.

Virus 40 (SV40) early promoter. It was decided to test whether cDNA encoding eukaryotic secretory proteins when cloned into these vectors could be expressed in oocytes. Garoff and his coworkers had constructed two recombinant vectors with inserts containing Semliki Forest Virus sequences; these were designated pSV2 SFV(d-1) and pSV S-SFV(d-1) (Garoff *et al*,1983, Kondor-Koch *et al*,1983), and their composition is given in Figure III.1. These vectors differed in the inclusion in one (pSV2 SFV(d-1)) of an intron and associated flanking sequences in the region encoding the 3' untranslated region of the transcripts initiating at the SV40 early promoter. It had been reported from early experiments on the expression of cloned genes in cultured cells that the presence of an intron was essential to the stability of the transcript, although the exact position of this intron is unimportant (Hamer & Leder,1979). As further genes were expressed in cultured cells this was not found to be a general observation and for several genes the presence of an intron was unecessary (eg. Gething & Sambrook,1982) or led to lower or abberant expression at the protein level (Kondor-Koch *et al*,1983). It was not clear whether there was a requirement for transcript processing for the expression of cloned genes in oocytes, so the Semliki Forest Virus sequences were excised from both the plasmids received from Garoff to generate pSV_1 and pSV_2 for testing as expression vectors in oocytes (see Fig.III.1).

Early experiments in this laboratory showed that the presence of an intron and/or its flanking regions reduced protein expression in oocytes. These results are described in Krieg *et al*(1984), see Appendix. cDNA encoding the secretory protein chick ovalbumin was cloned into the *HindIII* site of pSV_1 and pSV_2 and these recombinants were microinjected into the nucleus of *Xenopus* oocytes. The chick ovalbumin promoter itself does not function in oocytes (Wickens *et al*,1980) but expression of ovalbumin was seen after injection of both pSV_1 and pSV_2 containing ovalbumin cDNA; yet pSV_2 elicited the production of more protein than pSV_1. Therefore vectors which did not contain this small t intron region from SV40 were subsequently employed.

To compare the efficiency of expression from the SV40 early promoter with that of two other viral promoters, the region encoding the SV40 promoter was excised from pSV_2 and replaced by a fragment containing the Herpes Simplex Virus thymidine kinase (TK) promoter or the long terminal repeat (LTR) of Moloney Murine Sarcoma Virus, giving the vectors pTK_2 and $pLTR_2$ respectively (see Fig.III.1). The Herpes Simplex Virus TK promoter had been reported to function well in oocytes

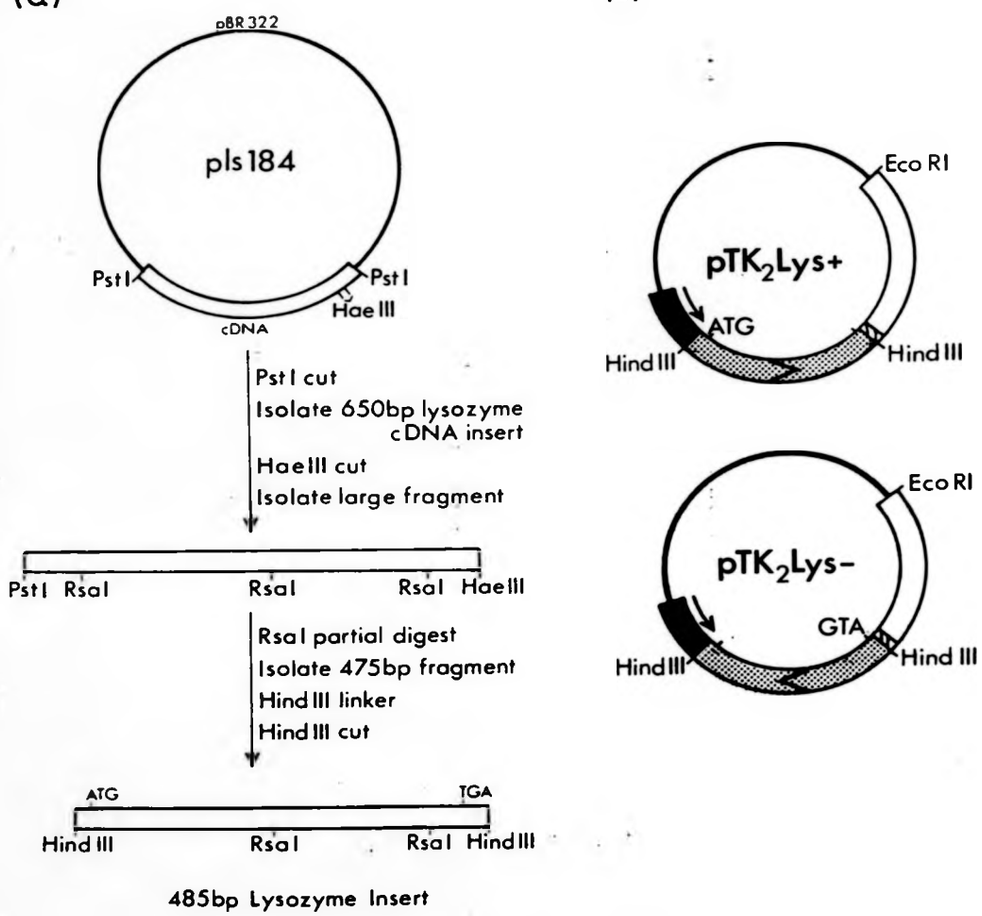


Figure III.2 Insertion of a lysozyme cDNA fragment into the expression vector pTK₂.
(a) Manipulation of lysozyme cDNA for insertion into pTK₂. Chick lysozyme cDNA was obtained from the pBR322 clone pls184, by excision of the *PstI* insert carrying the lysozyme sequence. To remove the ends of the insert bearing the G:C tails, a 475bp *RsaI-HaeIII* fragment was isolated as shown; this contained the entire coding region of lysozyme. *HindIII* linkers were then added to this fragment, which enabled *HindIII* ends to be generated for insertion into the vector pTK₂.
(b) The constructs obtained when this *HindIII* lysozyme fragment was ligated into the *HindIII*-cut vector pTK₂ (described in Fig.III.1). The stippled area represents the lysozyme cDNA insert shown in (a).

(McKnight & Kingsbury,1982) and the LTR promoter was known to be very active in mammalian cells (reviewed by Weiss *et al*,1984).

III.C. Expression of Lysozyme cDNA in *Xenopus* Oocytes

The vectors pSV₂, pTK₂ and pLTR₂, described above, were used to study the expression in oocytes of cDNA encoding lysozyme. This secretory protein with a cleaved N-terminal signal peptide (Palmiter *et al*,1977) was a suitable candidate for study in the context of the aims of this thesis. It was known from work carried out previously in this laboratory that lysozyme expressed from mRNA was processed and secreted well by oocytes (Cutler,1982). Antibodies to lysozyme had been raised for use in immunoprecipitation (Cutler,1982) and sequenced cDNA clones were also available (Jung *et al*,1980; Land *et al*,1981). A fragment containing the entire coding region of lysozyme cDNA (Fig.III.2a) was cloned into the three expression vectors, and those constructs in which the insert was in the 'plus' orientation relative to the promoter were selected (see Fig.III.2b); these were designated pSV₂Lys+, pTK₂Lys+ and pLTR₂Lys+ according to the vector used.

The constructs pSV₂Lys+, pTK₂Lys+ and pLTR₂Lys+ were microinjected into the nuclei of *Xenopus* oocytes as supercoiled plasmid DNA since circular DNA provides the best template for injection (Gurdon & Melton,1981). The injected oocytes were then cultured in the presence of [^{35}S]methionine and, after the medium had been collected (S fraction), the oocytes were homogenised and fractionated into a membrane fraction containing microsomes and other vesicles (M) and a soluble, cytosolic fraction (C), as described in the Methods (section H.). Samples of the S, M, and C fractions were then immunoprecipitated with rabbit anti-chick lysozyme sera and the precipitated radiolabelled proteins were analysed by SDS-PAGE (see II.F).

The injection of pTK₂Lys+ into oocytes resulted in good expression of lysozyme which was mostly secreted into the medium after the 24h incubation with [^{35}S]methionine (Fig.III.3). This result was consistent with the previous data from mRNA injection (Colman *et al*,1981; Cutler *et al*,1981; Cutler,1982), as was the unusual distribution of lysozyme between the C and M fractions within the oocyte, with apparently approximately equal amounts fractionating with the cytosol and vesicles. Generally on fractionation of oocytes expressing foreign secretory proteins, the protein found within the oocyte is in its signal-processed form and most is associated with the vesicle fraction; only a small amount of cleaved protein is found in the cytosolic fraction and this is

Figure III.3 Expression of lysozyme cDNA in oocytes
Oocytes were injected with pTK_2Lys+, $pLTR_2Lys+$ and pSV_2Lys+, and cultured for 24h in [^{35}S]methionine as described in the Materials & Methods. Oocytes were then fractionated into membrane (M) and cytosol (C) components. Samples of the incubation media (S) from 5 oocytes and fractions M and C (representing 2.5 oocytes) were then immunoprecipitated with anti-lysozyme antibody and analysed by SDS-PAGE on a gel containing 12.5% polyacrylamide (30% acrylamide:0.825% bisacrylamide) (see II.F), the dried gel was exposed to preflashed X-ray film. 'lys' indicates the [^{14}C]lysozyme present in the ^{14}C-labelled marker proteins. The arrow marks a band of similar mobility present only in the DNA injected oocytes. The additional bands seen in the S tracks are probably due to microbial contamination of the media.

thought to be due to leakage of translocated proteins from vesicles disrupted during homogenisation. The radiolabelled protein sequestered in microsomes can be subsequently chased out of the oocyte by newly synthesized unlabelled protein (Colman & Morser,1979). Investigations by Cutler(1982) showed that in the case of lysozyme, expressed from injected mRNA, a significant proportion (approx. 15%) of the lysozyme synthesized and processed by the oocyte was not available for secretion, but slowly entered a novel compartment of the oocyte which fractionated at high density on sucrose gradients.

When one compares the expression in oocytes of lysozyme produced under the control of the TK promoter, the retroviral LTR and the SV40 early promoter (see Fig.III.3), it is clear that the use of the LTR and TK promoters enhances the expression of lysozyme relative to the SV40 early promoter which was present in the original vectors; the levels of expression achieved with pTK_2Lys+ and $pLTR_2Lys+$ are roughly comparable. Once again as much lysozyme is seen in the C fraction as in the M fraction of the oocyte and also the bulk of the lysozyme is secreted. The additional bands seen in the S tracks in Fig.III.3 are probably due to microbial contamination of the media as they were also present in the S fraction of uninjected oocytes; they were not seen on other occasions when lysozyme cDNA was expressed in *Xenopus* oocytes (ie see Fig.III.7). The signal peptide appears to have been processed from the lysozyme immunoprecipitated in the M, S and C oocyte fractions as judged by comigration of these protein bands with the lysozyme in the molecular weight markers.

These results and those with ovalbumin (see Krieg *et al*,1984 in Appendix) show these recombinant vectors, particularly pTK_2 and $pLTR_2$, are suitable for expressing cDNA encoding secretory proteins in oocytes, giving a distribution of the foreign protein analogous to that seen following injection of the mRNA.

III.D. Secretion by Oocytes of Prochymosin Translated from Injected mRNA

As a result of collaborative work carried out between Dr. Alan Colman and a group at Celltech Ltd. we obtained in this laboratory mRNA, sequenced cDNA clones, and antibodies for another eukaryotic protein preprochymosin; which is presumed to be a secretory protein as it functions extracellularly as a digestive enzyme in the fourth stomach, or abomasum, of unweaned calves (see Introduction D). Furthermore, sequencing of cDNA clones (Harris *et al*,1982; Noir *et al*.1982) showed that the mRNA encodes a precursor with a N-terminal extension relative to the zymogen

Figure III.3 Expression of lysozyme cDNA in oocytes
Oocytes were injected with pTK$_2$Lys+, pLTR$_2$Lys+ and pSV$_2$Lys+, and cultured for 24h in [^{35}S]methionine as described in the Materials & Methods. Oocytes were then fractionated into membrane (M) and cytosol (C) components. Samples of the incubation media (S) from 5 oocytes and fractions M and C (representing 2.5 oocytes) were then immunoprecipitated with anti-lysozyme antibody and analysed by SDS-PAGE on a gel containing 12.5% polyacrylamide (30% acrylamide:0.825% bisacrylamide) (see II.F), the dried gel was exposed to preflashed X-ray film. 'lys' indicates the [^{14}C]lysozyme present in the ^{14}C-labelled marker proteins. The arrow marks a band of similar mobility present only in the DNA injected oocytes. The additional bands seen in the S tracks are probably due to microbial contamination of the media.

thought to be due to leakage of translocated proteins from vesicles disrupted during homogenisation. The radiolabelled protein sequestered in microsomes can be subsequently chased out of the oocyte by newly synthesized unlabelled protein (Colman & Morser,1979). Investigations by Cutler(1982) showed that in the case of lysozyme, expressed from injected mRNA, a significant proportion (approx. 15%) of the lysozyme synthesized and processed by the oocyte was not available for secretion, but slowly entered a novel compartment of the oocyte which fractionated at high density on sucrose gradients.

When one compares the expression in oocytes of lysozyme produced under the control of the TK promoter, the retroviral LTR and the SV40 early promoter (see Fig.III.3), it is clear that the use of the LTR and TK promoters enhances the expression of lysozyme relative to the SV40 early promoter which was present in the original vectors; the levels of expression achieved with pTK$_2$Lys+ and pLTR$_2$Lys+ are roughly comparable. Once again as much lysozyme is seen in the C fraction as in the M fraction of the oocyte and also the bulk of the lysozyme is secreted. The additional bands seen in the S tracks in Fig.III.3 are probably due to microbial contamination of the media as they were also present in the S fraction of uninjected oocytes; they were not seen on other occasions when lysozyme cDNA was expressed in *Xenopus* oocytes (ie see Fig.III.7). The signal peptide appears to have been processed from the lysozyme immunoprecipitated in the M, S and C oocyte fractions as judged by comigration of these protein bands with the lysozyme in the molecular weight markers.

These results and those with ovalbumin (see Krieg *et al*,1984 in Appendix) show these recombinant vectors, particularly pTK$_2$ and pLTR$_2$ are suitable for expressing cDNA encoding secretory proteins in oocytes, giving a distribution of the foreign protein analogous to that seen following injection of the mRNA.

III.D. Secretion by Oocytes of Prochymosin Translated from Injected mRNA

As a result of collaborative work carried out between Dr. Alan Colman and a group at Celltech Ltd. we obtained in this laboratory mRNA, sequenced cDNA clones, and antibodies for another eukaryotic protein preprochymosin; which is presumed to be a secretory protein as it functions extracellularly as a digestive enzyme in the fourth stomach, or abomasum, of unweaned calves (see Introduction D). Furthermore, sequencing of cDNA clones (Harris *et al*,1982; Moir *et al*.1982) showed that the mRNA encodes a precursor with a N-terminal extension relative to the zymogen

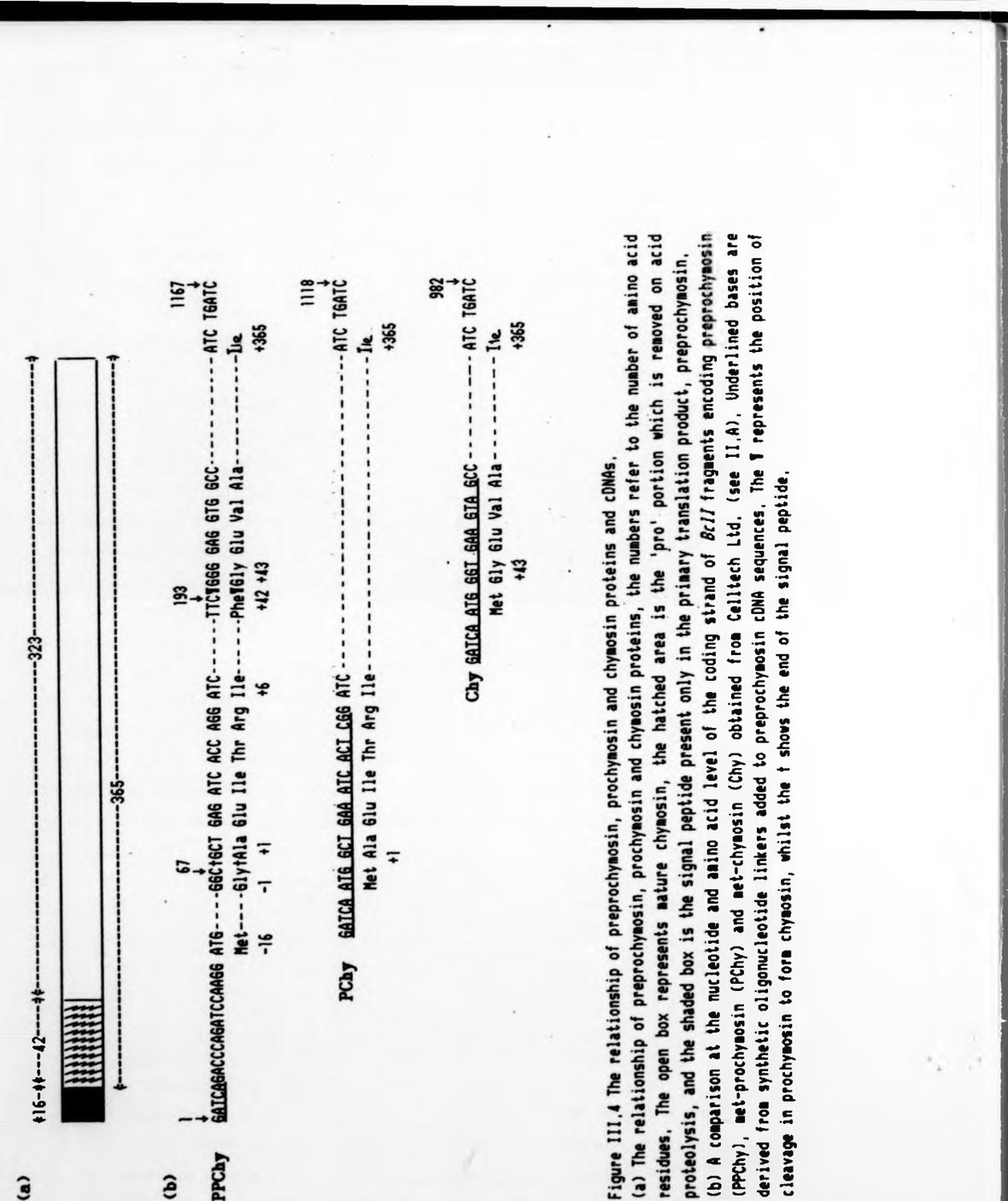


Figure III.4 The relationship of preprochymosin, prochymosin and chymosin proteins and cDNAs.
(a) The relationship of preprochymosin, prochymosin and chymosin proteins, the numbers refer to the number of amino acid residues. The open box represents mature chymosin, the hatched area is the 'pro' portion which is removed on acid proteolysis, and the shaded box is the signal peptide present only in the primary translation product, preprochymosin.
(b) A comparison at the nucleotide and amino acid level of the coding strand of *BclI* fragments encoding preprochymosin (PPChy), met-prochymosin (PChy) and met-chymosin (Chy) obtained from Celltech Ltd. (see II.A). Underlined bases are derived from synthetic oligonucleotide linkers added to preprochymosin cDNA sequences. The ▼ represents the position of cleavage in prochymosin to form chymosin, whilst the ↑ shows the end of the signal peptide.

prochymosin; this 16 amino acid extension has a structure similar to other signal peptides and was presumed to be the signal sequence for translocation.

Experiments carried out by Alan Colman at Warwick demonstrated that injection into oocytes of polyA RNA extracted from the mucosal layer of the abomasum of calves resulted in the expression and secretion of a protein with the characteristics of prochymosin (see Krieg *et al*,1984 in Appendix). The preprochymosin mRNA expressed a protein which was precipitated by antibodies to calf prochymosin and showed the same mobility on SDS-PAGE as authentic prochymosin. In common with observations on the expression in oocytes of proinsulin (Rapoport,1981) and promelittin (Lane *et al*,1981), the prochymosin zymogen is not processed to its active form, chymosin, by the oocyte. This is not surprising since this cleavage is an autocatalytic process occurring at pH<4.0, and is induced in the acidic environment of the calf's abomasum. The prochymosin secreted by oocytes was partially cleaved on acidification of the medium which had surrounded the injected oocytes, giving a polypeptide immunoprecipitated by prochymosin antisera which migrated as authentic chymosin on SDS-PAGE. As this work showed that preprochymosin is secreted by oocytes it was decided to test the expression of this secretory protein from cDNA inserted into the vector pTK_2, to compare with lysozyme; since the availability of mRNA, cDNA and antibodies provided material necessary for the objective of generating signal sequence mutants for study *in vivo* and *in vitro*.

Further studies on the expression of preprochymosin mRNA in oocytes and *in vitro* are presented in Chapter IV.

III.E. Oocyte Expression of cDNAs Encoding Preprochymosin, Prochymosin and Chymosin

III.E.1 Cloning of Chymosin cDNAs into the Expression Vector pTK_2

The Celltech group had isolated and sequenced a full-length cDNA clone of preprochymosin (Harris *et al*,1982). The sequence data indicated that this cDNA represented the preprochymosin B gene (see Introduction I.D). Two derivatives of the preprochymosin cDNA had been constructed by Celltech and their collaborators for use in expression studies in bacteria and yeast (Emtage *et al*,1983; Mellor *et al*,1983; described in sections I.D.3 & 4); these encoded methionyl-prochymosin and methionyl-chymosin polypeptides, and their structure is shown in Figure III.4 compared with the full-length preprochymosin cDNA. The expression in oocytes of the

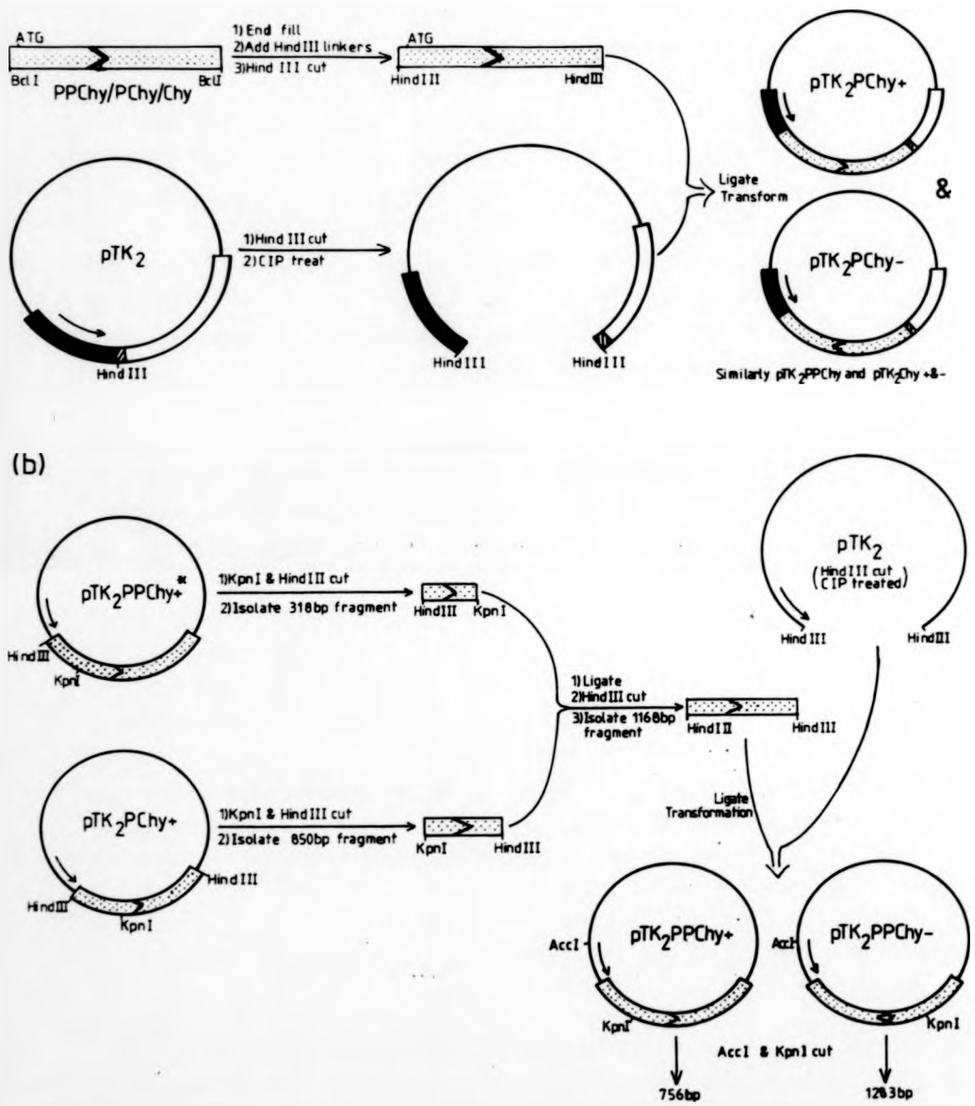


Figure III.5 Manipulation of chymosin cDNAs and insertion into pTK_2
(a) After filling in the recessed ends of *BclI* fragments (described in Fig.III.4b) encoding methionyl-chymosin (Chy), methionyl-prochymosin (PChy) or preprochymosin (PPChy), shown as stippled boxes, *HindIII* linkers were then added to each blunt ended fragment. These linkered fragments were digested with *HindIII* and ligated into the phosphatase treated *HindIII* cut vector pTK_2, using methods described in II.B.
(b) The recombinant $pTK_2PPChy+$, encoding preprochymosin, was reconstructed from pTK_2PChy+ and the mutant $pTK_2PPChy+^*$ using the strategy indicated in the figure and techniques described in the Materials & Methods (section II.B).

preprochymosin cDNA cloned into pTK_2 would provide information on the behaviour of a second secretory protein with a cleaved signal in this cDNA expression system; and preprochymosin could prove more suitable than lysozyme for the proposed signal sequence mutagenesis studies. In this context it was of interest to also characterise the expression and localisation in oocytes of the two other chymosin cDNA constructs which lacked the signal sequence region. In the absence of the signal peptide prochymosin and chymosin are expected to exhibit a nonsegregating phenotype and remain cytosolic, and this is the anticipated localisation of mutant preprochymosin molecules with a functionally defective signal sequence.

All three chymosin cDNA variants were obtained from Celltech as *BclI* restriction fragments (see Fig.III.4) and these were cloned into the *HindIII* site of pTK_2 using the strategy outlined in Figure III.5a and techniques described in the Materials & Methods. Plasmids containing each of the different chymosin cDNA constructs in the 'plus' orientation relative to the TK promoter (see Fig.III.2b) were selected; these were designated $pTK_2PPChy+$, pTK_2PChy+ and pTK_2Chy+ and encoded preprochymosin, prochymosin and chymosin respectively. The initial cloning of the 'PPChy' insert, encoding preprochymosin, proved abberant as the *HindIII* site downstream of the insert was missing, this abberant plasmid was designated $pTK_2PPChy+^*$. Since it would be necessary to be able to excise the preprochymosin insert for manipulation $pTK_2PPChy+$ was reconstructed from pTK_2PChy+ and $pTK_2PPChy+^*$ as shown in Figure III.5b.

III.E.2. Expression of Chymosin cDNA Constructs in Oocytes

The recombinants $pTK_2PPChy+$, pTK_2PChy+ and pTK_2Chy+ were injected into oocyte nuclei and the injected oocytes were labelled and fractionated as described earlier (II.H). Samples of the oocyte fractions were immunoprecipitated with prochymosin antisera raised at Celltech, αpC_{CT}, and the precipitated proteins were analysed by SDS-PAGE. In agreement with the results obtained with preprochymosin mRNA, described in III.D, injection of $pTK_2PPChy+$ DNA, encoding preprochymosin, gave good expression of a protein which is secreted by the oocyte and is precipitated by αpC_{CT} (Fig.III.6a). Within the oocyte most prochymosin is segregated in the membrane fraction and only a small amount of immunospecific protein is seen in the cytosol fraction; this is consistent with the behaviour of most other foreign secretory proteins in *Xenopus* oocytes (Colman & Morser,1979; as discussed in III.C) and is in contrast to the anomalous distribution observed on expression of lysozyme cDNA

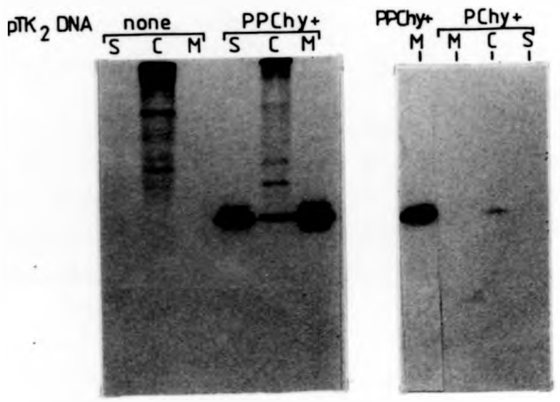


Figure III.6 Expression of chymosin cDNAs in oocytes
The constructs pTK₂PPChy+ (encoding preprochymosin) and pTK₂PChy+ (encoding methionyl-prochymosin) (described in Figs.III.4 & 5) were injected into oocytes which were cultured and processed as described in Fig.III.3 except that anti-prochymosin sera ($ap^{C_{CT}}$, see II.A & II.6) was used for the immunoprecipitating the oocyte fractions.

and mRNA in oocytes. As expected the polypeptide expressed from the PChy cDNA insert, encoding Met-prochymosin, was only found in the cytosol fraction of the oocyte and it showed the same migration on PAGE as the pTK₂PPChy+ product (Fig.III.6b). But expression from pTK₂PChy+ was very poor and was only ever seen in experiments where the expression from pTK₂PPChy+ was very high. It was thought that this low expression may be due to poor stability of the prochymosin in the cytosol compartment where it is not usually localised. On no occasion was any chymosin-specific peptide detected from pTK₂Chy+ which suggests that the chymosin polypeptide encoded by the 'Chy' insert is unstable in oocytes. Evidence is presented later that this insert is capable of producing a translation product *in vitro* (Chapter V).

It should be noted that the sequences flanking the initiation codon in the three cDNA fragments is not the same, and in the PChy & Chy constructs this region differs from the sequence of the preprochymosin cDNA (see Fig.III.4). It is known that the translational efficiency from an initiation codon can be affected by its flanking sequences as well as its position (Kozak,1983b and 1984). From an analysis of sequences upstream from the translational start site in 211 eukaryotic mRNAs Kozak derived a consensus sequence for eukaryotic initiator sites of $CC^{A}/_{G}CCAUGG$; although she comments that there is quite a variation in the extent to which a given mRNA matches the -1 to -5 consensus, and notes that few of the mRNAs studied conformed perfectly (Kozak,1983b). In this context the sequence around the initiation codon in the PPChy cDNA is CAAGGATGG, which matches the consensus sequence at positions -3 and -5 to the initiation ATG. However both the PChy and Chy cDNAs have the sequence GATCAATGG in which none of the -1 to -5 bases match the consensus sequence, so it is likely that this is a poor ribosome binding site which may contribute to the poor expression of the PChy and Chy cDNAs in the *Xenopus* oocyte.

The results we obtained with the PPChy, PChy and Chy cDNAs expressed in the oocyte are in contrast to the work by Mellor *et al*(1983) on the expression of these same inserts in another eukaryotic system, yeast (described in I.D.4). Unlike the localisation observed in *Xenopus* oocytes the chymosin variants did not display their expected compartmentation in yeast. The prochymosin expressed from preprochymosin cDNA in yeast was not secreted but was associated with the cell wall. Chymosin-specific peptides were also expressed from both the Met-prochymosin and Met-chymosin cDNAs but these too were not detected in the cytoplasm and were associated with the yeast cell wall. In

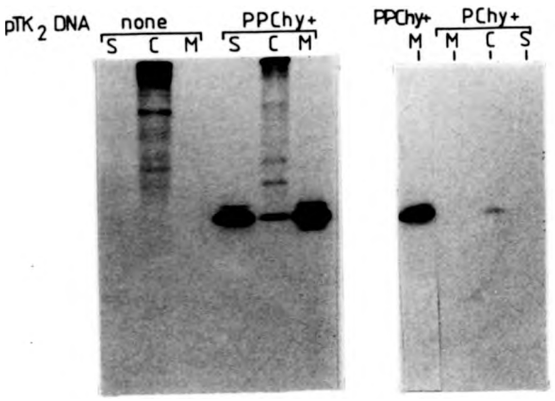


Figure III.6 Expression of chymosin cDNAs in oocytes
The constructs $pTK_2PPChy+$ (encoding preprochymosin) and pTK_2PChy+ (encoding methionyl-prochymosin) (described in Figs.III.4 & 5) were injected into oocytes which were cultured and processed as described in Fig.III.3 except that anti-prochymosin sera (αpC_{CT}, see II.A & II.G) was used for the immunoprecipitating the oocyte fractions.

and mRNA in oocytes. As expected the polypeptide expressed from the PChy cDNA insert, encoding Met-prochymosin, was only found in the cytosol fraction of the oocyte and it showed the same migration on PAGE as the $pTK_2PPChy+$ product (Fig.III.6b). But expression from pTK_2PChy+ was very poor and was only ever seen in experiments where the expression from $pTK_2PPChy+$ was very high. It was thought that this low expression may be due to poor stability of the prochymosin in the cytosol compartment where it is not usually localised. On no occasion was any chymosin-specific peptide detected from pTK_2Chy+ which suggests that the chymosin polypeptide encoded by the 'Chy' insert is unstable in oocytes. Evidence is presented later that this insert is capable of producing a translation product *in vitro* (Chapter V).

It should be noted that the sequences flanking the initiation codon in the three cDNA fragments is not the same, and in the PChy & Chy constructs this region differs from the sequence of the preprochymosin cDNA (see Fig.III.4). It is known that the translational efficiency from an initiation codon can be affected by its flanking sequences as well as its position (Kozak,1983b and 1984). From an analysis of sequences upstream from the translational start site in 211 eukaryotic mRNAs Kozak derived a consensus sequence for eukaryotic initiator sites of $CC^{A}/_{G}CC$**AUGG**; although she comments that there is quite a variation in the extent to which a given mRNA matches the -1 to -5 consensus, and notes that few of the mRNAs studied conformed perfectly (Kozak,1983b). In this context the sequence around the initiation codon in the PPChy cDNA is CAAGG**ATGG**, which matches the consensus sequence at positions -3 and -5 to the initiation ATG. However both the PChy and Chy cDNAs have the sequence GATC**AATGG** in which none of the -1 to -5 bases match the consensus sequence, so it is likely that this is a poor ribosome binding site which may contribute to the poor expression of the PChy and Chy cDNAs in the *Xenopus* oocyte.

The results we obtained with the PPChy, PChy and Chy cDNAs expressed in the oocyte are in contrast to the work by Mellor *et al*(1983) on the expression of these same inserts in another eukaryotic system, yeast (described in I.D.4). Unlike the localisation observed in *Xenopus* oocytes the chymosin variants did not display their expected compartmentation in yeast. The prochymosin expressed from preprochymosin cDNA in yeast was not secreted but was associated with the cell wall. Chymosin-specific peptides were also expressed from both the Met-prochymosin and Met-chymosin cDNAs but these too were not detected in the cytoplasm and were associated with the yeast cell wall. In

addition when Goff *et al*(1984) expressed a Met-prochymosin construct in yeast they too found the prochymosin was not freely soluble in the cytoplasm, while prochymosin expressed in *E.coli* appeared to be localised in cell membranes (Emtage *et al*,1983; Nishimori *et al*,1982b).

While Mellor *et al* did detect a translation product in yeast from the Met-chymosin insert its expression was very poor and over tenfold more immunospecific product was obtained from the insert encoding preprochymosin. In contrast to the relative levels of protein expression seen in oocytes, the Met-prochymosin cDNA gave the best expression in yeast, fivefold better than the preprochymosin cDNA. Mellor *et al* found the differences in levels of expression from the three constructs was not a consequence of differences in amounts of chymosin-specific RNA. They also comment that the low levels of chymosin detected from the Met-chymosin cDNA in yeast may reflect the observed instability of authentic chymosin at pH>5.0; this may also account for our inability to detect any chymosin in oocytes following injection of pTK_2Chy+.

It appears that, unlike yeast, the *Xenopus* oocyte provides a eukaryotic system in which the preprochymosin cDNA displays its expected secretory phenotype. Furthermore the results described in this section also demonstrate that a secretory protein which lacks its signal peptide will not be translocated in the oocyte, although such a nonsegregating polypeptide can be unstable in the cytosol and it may not necessarily be possible to detect the protein at all in the oocyte. The implications of these findings to the objectives of this project are discussed later (III.G).

III.F. Comparison of the Expression of Prelysozyme and Preprochymosin cDNA in Oocytes

To provide a direct comparison of the expression of cDNA encoding the two secretory proteins, lysozyme and prochymosin, the constructs pTK_2Lys+ and $pTK_2PPChy+$ were coinjected into oocytes in similar amounts; this avoids any effect of the noted variation in the translational response to injected DNA of different batches of oocytes (Colman,1984b). The injected oocytes were incubated with [^{35}S]methionine, fractionated and analysed as before, except that in these experiments one set of samples of the C, M & S fractions were immunoprecipitated with αpC_{CT} and a duplicate set were precipitated with the anti-lysozyme sera. These coinjection experiments highlighted several differences in the expression and compartmentalisation of these two secretory proteins.

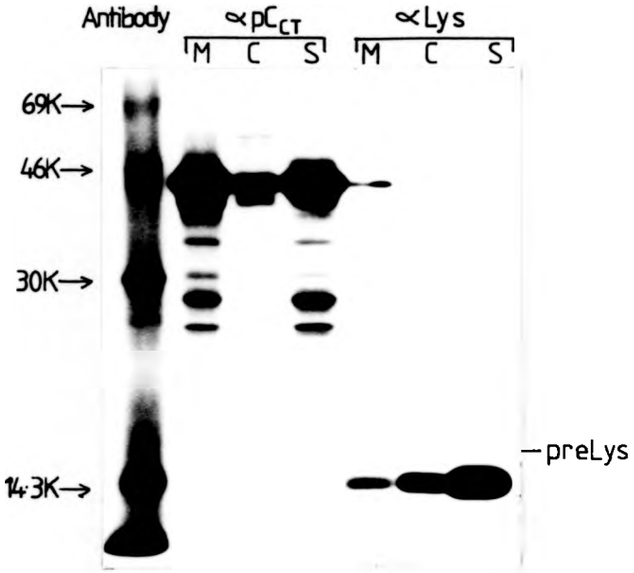


Figure III.7 Synthesis of lysozyme and prochymosin in oocytes after coinjection of DNA
Oocytes were injected with a mixture of equal amounts of the constructs $pTK_2PPChy+$ (encoding preprochymosin) and pTK_2Lys+ (encoding prelysozyme). They were then cultured and analysed as in Fig.III.3 except that one set of samples of the oocyte fractions were immunoprecipitated with antiprochymosin sera (αpC_{CT}) and a duplicate set with anti-lysozyme (∝lys). The membrane (M) and cytosol (C) fraction tracks represent 2.5 oocytes and the secreted protein track (S) 5 oocytes. 'preLys' marks the expected position on this gel of a polypeptide with the molecular weight of unprocessed prelysozyme (M_r=16,200).

Firstly it was observed from a number of experiments that $pTK_2PPChy+$ was always well expressed and consistently gave better expression than pTK_2Lys+, while the expression of lysozyme showed more susceptibility to the batch variation of the oocytes, noted above. The different distribution of the two proteins between the C, M & S fractions of the injected oocytes is clear (Fig.III.7). As before lysozyme is well secreted by the oocytes, and within the oocyte more lysozyme is associated with the cytosol fraction than with the membrane fraction. In comparison to the compartmentation observed for lysozyme, less of the prochymosin expressed from $pTK_2PPChy+$ in the same oocytes is secreted over the same time period, and more prochymosin is detected in the M fraction than in the C fraction (bearing in mind that in Fig.III.7 the S track represents twice as many oocytes as the C & M tracks). Different rates of secretion of foreign secretory proteins from oocytes has been observed previously by Cutler *et al*(1981). In that instance more detailed study showed the intrinsic rate of lysozyme secretion from oocytes was 12 times that of ovalbumin, following injection of chick oviduct mRNA which encodes both proteins, and that the differential rates of transport were not a consequence of competition for either *Xenopus* or avian factors. It appears from these results with oviduct mRNA and those presented in this chapter that lysozyme is particularly efficiently secreted by oocytes compared with other secretory proteins; however these too may display differential rates of secretion. Such non-parallel kinetics of secretion argues against bulk intracellular transport of proteins destined for export as the sole means by which proteins are transported to the cell surface for secretion from *Xenopus* oocytes; but the basis for the different rates of intracellular transport is not known.

The localisation of most prochymosin within the oocyte in the vesicle fraction is as expected and, as noted earlier, it can be argued that the prochymosin detected in the cytosol fraction is due to the crude fractionation procedure used, which disrupts some of the microsomal vesicles. However the relative proportions of prochymosin in the C and M fractions provides an internal reference against which to compare the abnormally large amounts of (apparently processed) lysozyme found in the cytosolic fraction. Hence the large proportion of cytosolic lysozyme cannot just be due to leakage from ER microsomal vesicles disrupted during fractionation, in the same way as the prochymosin seen in the cytosol fraction. It is possible, however, that the lysozyme associated with the cytosol fraction of the oocyte was sequestered in a membrane-bound 'compartment', distinct from the ER, and this unidentified

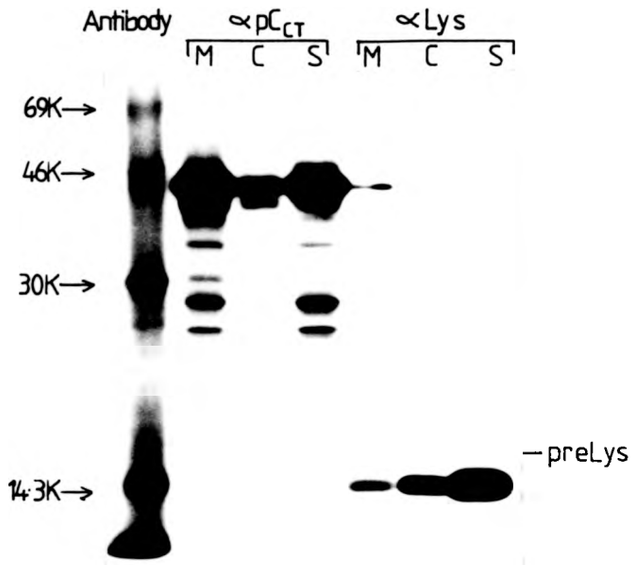


Figure III.7 Synthesis of lysozyme and prochymosin in oocytes after coinjection of DNA
Oocytes were injected with a mixture of equal amounts of the constructs pTK$_2$PPChy+ (encoding preprochymosin) and pTK$_2$Lys+ (encoding prelysozyme). They were then cultured and analysed as in Fig.III.3 except that one set of samples of the oocyte fractions were immunoprecipitated with antiprochymosin sera (∝pC$_{CT}$) and a duplicate set with anti-lysozyme (∝lys). The membrane (M) and cytosol (C) fraction tracks represent 2.5 oocytes and the secreted protein track (S) 5 oocytes. 'preLys' marks the expected position on this gel of a polypeptide with the molecular weight of unprocessed prelysozyme (M_r=16,200).

Firstly it was observed from a number of experiments that pTK$_2$PPChy+ was always well expressed and consistently gave better expression than pTK$_2$Lys+, while the expression of lysozyme showed more susceptibility to the batch variation of the oocytes, noted above. The different distribution of the two proteins between the C, M & S fractions of the injected oocytes is clear (Fig.III.7). As before lysozyme is well secreted by the oocytes, and within the oocyte more lysozyme is associated with the cytosol fraction than with the membrane fraction. In comparison to the compartmentation observed for lysozyme, less of the prochymosin expressed from pTK$_2$PPChy+ in the same oocytes is secreted over the same time period, and more prochymosin is detected in the M fraction than in the C fraction (bearing in mind that in Fig.III.7 the S track represents twice as many oocytes as the C & M tracks). Different rates of secretion of foreign secretory proteins from oocytes has been observed previously by Cutler *et al*(1981). In that instance more detailed study showed the intrinsic rate of lysozyme secretion from oocytes was 12 times that of ovalbumin, following injection of chick oviduct mRNA which encodes both proteins, and that the differential rates of transport were not a consequence of competition for either *Xenopus* or avian factors. It appears from these results with oviduct mRNA and those presented in this chapter that lysozyme is particularly efficiently secreted by oocytes compared with other secretory proteins; however these too may display differential rates of secretion. Such non-parallel kinetics of secretion argues against bulk intracellular transport of proteins destined for export as the sole means by which proteins are transported to the cell surface for secretion from *Xenopus* oocytes; but the basis for the different rates of intracellular transport is not known.

The localisation of most prochymosin within the oocyte in the vesicle fraction is as expected and, as noted earlier, it can be argued that the prochymosin detected in the cytosol fraction is due to the crude fractionation procedure used, which disrupts some of the microsomal vesicles. However the relative proportions of prochymosin in the C and M fractions provides an internal reference against which to compare the abnormally large amounts of (apparently processed) lysozyme found in the cytosolic fraction. Hence the large proportion of cytosolic lysozyme cannot just be due to leakage from ER microsomal vesicles disrupted during fractionation, in the same way as the prochymosin seen in the cytosol fraction. It is possible, however, that the lysozyme associated with the cytosol fraction of the oocyte was sequestered in a membrane-bound 'compartment', distinct from the ER, and this unidentified

compartment is particularly fragile and does not remain intact in the crude fractionation procedure.

The signal peptide has apparently been cleaved from the lysozyme detected in the oocyte fractions as judged by its migration to the same position as the lysozyme in the marker track (M_r=14,300), and its greater mobility than that expected of prelysozyme (M_r=16,200), see Fig.III.7. The resolution of the gel in Fig.III.7 is insufficient to make a similar assessment as to whether the major protein immunoprecipitated by antiprochymosin has a mobility consistent with prochymosin (40,800) or preprochymosin (42,600). Evidence is presented later that both the lysozyme and prochymosin proteins encoded by the Lys and PPChy inserts are processed relative to their respective precursor forms when expressed in oocytes (see Chapters V & VI).

III.G. Discussion

As a prerequisite to the site specific mutagenesis of cDNA encoding the signal sequence of a eukaryotic secretory protein it was necessary to demonstrate that the wild-type cDNA expressed a protein that was correctly compartmented in *Xenopus* oocytes, which was to be the *in vivo* assay system employed initially to characterise the translocation phenotype of signal sequence mutants. The results described in this chapter establish that cDNAs encoding two secretory proteins, lysozyme and prochymosin, cloned into the vector pTK_2 were expressed after injection into oocyte nuclei, and the foreign proteins showed the same compartmentation pattern they exhibit following cytoplasmic injection of the corresponding mRNAs. Although lysozyme was very efficiently secreted from oocytes, the expression of pTK_2Lys+ was quite variable in different batches of oocytes and resulted in an unusual compartmentation of lysozyme in the cytosolic fraction of the injected oocytes. This could complicate analysis of the effect of signal sequence mutations of lysozyme on translocation. The construct pTK_2PPChy+, encoding preprochymosin, gave better and more consistent expression; the prochymosin expressed displayed the usual fractionation pattern in the oocyte observed for most foreign secretory proteins and, although less efficiently secreted than lysozyme, it was readily detected in medium surrounding injected oocytes. These considerations suggest that preprochymosin may be a better choice than prelysozyme for the intended mutagenesis studies.

In this context the expression in oocytes of pTK_2PChy+, which contains cDNA encoding methionyl-prochymosin, showed that the absence of

the signal peptide prevents the translocation of prochymosin in oocytes and the protein remains in the cytosol. However this plasmid gave poor expression, and no chymosin-specific product was detected from pTK_2Chy+, encoding methionyl-chymosin. A similar situation was encountered in our work towards identifying the signal sequence region of ovalbumin, using the oocyte as an assay system for mutated ovalbumin cDNAs (outlined briefly in I.E. and see Tabe *et al*,1984 in Appendix). In this case one construct, lacking codons 20 to 145, resulted in the expression in oocytes of a truncated ovalbumin which was localised exclusively in the cyosol, and hence this mutant must lack a functional signal sequence. However for another mutant, in which the first 40 amino acid residues of ovalbumin were deleted, although transcripts were detected from this construct no protein was immunoprecipitated from the injected oocytes by anti-ovalbumin sera. These observations raised a number of points to be considered in using the *Xenopus* oocyte as an assay system to characterise the translocation phenotype of signal sequence mutants of preprochymosin. The results of this chapter imply that a nonsegregated preprochymosin may be unstable in the oocyte cytosol and the protein may not be detected at all following injection of the mutated cDNA. If no chymosin-specific polypeptide were detected it would raise the question of whether the cDNA had been expressed but the protein rapidly degraded in the cytoplasm, or the mutagenesis had affected sequences involved in translational signals. This latter situation could occur if methods were used in which the 5' untranslated sequence as well as the signal sequence in the preprochymosin cDNA were exposed to mutagenesis (see Chapter VII). Therefore *in vitro* systems may need to be adopted to determine if a particular mutant which is not detected in the oocyte is translated, but carries a mutation which prevents its translocation. As described earlier (I.E and at the beginning of this chapter) it is intended to use *in vitro* systems in the characterisation of signal sequence mutants. It is anticipated, however, that those proteins which are translocated will be detected following expression of the mutated preprochymosin cDNA in the oocyte, and these will be secreted.

III.H. Summary

Plasmids derived from pBR/SV40 'shuttle' vectors constructed by Mulligan & Berg(1980) were used as vectors for the expression of cDNAs encoding eukaryotic secretory proteins in *Xenopus* oocytes. It was found that the TK promoter and a retroviral LTR gave better expression of lysozyme cDNA than the original SV40 early promoter. The vector pTK_2, which contained the TK promoter of *Herpes Simplex* Virus and lacked an SV40 intron, was selected for further studies.

With a view to selecting a eukaryotic secretory protein for studies on the effect of mutagenesis of the signal sequence on translocation, a study was made of the expression in oocytes of cDNA encoding preprochymosin and prelysozyme, cloned into pTK_2; and the compartmentation of these proteins was characterised. Both proteins were secreted from oocytes. Whereas prochymosin fractionated as expected mainly with the membranes and vesicles within the oocyte, more lysozyme was found in the cytosolic compartment than in the membrane fraction; this was consistent with published results of the injection of lysozyme mRNA into oocytes (Cutler *et al*,1981). However the expression of preprochymosin cDNA in yeast (Mellor *et al*,1983) had been reported not to result in the secretion of prochymosin.

Injection of cDNA encoding prochymosin without its signal sequence resulted in the synthesis of prochymosin protein which was exclusively localised in the oocyte cytosol. This is in contrast to published reports of the expression of prochymosin cDNA constructs in other *in vivo* systems which observe the prochymosin was not located in the cytoplasm but was associated with cell membranes (Emtage *et al*,1983; Goff *et al*,1984; Mellor *et al*,1983; Nishimori *et al*,1982b). No chymosin-specific polypeptide was detected in oocytes from another construct derived from preprochymosin cDNA, which also lacked the signal sequence and encoded a methionyl-chymosin polypeptide.

This oocyte coupled transcription-translation assay for cloned cDNAs, and the compartmentation of their translation products, was also used in this laboratory in a project to identify the region of ovalbumin which functions as the signal sequence. This work, in which I was also involved, is described in Tabe *et al*(1984) which is included in the Appendix.

IV. PROCESSING OF PREPROCHYMOSIN EXPRESSED FROM mRNA IN *XENOPUS* OOCYTES AND *IN VITRO*

IV.A. Introduction

Since it was decided that preprochymosin was a suitable candidate for the proposed signal sequence mutagenesis, further studies were made of the segregation of the wild-type protein, using preprochymosin mRNA which could be expressed in the oocyte and also *in vitro*. Prior to the SP6 transcription system becoming available the preprochymosin mRNA provided a means of characterizing the translation and translocation of wild-type preprochymosin in cell-free systems, which would later be used to analyse the translocation of signal sequence mutants. It was particularly hoped to demonstrate that, in common with almost all other eukaryotic secretory proteins, the signal peptide of preprochymosin is processed on translocation of the precursor *in vitro* and *in vivo*.

The presence of a transient N-terminal signal sequence for prochymosin has been implied from the following experimental observations. As described in the Introduction (I.D.) the nucleotide sequence of cDNA copies of the mRNA shows that after the initiation ATG there are another 15 codons before the mRNA encodes the known amino acid sequence of prochymosin, which has been determined by sequencing the protein isolated from calf stomachs (Foltmann *et al*,1977; Harris *et al*,1982; Moir *et al*,1982). It was anticipated, therefore, that the primary translation product from the mRNA would be a precursor, preprochymosin, with an expected molecular weight of 42.5K as compared with the molecular weight of prochymosin, 40.8K, calculated from the amino acid sequence. Several groups have studied the expression *in vitro* of polyA RNA extracted from calf abomasum (Nishimori *et al*,1981; Harris *et al*,1982; Moir *et al*,1982; Nicholson & Jones,1984; McConnell *et al*,1984) and found that it produced one major polypeptide, which is precipitated by antibodies to prochymosin, and had an apparent molecular weight from SDS-PAGE variously estimated to be between 40K and 43K. This was assumed to be the precursor preprochymosin from which the signal sequence would be processed on cotranslational translocation across ER membranes. However none of the above groups have actually shown this to be the case using *in vitro* translation in the presence of microsomes. Also the protein expressed in yeast from cloned preprochymosin cDNA (Mellor *et al*,1983), which was thought to be prochymosin and was associated with membranes (but not secreted - see III.E.), was not shown experimentally

to be processed compared with the putative full-length precursor seen in *in vitro* translations of the mRNA. The initial experiments carried out in this laboratory on the expression of preprochymosin mRNA (described earlier, III.D.), showed that following injection of polyA RNA from calf abomasum into oocytes a protein was secreted with the characteristics of prochymosin. Unless prochymosin is post-translationally modified in oocytes then the secreted prochymosin should be smaller than the primary translation product of the mRNA. However, in these experiments there was no detectable size difference on SDS-PAGE between the polypeptide produced by *in vitro* translation of preprochymosin mRNA and that segregated and secreted by injected oocytes (Krieg *et al*,1984). Similarly the experiments described on the expression of cloned chymosin genes in the oocyte (III.E. & F.) did not rigorously show that the product secreted from oocytes injected with $pTK_2PPChy+$ (which encodes preprochymosin) was the processed prochymosin. As noted earlier (I.D. & III.E.), published reports on the expression of cDNA encoding prochymosin without its signal sequence, suggested that this signal-minus polypeptide contained a topogenic sequence which directed its association with membranes in *E.coli* and in yeast (Nishimori *et al*,1982b; Emtage *et al*,1983; Mellor *et al*,1983; Goff *et al*,1984); although work presented in this thesis (III.E.) showed that in oocytes injection of a cDNA encoding methionyl-prochymosin resulted in the expression of a prochymosin protein localised only in the cytoplasm.

It was important to resolve these discrepancies and demonstrate that the precursor form of prochymosin encoded by the mRNA does contain a signal sequence which is cleaved on translocation across ER membranes, as expected and in common with all eukaryotic secretory proteins examined to date, with the already noted exception of ovalbumin (Palmiter *et al*,1978; Lingappa *et al*,1978b; see I.E.). In terms of characterizing the effect on translocation of introducing site-specific mutations into the preprochymosin signal sequence the demonstration of signal cleavage provides a useful indicator for at least the initiation of translocation. Although it is noted that it is possible that the introduction of mutations towards the end of the signal sequence could destroy the signal peptidase recognition site and abolish processing without preventing the initiation of translocation (Hortin & Boime 1980 and 1981; Schauer *et al*,1985; Perlman & Halvorson,1983; Von Heijne,1983 and 1984a; discussed in the Introduction, see I.A.2 and I.B.3).

It was therefore necessary to examine whether the signal sequence of preprochymosin is cleaved by the oocyte and by *in vitro* translocation

systems. Owing to the unexpected results obtained in characterizing the processing of preprochymosin expressed from the mRNA, the experiments presented in this chapter were carried out over an extended period; during which time the SP6 transcription system became available in this laboratory. Although synthetic SP6 RNAs were used in experiments included here the SP6 system will be introduced and described in the next chapter(V); together with further experiments on the expression of SP6 RNAs derived from the chymosin cDNAs.

IV.B. Processing of Preprochymosin in *Xenopus* Oocytes

Oocytes were injected with preprochymosin mRNA, then labelled with [^{35}S]methionine for 7h, after which the incubation medium was removed for analysis of secreted proteins, then the oocytes were homogenised and fractionated to separate cytosolic proteins from those associated with membranes (as described in the Methods II.H). Samples of the oocyte fractions were immunoprecipitated with prochymosin antibodies raised at Warwick (αpC_{RB}, the raising and characterization of these antibodies is described in II.G.) and then separated by SDS-PAGE on a 9%(w/v) polyacrylamide gel, containing 30% acrylamide:1.6% bisacrylamide (30:1.6 cross-linked), as described in II.F.2. It was hoped that this gel composition would provide a better resolution of prochymosin from its precursor than the SDS-polyacrylamide gels used previously (12.5% polyacrylamide, 30:0.825 cross-linked); the alternative composition was chosen since it was the same as that used by Mellor *et al*(1983) in studies of expression of cloned chymosin genes in yeast, although these workers do not show that this gel system will resolve the prochymosin expressed in yeast from the *in vitro* translation product of preprochymosin mRNA. To provide a marker in our experiments for the primary translation product of preprochymosin mRNA, polyA RNA extracted from calf abomasum (II.A) was translated in the rabbit reticulocyte lysate cell-free system, as described in the Methods section (II.E.2). An aliquot from this *in vitro* translation was electrophoresed, without immunoprecipitation, on the same gel as the precipitated samples from injected oocytes.

Analysis of Figure IV.1 shows that preprochymosin mRNA gave a doublet of two translation products both when translated *in vitro* (track 2) and in the oocyte (tracks 3-5). This was unexpected since, as noted earlier (in section A) and in the Introduction (I.D.2), other workers have reported that *in vitro* translation of preprochymosin mRNA in the reticulocyte lysate system gives a single major polypeptide

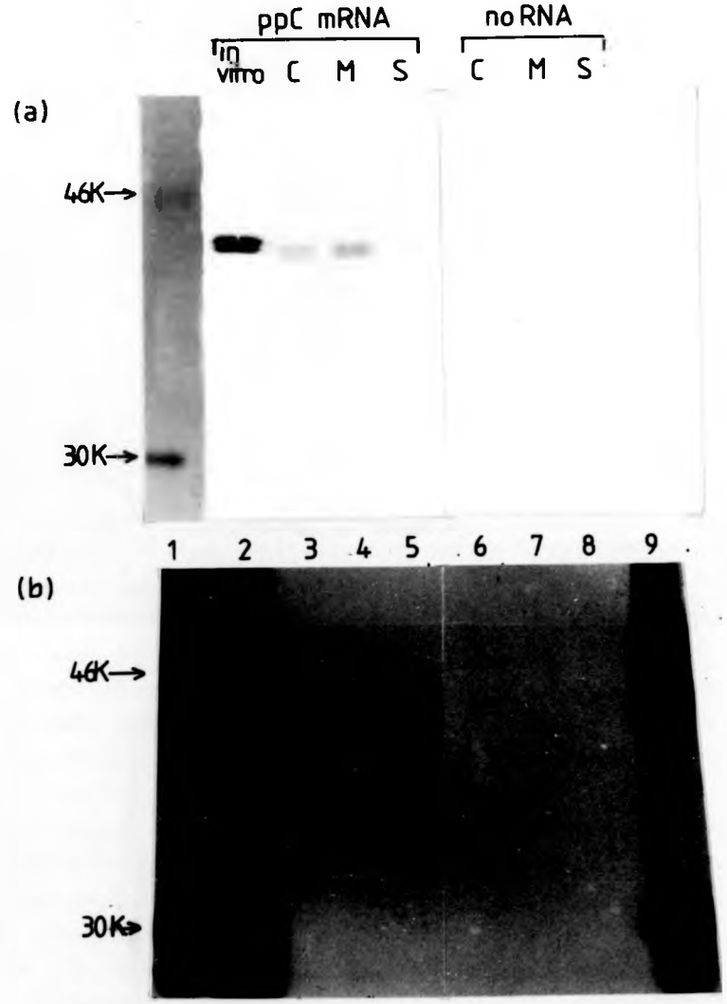


Figure IV.1 Expression of preprochymosin mRNA in *Xenopus* oocytes and *in vitro*
Oocytes were injected with preprochymosin mRNA (ppC mRNA) and incubated for 7h with [^{35}S]methionine, together with uninjected control oocytes (no RNA). The oocytes were fractionated into secreted(S), cytosolic(C) and membrane(M) components, as described in the Methods (II.H). Samples of the different fractions, equivalent to 1 oocyte, were immunoprecipitated with antiprochymosin antibodies (αpC_{ca}, see II.G) and analysed by SDS-PAGE on a 9% polyacrylamide gel (30:1.6 cross-linked). The ppC mRNA was also translated in a reticulocyte lysate *in vitro* system (at a ppC mRNA concentration of 12.5ng/µl, and 8µCi/µl [^{35}S]methionine), as described in the Methods II.E.2. An aliquot (2µl) from this translation was loaded on the same gel (track 2). The gel was fluorographed and exposed to X-ray sensitive film for (a) 17.5h (b) 3 days (II.F.3). When preparing the print in (a), in order to show clearly the radiolabelled protein bands, track 2 was exposed for a shorter period than tracks 3-8; and tracks 1 & 9, containing ^{14}C-labelled protein molecular weight markers (see II.F.2), were exposed for longer than the rest to bring up the marker bands. The arrows mark the position of two of the molecular weight markers, M_r=46,000 and M_r=30,000.

product on SDS-PAGE; Moir *et al*(1982) and Nicholson & Jones(1984) used a 10%(w/v) polyacrylamide gel (cross linking ratio not given), but Harris *et al*(1982) do not specify the composition of SDS-polyacrylamide gels they used. The significance of the doublet of products is discussed fully later (IV.D.). The apparent molecular weight of the *in vitro* polypeptides were approx. M_r=43,000 and M_r=42,500, whilst those in the secreted, membrane and cytosol fractions of the oocyte were M_r=42,000 and 41,500. Thus it appears that both species produced from translation of preprochymosin mRNA *in vitro* can be translocated, cleaved and secreted when expressed by the oocyte. It should be noted that in this experiment the injected oocytes were incubated for a relatively short period, 7h, with radiolabelled methionine (compared with the 24h labelling period used in previous experiments), and this limited the chase of labelled prochymosin into the medium, hence the amount of protein immunoprecipitated in the S fraction is small but distinct (Fig.IV.1b, track 5). It is likely that the relatively large amounts of the polypeptide doublet seen in the cytosol fraction (track 3), which is of the same size as that in the membranes, represents particularly bad leakage from membranes during fractionation of the injected oocytes; this is probably oocyte batch dependent since far less 'leakage' was seen in other experiments and more commonly within the oocyte the membrane fraction contained most of the prochymosin (eg. Fig.III.6). Apparently the two species encoded by the preprochymosin mRNA could not be resolved on the 12.5% polyacrylamide (30:0.825 cross-linked) gels used previously for the analysis of translation products (see Fig.II.1 and Krieg *et al*,1984 in Appendix). From the gel shown in Fig.IV.1 the apparent size reduction on cleavage of each preprochymosin is approx. M_r=1,000 and not the 1,600 expected from the amino acid composition of the signal peptide. This could be due to glycosylation of the signal-processed, translocated prochymosins.

A closer comparison was later made of the *in vitro* and oocyte translation products of preprochymosin mRNA and the prochymosin cDNA construct 'PChy' (described in III.E.). The PChy translation product will provide a marker for prochymosin, without its signal peptide; also the prochymosin expressed in *Xenopus* oocytes from the PChy construct will not be glycosylated since it is not translocated (see III.E). Figure IV.2a. shows the separation of material immunoprecipitated by prochymosin antibodies (αpC_{ca}) from the oocyte and *in vitro* translation products of preprochymosin mRNA (tracks 5 & 6, and track 2 respectively); along with that obtained from expression of synthetic RNA encoding

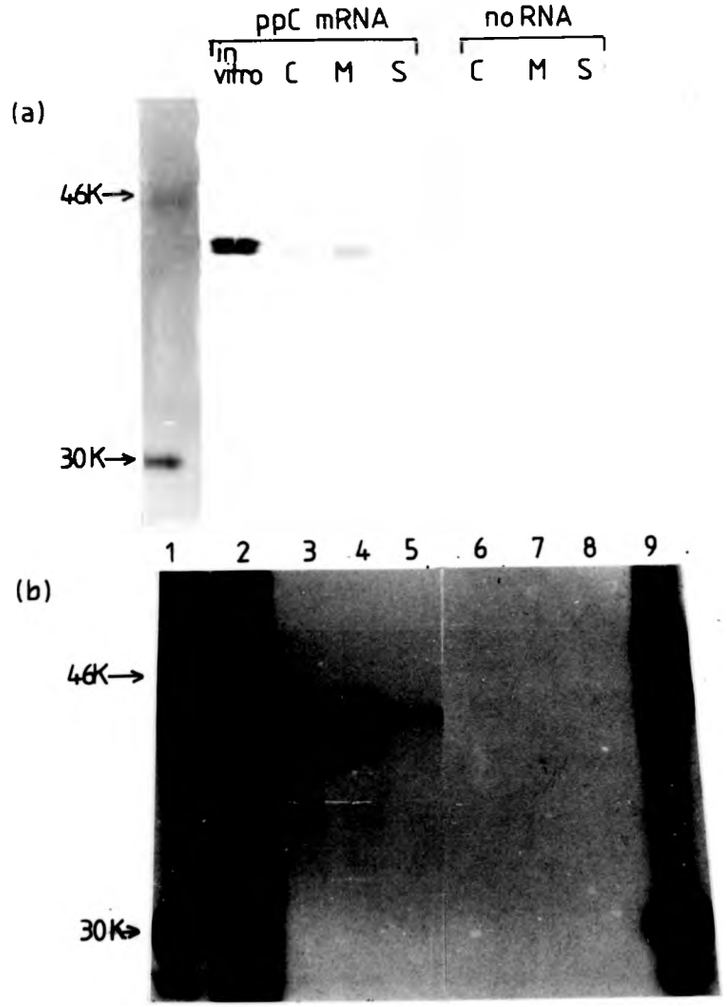


Figure IV.1 Expression of preprochymosin mRNA in *Xenopus* oocytes and *in vitro*
Oocytes were injected with preprochymosin mRNA (ppC mRNA) and incubated for 7h with [^{35}S]methionine, together with uninjected control oocytes (no RNA). The oocytes were fractionated into secreted(S), cytosolic(C) and membrane(M) components, as described in the Methods (II,H). Samples of the different fractions, equivalent to 1 oocyte, were immunoprecipitated with antiprochymosin antibodies (αpC_{Ms}, see II,G) and analysed by SDS-PAGE on a 9% polyacrylamide gel (30:1,6 cross-linked). The ppC mRNA was also translated in a reticulocyte lysate *in vitro* system (at a ppC mRNA concentration of 12,5ng/µl, and 8µCi/µl [^{35}S]methionine), as described in the Methods II,E,2. An aliquot (2µl) from this translation was loaded on the same gel (track 2). The gel was fluorographed and exposed to X-ray sensitive film for (a) 17,5h (b) 3 days (II,F,3). When preparing the print in (a), in order to show clearly the radiolabelled protein bands, track 2 was exposed for a shorter period than tracks 3-8; and tracks 1 & 9, containing ^{14}C-labelled protein molecular weight markers (see II,F,2), were exposed for longer than the rest to bring up the marker bands. The arrows mark the position of two of the molecular weight markers, M_r=46,000 and M_r=30,000.

product on SDS-PAGE; Moir *et al*(1982) and Nicholson & Jones(1984) used a 10%(w/v) polyacrylamide gel (cross linking ratio not given), but Harris *et al*(1982) do not specify the composition of SDS-polyacrylamide gels they used. The significance of the doublet of products is discussed fully later (IV.D.). The apparent molecular weight of the *in vitro* polypeptides were approx. M_r=43,000 and M_r=42,500, whilst those in the secreted, membrane and cytosol fractions of the oocyte were M_r=42,000 and 41,500. Thus it appears that both species produced from translation of preprochymosin mRNA *in vitro* can be translocated, cleaved and secreted when expressed by the oocyte. It should be noted that in this experiment the injected oocytes were incubated for a relatively short period, 7h, with radiolabelled methionine (compared with the 24h labelling period used in previous experiments), and this limited the chase of labelled prochymosin into the medium, hence the amount of protein immunoprecipitated in the S fraction is small but distinct (Fig.IV.1b, track 5). It is likely that the relatively large amounts of the polypeptide doublet seen in the cytosol fraction (track 3), which is of the same size as that in the membranes, represents particularly bad leakage from membranes during fractionation of the injected oocytes; this is probably oocyte batch dependent since far less 'leakage' was seen in other experiments and more commonly within the oocyte the membrane fraction contained most of the prochymosin (eg. Fig.III.6). Apparently the two species encoded by the preprochymosin mRNA could not be resolved on the 12.5% polyacrylamide (30:0.825 cross-linked) gels used previously for the analysis of translation products (see Fig.II.1 and Krieg *et al*,1984 in Appendix). From the gel shown in Fig.IV.1 the apparent size reduction on cleavage of each preprochymosin is approx. M_r=1,000 and not the 1,600 expected from the amino acid composition of the signal peptide. This could be due to glycosylation of the signal-processed, translocated prochymosins.

A closer comparison was later made of the *in vitro* and oocyte translation products of preprochymosin mRNA and the prochymosin cDNA construct 'PChy' (described in III.E.). The PChy translation product will provide a marker for prochymosin, without its signal peptide; also the prochymosin expressed in *Xenopus* oocytes from the PChy construct will not be glycosylated since it is not translocated (see III.E). Figure IV.2a. shows the separation of material immunoprecipitated by prochymosin antibodies (αpC_{Ms}) from the oocyte and *in vitro* translation products of preprochymosin mRNA (tracks 5 & 6, and track 2 respectively); along with that obtained from expression of synthetic RNA encoding

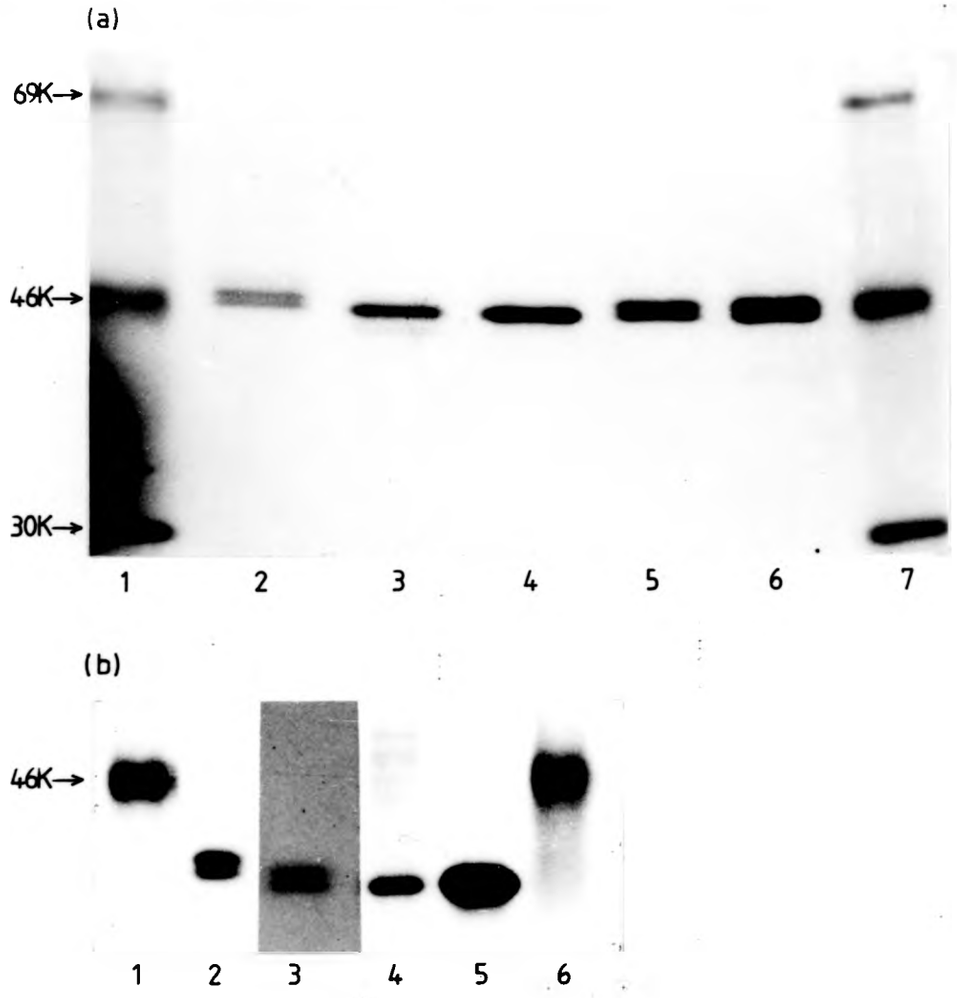


Figure IV.2 Comparison of proteins expressed from preprochymosin mRNA and chymosin cDNAs. Samples of the translation products of preprochymosin mRNA and cloned cDNAs encoding preprochymosin and prochymosin were immunoprecipitated with antiprochymosin antibodies (αpC_{ms}, see II.6) and then separated by SDS-PAGE as described below.
(a) Shows a linear gradient (10-30%) polyacrylamide gel fluorographed using 'Enhance' and exposed for 3 days to X-ray sensitive film. Track 2 contains an aliquot of the reticulocyte lysate translation of preprochymosin mRNA described in Fig.IV.1. Tracks 5 & 6 are the cytosol and membrane fractions, respectively, of oocytes injected with preprochymosin mRNA (see Fig.IV.1). Track 3 is capped SP6PChy RNA (encoding methionyl-prochymosin, described in V.B) translated in a wheat germ cell-free system (II.E.1) and track 4 is the product of the same synthetic RNA expressed in the oocyte (cytosol fraction). The molecular weights of the radiolabelled marker proteins in tracks 1 & 7 are indicated.
(b) Shows the autoradiogram of a 9% polyacrylamide gel (30:1.6 cross-linked), fluorographed using the method of Bonner & Laskey(1974) and exposed to preflashed X-ray film for 3 days. Track 2 is the immunospecific products of preprochymosin mRNA translated *in vitro* (as track 2 above) and track 3 is the same mRNA expressed in the oocyte (membrane fraction). Track 4 is material immunoprecipitated from total homogenised oocytes injected with $pTK_2PPChy+$ DNA encoding preprochymosin (see III.E). Track 5 is the oocyte product of SP6PChy RNA (as track 4 above). Tracks 1 & 6 are the ^{14}C-labelled protein markers. To show the bands in track 3 more clearly that area of the negative was exposed longer during printing.

methionyl-prochymosin, SP6PChy (see V.B), *in vitro* (track 3) and in the oocyte (track 4). In this experiment a 10-30%(w/v) linear gradient SDS-polyacrylamide gel was used as it was decided to test if this would result in greater separation of prochymosin from its precursor preprochymosin, compared with that seen on the 9%(w/v) polyacrylamide gel used above. The preparation of this gel, its electrophoresis and fluorography using 'Enhance' (NEN) was kindly carried out by A.S. Carver. This gel (Fig.IV.2a) did not give better resolution of the preprochymosin(s) and prochymosin(s) than the 9%(w/v) polyacrylamide gel (30:1.6 cross-linked). As expected there no was detectable difference in size between the *in vitro* and oocyte products expressed from the SP6PChy RNA, encoding methionyl-prochymosin (tracks 3 & 4, respectively). It was shown, however, that the polypeptide translated from this synthetic RNA appeared to display the same mobility as the faster migrating prochymosin protein of the two species expressed in oocytes from the preprochymosin mRNA, and probably had a marginally greater mobility than either of the *in vitro* preprochymosin mRNA proteins, although the comparison of the small differences in relative migration of the polypeptides is difficult between tracks. These observations were confirmed when a similar set of immunoprecipitated samples were electrophoresed on a nongradient 9%(w/v) polyacrylamide gel (ratio %acrylamide to %bisacrylamide 30:1.6), shown in Fig.IV.2b. This gel demonstrated more clearly that the oocyte SP6PChy translation product (track 5) has a mobility greater than that displayed by either of the primary translation products translated from the mRNA *in vitro* (track 2). The prochymosin expressed from the SP6PChy RNA again migrated to the same position as the lower band of the doublet expressed from preprochymosin mRNA in the oocyte (track 3). This suggests that glycosylation of the translocated prochymosins expressed from the mRNA in oocytes does not account for their smaller than expected difference in the migration on SDS-PAGE to that of the precursor preprochymosins. Furthermore the protein expressed in oocytes from $pTK_2PPChy+$ DNA (track 4) also showed the same migration as the faster migrating 'oocyte' preprochymosin mRNA product and the SP6PChy prochymosin protein. $pTK_2PPChy+$, which contains full-length preprochymosin cDNA insert (see III.E.), is known to express a protein in oocytes which is translocated and secreted (Fig.III.6); the results shown in Fig.IV.2b suggests that the product immunoprecipitated from oocytes injected with $pTK_2PPChy+$ does not contain the signal peptide. At this time the corresponding SP6 RNA of the 'PPChy' cDNA was not available. Further experiments on the expression of

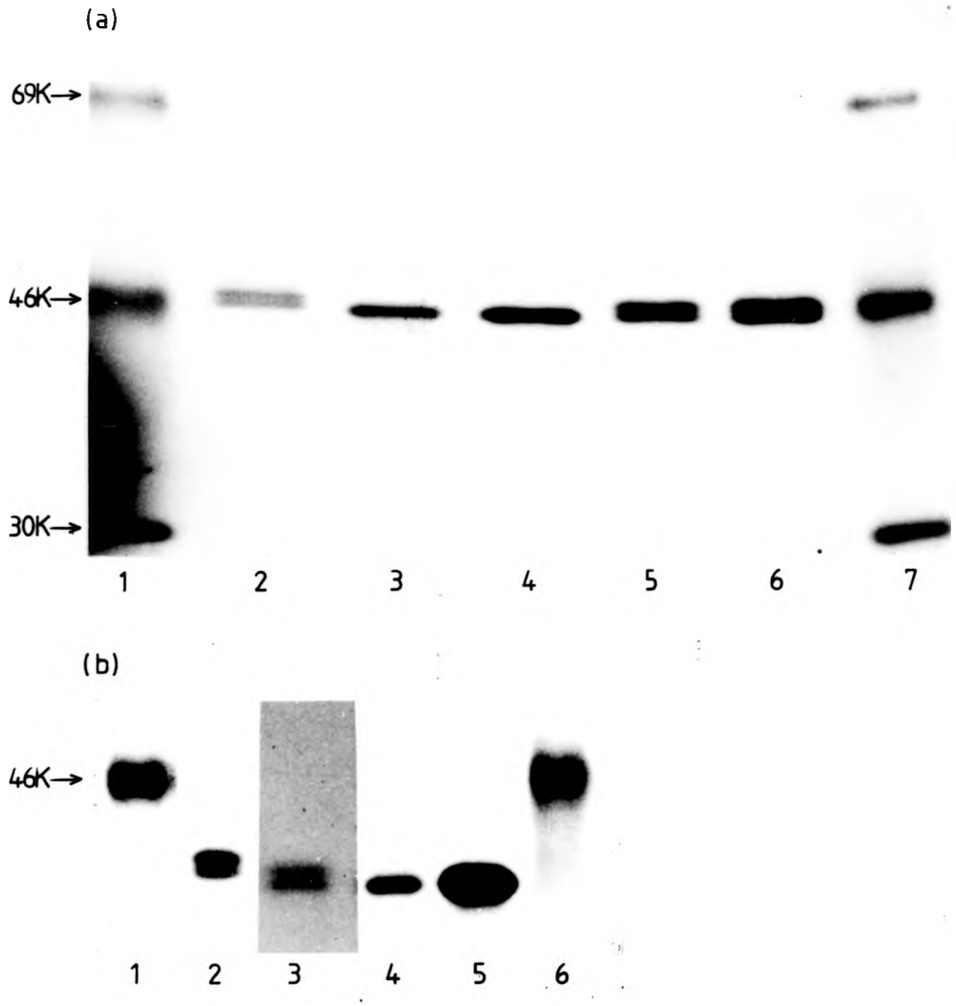


Figure IV.2 Comparison of proteins expressed from preprochymosin mRNA and chymosin cDNAs. Samples of the translation products of preprochymosin mRNA and cloned cDNAs encoding preprochymosin and prochymosin were immunoprecipitated with antiprochymosin antibodies (αpC_{ms}, see II.G) and then separated by SDS-PAGE as described below.
(a) Shows a linear gradient (10-30%) polyacrylamide gel fluorographed using 'Enhance' and exposed for 3 days to X-ray sensitive film. Track 2 contains an aliquot of the reticulocyte lysate translation of preprochymosin mRNA described in Fig.IV.1. Tracks 5 & 6 are the cytosol and membrane fractions, respectively, of oocytes injected with preprochymosin mRNA (see Fig.IV.1). Track 3 is capped SP6PChy RNA (encoding methionyl-prochymosin, described in V.B) translated in a wheat germ cell-free system (II.E.1) and track 4 is the product of the same synthetic RNA expressed in the oocyte (cytosol fraction). The molecular weights of the radiolabelled marker proteins in tracks 1 & 7 are indicated.
(b) Shows the autoradiogram of a 9% polyacrylamide gel (30:1.6 cross-linked), fluorographed using the method of Bonner & Laskey(1974) and exposed to preflashed X-ray film for 3 days. Track 2 is the immunospecific products of preprochymosin mRNA translated *in vitro* (as track 2 above) and track 3 is the same mRNA expressed in the oocyte (membrane fraction). Track 4 is material immunoprecipitated from total homogenised oocytes injected with $pTK_2PPChy+$ DNA encoding preprochymosin (see III.E). Track 5 is the oocyte product of SP6PChy RNA (as track 4 above). Tracks 1 & 6 are the ^{14}C-labelled protein markers. To show the bands in track 3 more clearly that area of the negative was exposed longer during printing.

methionyl-prochymosin, SP6PChy (see V.B), *in vitro* (track 3) and in the oocyte (track 4). In this experiment a 10-30%(ᵂ/v) linear gradient SDS-polyacrylamide gel was used as it was decided to test if this would result in greater separation of prochymosin from its precursor preprochymosin, compared with that seen on the 9%(ᵂ/v) polyacrylamide gel used above. The preparation of this gel, its electrophoresis and fluorography using 'Enhance' (NEN) was kindly carried out by A.S. Carver. This gel (Fig.IV.2a) did not give better resolution of the preprochymosin(s) and prochymosin(s) than the 9%(ᵂ/v) polyacrylamide gel (30:1.6 cross-linked). As expected there no was detectable difference in size between the *in vitro* and oocyte products expressed from the SP6PChy RNA, encoding methionyl-prochymosin (tracks 3 & 4, respectively). It was shown, however, that the polypeptide translated from this synthetic RNA appeared to display the same mobility as the faster migrating prochymosin protein of the two species expressed in oocytes from the preprochymosin mRNA, and probably had a marginally greater mobility than either of the *in vitro* preprochymosin mRNA proteins, although the comparison of the small differences in relative migration of the polypeptides is difficult between tracks. These observations were confirmed when a similar set of immunoprecipitated samples were electrophoresed on a nongradient 9%(ᵂ/v) polyacrylamide gel (ratio %acrylamide to %bisacrylamide 30:1.6), shown in Fig.IV.2b. This gel demonstrated more clearly that the oocyte SP6PChy translation product (track 5) has a mobility greater than that displayed by either of the primary translation products translated from the mRNA *in vitro* (track 2). The prochymosin expressed from the SP6PChy RNA again migrated to the same position as the lower band of the doublet expressed from preprochymosin mRNA in the oocyte (track 3). This suggests that glycosylation of the translocated prochymosins expressed from the mRNA in oocytes does not account for their smaller than expected difference in the migration on SDS-PAGE to that of the precursor preprochymosins. Furthermore the protein expressed in oocytes from $pTK_2PPChy+$ DNA (track 4) also showed the same migration as the faster migrating 'oocyte' preprochymosin mRNA product and the SP6PChy prochymosin protein. $pTK_2PPChy+$, which contains full-length preprochymosin cDNA insert (see III.E.), is known to express a protein in oocytes which is translocated and secreted (Fig.III.6); the results shown in Fig.IV.2b suggests that the product immunoprecipitated from oocytes injected with $pTK_2PPChy+$ does not contain the signal peptide. At this time the corresponding SP6 RNA of the 'PPChy' cDNA was not available. Further experiments on the expression of

the chymosin cDNA constructs are presented in Chapter V. The implications of the above observations are discussed later (IV.D.).

The experiments described in this section suggest that the mRNA isolated from calf abomasum does encode not one, but two precursors of prochymosin which are both processed following translocation in *Xenopus* oocytes and are secreted. However the presence of the closely migrating doublet of protein bands on the SDS-polyacrylamide gels and the small differences in the mobilities of the various chymosin polypeptides makes it difficult to demonstrate unequivocally that this is the case. It could also be argued that the pattern of radiolabelled polypeptide bands observed on the fluorographed gels is consistent with the hypothesis that only one of the mRNA '*in vitro* products' is cleaved in the oocyte; the primary translation product with the slower mobility on SDS-PAGE is processed to the faster migrating species of the doublet seen in the oocyte fractions.

IV.C. Translocation and Processing *In Vitro* of Preprochymosin Translated from mRNA

The previous section described studies on the expression of preprochymosin mRNA on injection into *Xenopus* oocytes compared with the primary translation products obtained on *in vitro* translation. These suggested that both chymosin polypeptides encoded by the mRNA do contain a signal sequence which is cleaved on translocation across the ER membrane in oocytes. As mentioned earlier (section A this chapter, III.A. and in the Introduction, I.E.) in order to establish the particular step affected by a signal sequence mutation it will be necessary to complement studies using the oocyte as an assay system with experiments using *in vitro* translation-translocation systems, in which translation is carried out in the presence of isolated microsomal vesicles. The recent development of such *in vitro* systems, reconstituted from a variety of components, will allow the affect of signal sequence mutations to be studied in a more detailed way. The experiments presented in this section concern the translation of preprochymosin mRNA *in vitro* and the translocation and processing of the translation products in the presence of dog pancreatic microsomes. As well as fulfilling the objective of characterising the behaviour of wild-type preprochymosin in the *in vitro* systems, these studies have enabled me to examine in further detail the 'doublet' of prochymosin proteins obtained on translation of preprochymosin mRNA, since this had not previously been described by others.

the chymosin cDNA constructs are presented in Chapter V. The implications of the above observations are discussed later (IV.D.).

The experiments described in this section suggest that the mRNA isolated from calf abomasum does encode not one, but two precursors of prochymosin which are both processed following translocation in *Xenopus* oocytes and are secreted. However the presence of the closely migrating doublet of protein bands on the SDS-polyacrylamide gels and the small differences in the mobilities of the various chymosin polypeptides makes it difficult to demonstrate unequivocally that this is the case. It could also be argued that the pattern of radiolabelled polypeptide bands observed on the fluorographed gels is consistent with the hypothesis that only one of the mRNA '*in vitro* products' is cleaved in the oocyte; the primary translation product with the slower mobility on SDS-PAGE is processed to the faster migrating species of the doublet seen in the oocyte fractions.

IV.C. Translocation and Processing *In Vitro* of Preprochymosin Translated from mRNA

The previous section described studies on the expression of preprochymosin mRNA on injection into *Xenopus* oocytes compared with the primary translation products obtained on *in vitro* translation. These suggested that both chymosin polypeptides encoded by the mRNA do contain a signal sequence which is cleaved on translocation across the ER membrane in oocytes. As mentioned earlier (section A this chapter, III.A. and in the Introduction, I.E.) in order to establish the particular step affected by a signal sequence mutation it will be necessary to complement studies using the oocyte as an assay system with experiments using *in vitro* translation-translocation systems, in which translation is carried out in the presence of isolated microsomal vesicles. The recent development of such *in vitro* systems, reconstituted from a variety of components, will allow the affect of signal sequence mutations to be studied in a more detailed way. The experiments presented in this section concern the translation of preprochymosin mRNA *in vitro* and the translocation and processing of the translation products in the presence of dog pancreatic microsomes. As well as fulfilling the objective of characterising the behaviour of wild-type preprochymosin in the *in vitro* systems, these studies have enabled me to examine in further detail the 'doublet' of prochymosin proteins obtained on translation of preprochymosin mRNA, since this had not previously been described by others.

Preprochymosin mRNA was translated in a nuclease treated reticulocyte lysate system (obtained from Amersham) supplemented with [^{35}S]methionine, in the presence and absence of dog pancreatic microsomes (as described in the Materials & Methods II.E.2 & 3). Two other mRNAs, human lactogen and chick oviduct, were translated under the same conditions to provide positive controls for the translocation and processing activity of the microsomes. Following the initial translation an aliquot from the translations containing microsomes was then incubated on ice in the presence of chymotrypsin and trypsin (50μg/ml each) for a further 60min. Another aliquot from the same translation was incubated under the same conditions but in the presence of Triton-X-100 (1%) in addition to the proteases; and a third aliquot was simply incubated untreated for 60min on ice. It was expected that the untreated aliquot from the translation in the presence of microsomes would contain both full-length and the cleaved forms of the secretory proteins encoded by the mRNAs, the relative proportions of each depending on the translocation and processing efficiency of the microsomes. In the aliquots incubated only with proteases those proteins fully translocated within the membranes would be protected from digestion by the enzymes. The addition of the detergent Triton-X-100 will disrupt the membranes allowing protease digestion of proteins both outside and those translocated into the microsomes.

Since it had been established that the nuclease treated reticulocyte lysate gave a low background level of endogenous protein synthesis, samples were not immunoprecipitated prior to electrophoresis on SDS-polyacrylamide gels. Equivalent aliquots from the various treatments of the preprochymosin mRNA translations were separated on a 9%(w/v) polyacrylamide gel (30:0.825 cross-linked), whilst 15% gels (with a cross-linking ratio of 40:0.2) were used for samples from the lactogen and oviduct mRNA translations.

It can be seen from Figure IV.3a that the microsomes were functional in terms of processing of the lactogen and the oviduct lysozyme proteins (tracks 3 & 7). *In vitro* translation of human placental lactogen RNA results in the expression of a precursor protein (Fig.IV.3a, track 2) with an apparent molecular weight M_r=28,000 which, when microsomes are present in the translation, is cleaved to the mature form, apparent molecular weight M_r=22,000 (track 3). Szczesna & Boime(1976) using an ascites cell-free translation system first demonstrated the conversion of the precursor form of lactogen, which they found to have an apparent molecular weight of M_r=25,000, to

Figure IV.3 *In vitro* translation in a reticulocyte lysate of lactogen and oviduct mRNA
Oviduct (Ovd) and lactogen (Lac) mRNAs were translated in the absence and presence of dog pancreatic microsomes in a rabbit reticulocyte lysate cell-free system, as described in the Methods II.E.2 & 3; and assays with no added RNA (tracks marked 'none') and preprochymosin mRNA (see Fig.IV.4) were also included. After translation the assays containing microsomes were divided into 3 aliquots which were treated as follows: one aliquot was treated with proteases (50µg/ml each of chymotrypsin and trypsin), another with Triton-X-100 (1%) and proteases, and the last was left untreated; all three were then incubated on ice for 60min. Samples equivalent to the same volume (2µl) of the translation assay were taken and electrophoresed on SDS-polyacrylamide gels containing 15% polyacrylamide (40:0.2 cross-linked), as described in II.F.2 & 3. The dried fluorographed gels were exposed to X-ray film for (a) 7days (b) 2d. The tracks contain samples of the different translations and treatments as indicated. In (b) track 10 contains an aliquot of a reticulocyte lysate translation primed with synthetic SP6Lys RNA, encoding prelysozyme (described in VI.D). The bands corresponding to [^{35}C]methylated Ovalbumin (Ov) and Lysozyme (Lys) in the molecular weight markers (tracks 1 & 10 in a, and tracks 5 & 11 in b) are indicated.

the mature form (estimated M_r=22,000). The discrepancies in
molecular weights of the precursor proteins is probably a co
the different gel systems used by myself and these worke
relative intensities of the lactogen and prelactogen bands
track 3 it appears that the microsomes have been very
translocating and processing the prelactogen expressed, l
uncleaved precursor is seen. Translation of the oviduct mRN
in the case of prelysozyme, processing by the microsomes app
been slightly less efficient than that seen with prelac
estimated that roughly 50% of the prelysozyme has been
lysozyme, which comigrates with the lysozyme in the radiolab
markers (see Fig.IV.3a track 7, and more clearly in Fig.IV.
The differential efficiency of processing may be due to
quantities of secretory proteins synthesized from the c
compared with the lactogen mRNA translation (compare tracks
6 & 7 in Fig.IV.3a), and hence the microsomes may be satu
oviduct mRNA translation. The major proteins encoded by the
include not only lysozyme but also ovalbumin which is no
translocation, and ovomucoid and conalbumin. It is n
therefore, to use signal processing to assess the e
translocation of ovalbumin by the microsomes. The ovalbum
from the oviduct mRNA migrated slighly faster than the
ovalbumin in the molecular weight markers, this is probab
differences in the extent of glycosylation of the *in vitro*
ovalbumin. There is a less intense band in the marker tracks
position as the *in vitro* ovalbumin; this may correspond to
ovalbumin which has a more similar glycosylation pattern to
oviduct mRNA product. Native ovalbumin carries both N-linked
carbohydrate chains (Glabe *et al*,1980; Hanover & Lennar
et al,1984). It is possible for canine pancreatic microsomes
some N-linked glycosylation of translocated proteins *in vi*
there is no clear indication in the experiment shown in Fig.
ovalbumin synthesized in the presence of microsomes (Fig.IV
has a slower mobility, due to glycosylation, than th
synthesized in the absence of microsomal vesicles (3b track 9

There is also a difference in the effectiveness of
resistance assay for the different translocated proteins tra
the control mRNAs. None of the processed lactogen appears to
to digestion by chymotrypsin and trypsin under the condition
no band was visible at the position corresponding to mature

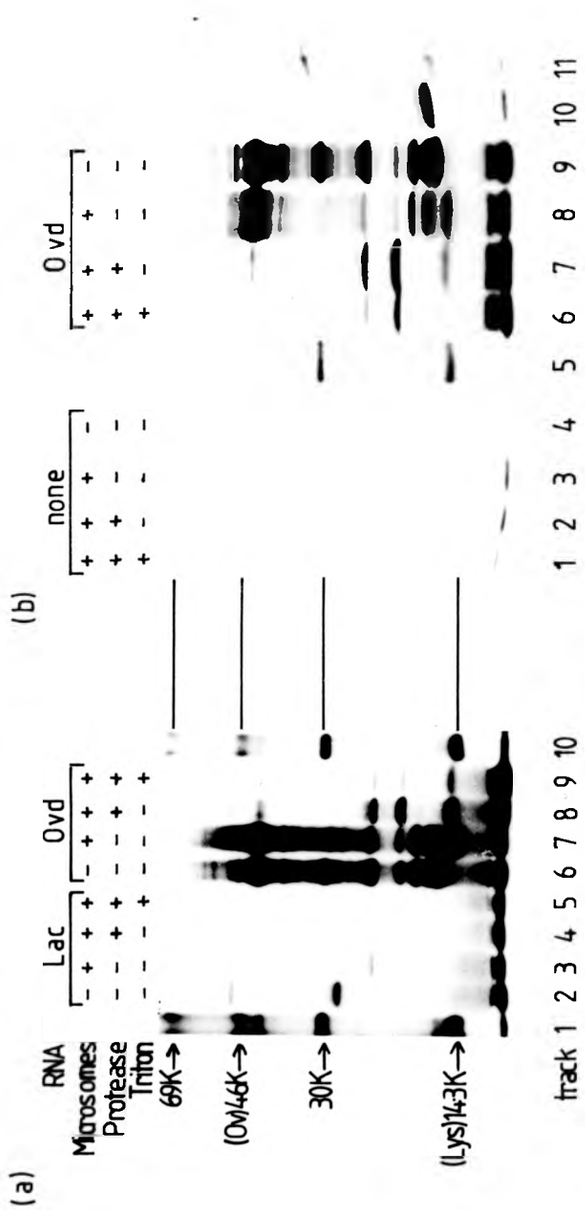


Figure IV.3 *In vitro* translation in a reticulocyte lysate of lactogen and oviduct mRNA
Oviduct (Ovd) and lactogen (Lac) mRNAs were translated in the absence and presence of dog pancreatic microsomes in a rabbit reticulocyte lysate cell-free system, as described in the Methods II.E.2 & 3; and assays with no added RNA (tracks marked 'none') and preprochymosin mRNA (see Fig.IV.4) were also included. After translation the assays containing microsomes were divided into 3 aliquots which were treated as follows; one aliquot was treated with proteases (50µg/ml each of chymotrypsin and trypsin), another with Triton-X-100 (1%) and proteases, and the last was left untreated; all three were then incubated on ice for 60min. Samples equivalent to the same volume (2µl) of the translation assay were taken and electrophoresed on SDS-polyacrylamide gels containing 15% polyacrylamide (40:0.2 cross-linked), as described in II.F.2 & 3. The dried fluorographed gels were exposed to X-ray film for (a) 7days (b) 2d.
The tracks contain samples of the different translations and treatments as indicated. In (b) track 10 contains an aliquot of a reticulocyte lysate translation primed with synthetic SP6Lys RNA, encoding prelysozyme (described in VI.D).
The bands corresponding to [^{14}C]methylated Ovalbumin (Ov) and Lysozyme (Lys) in the molecular weight markers (tracks 1 & 10 in a, and tracks 5 & 11 in b) are indicated.

the mature form (estimated M_r=22,000). The discrepancies in the estimated molecular weights of the precursor proteins is probably a consequence of the different gel systems used by myself and these workers. From the relative intensities of the lactogen and prelactogen bands in Fig.IV.3a, track 3 it appears that the microsomes have been very efficient in translocating and processing the prelactogen expressed, little of the uncleaved precursor is seen. Translation of the oviduct mRNA shows that in the case of prelysozyme, processing by the microsomes appears to have been slightly less efficient than that seen with prelactogen. It is estimated that roughly 50% of the prelysozyme has been cleaved to lysozyme, which comigrates with the lysozyme in the radiolabelled protein markers (see Fig.IV.3a track 7, and more clearly in Fig.IV.3b track 8). The differential efficiency of processing may be due to the larger quantities of secretory proteins synthesized from the oviduct mRNA compared with the lactogen mRNA translation (compare tracks 2 & 3 with 6 & 7 in Fig.IV.3a), and hence the microsomes may be saturated in the oviduct mRNA translation. The major proteins encoded by the oviduct mRNA include not only lysozyme but also ovalbumin which is not cleaved on translocation, and ovomucoid and conalbumin. It is not possible, therefore, to use signal processing to assess the efficiency of translocation of ovalbumin by the microsomes. The ovalbumin expressed from the oviduct mRNA migrated slighly faster than the ^{14}C-labelled ovalbumin in the molecular weight markers, this is probably due to a differences in the extent of glycosylation of the *in vitro* and 'marker' ovalbumin. There is a less intense band in the marker tracks at the same position as the *in vitro* ovalbumin; this may correspond to ^{14}C-labelled ovalbumin which has a more similar glycosylation pattern to that of the oviduct mRNA product. Native ovalbumin carries both N-linked and O-linked carbohydrate chains (Glabe *et al*,1980; Hanover & Lennarz,1981; Kato *et al*,1984). It is possible for canine pancreatic microsomes to carry out some N-linked glycosylation of translocated proteins *in vitro*; however, there is no clear indication in the experiment shown in Fig.IV.3 that the ovalbumin synthesized in the presence of microsomes (Fig.IV.3b, track 8) has a slower mobility, due to glycosylation, than the ovalbumin synthesized in the absence of microsomal vesicles (3b track 9).

There is also a difference in the effectiveness of the protease resistance assay for the different translocated proteins translated from the control mRNAs. None of the processed lactogen appears to be resistant to digestion by chymotrypsin and trypsin under the conditions used, since no band was visible at the position corresponding to mature lactogen in

the aliquot of the lactogen mRNA translation in the presence of microsomes which was treated with the proteases alone (Fig.IV.3a,track 4), not even on a longer exposure of the fluorographed gel to preflashed X-ray film (not shown). This suggests the proteolysis conditions were too severe, unless for some reason the lactogen is not fully translocated within the microsomes. However when an aliquot of the oviduct translation was incubated with the same concentrations of proteases this resulted in digestion of the prelysozyme, whilst the mature form of lysozyme was mostly resistant to protease digestion (see Fig.IV.3b track 7 and compare with track 8). The extent of the protection of cleaved lysozyme cannot be determined as in this case the incubation with Triton-X-100 and proteases did not result in complete digestion of all proteins, and some mature lysozyme resisted digestion even under these conditions (Fig.IV.3b track 6, also Fig.IV.3a track 9). Therefore, in contrast to the lactogen translation, more severe conditions are required for the lysozyme protease resistance assays - perhaps a higher concentration of Triton in the detergent plus protease incubation may suffice. The variation in the sensitivity of different proteins to the same protease assay conditions is demonstrated further when the fate of ovalbumin translated from the oviduct mRNA is considered. As ovalbumin does not contain a cleaved signal sequence the degree to which the protein synthesized in the presence of microsomes is protected from exogenous proteases provides a means of assessing the efficiency of its translocation; on this basis it appears that little of the ovalbumin has been translocated (Fig.IV.3b track 6, also Fig.IV.3a track 7) such that it is protected from digestion by the chymotrypsin and trypsin. As with lysozyme not all the protected protein is then digested when detergent is also included in the incubation (Fig.IV.3b track 7 and Fig.IV.3a track 8).

The translation of the lactogen and oviduct mRNAs, known to encode proteins with cleavable signal peptides, was designed to provide positive controls for translocation, processing and protease protection of segregated proteins in the experimental conditions used. The different results obtained with these controls highlights the necessity for adjustment of the conditions used in these translocation assays according to the particular protein to be studied and the amounts of protein produced in the *in vitro* translation. With these points in mind we can look at the results obtained from the translation of preprochymosin mRNA in the same experiment.

Figure IV.4a is of a short exposure autoradiogram of the fluorographed gel from the electrophoresis of the preprochymosin mRNA

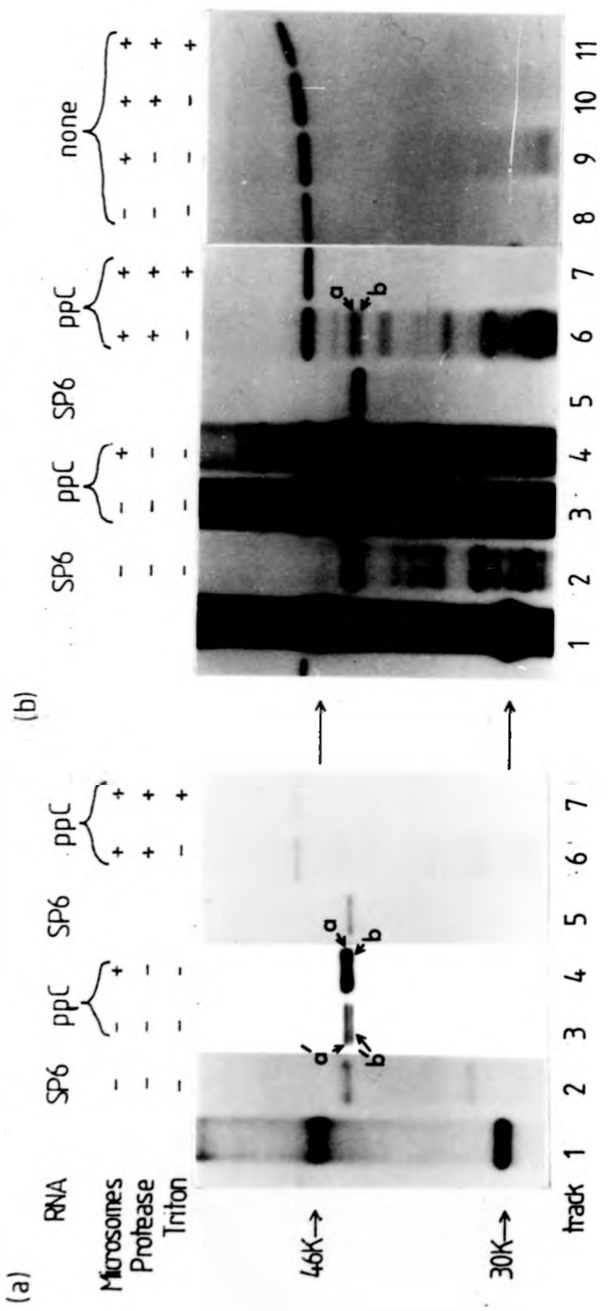


Figure IV.4 Protease resistance of proteins expressed from preprochymosin mRNA translated *in vitro* in the presence of microsomes

Preprochymosin mRNA (ppC) was translated *in vitro*, in the presence and absence of microsomes, in a reticulocyte lysate system in the same experiment as described in Fig.IV.3. (ppC mRNA concentration, 7ng/µl; 1.4µCi/µl [^{35}S]methionine). Aliquots from the translation containing microsomes were treated with protease as described in Fig.IV.3. Equivalents of the translations were analysed by SDS-PAGE on a 9% polyacrylamide (30:0.825 cross-linked) gel. Synthetic RNA SP6PPChy (SP6) (described in V.B) was expressed in a wheat germ cell-free system and in oocytes (O); samples of these translations were also included on the gel, tracks 2 & 5, to provide markers for the full-length and processed products, respectively, of cloned preprochymosin B cDNA. Track 1 is an aliquot of ^{14}C-labelled molecular weight marker proteins. In (b) tracks 8-11 are from the no RNA control reticulocyte lysate translation described in Fig.IV.3. The fluorographed gel was exposed to X-ray film for (a) 20h (b) 14days. In order to give clarity to the protein bands seen on the shorter exposure autoradiogram (a) different parts of the negative were exposed for different durations during the printing.

translation products. This autoradiogram showed that in the presence of microsomes (track 4) a triplet pattern of protein bands is observed, an intense middle band is flanked by bands of lower intensity. This indicates that both the primary translation products expressed from the mRNA in the absence of microsomes, designated and marked a' and b' (track 3, Fig.IV.4a), are partially processed when translation of the mRNA is carried out in the presence of microsomes; with b' being cleaved to the protein with the greatest mobility in track 4 (designated b), whilst the middle band of the triplet consists of the unprocessed b' migrating with the processed form (which I will call a) of the precursor species with the slower mobility a'. It is the unprocessed a' polypeptide which forms the slowest migrating band of the triplet in track 4 of Fig.IV.4a. If this interpretation of the radiolabelled protein bands seen on the autoradiogram of the gel is correct then it is expected that only bands representing the translocated and processed form of prochymosin will be resistant to protease digestion, ie. bands a and b. This was found to be the case when a longer exposure was made of the same gel (Fig.IV.4b). Track 6 is the sample of the translation in the presence of microsomes which was subsequently incubated with chymotrypsin and trypsin; this shows a doublet of protected proteins which migrate to the same postion as band a and b, described above. No radiolabelled proteins are seen in this part of the gel either in the sample treated with Triton and proteases (track 7) or the no RNA control translations (tracks 8-11). The radiolabelled band of approximately 50K which is present in the samples of *in vitro* translations in both Figs.IV.3 & 4, but is particularly prominent in Fig.IV.4b, has been noted to occur in reticulocyte lysate cell-free translation assays even in the absence of exogenous RNAs (Clements,1984). This experiment shows firstly that the proteolysis conditions used enable one to show the protection of prochymosin segregated within microsomes from exogenous proteases, although some digestion of the processed prochymosin has occurred. In other experiments using diffent proteolysis conditions (incubation at 15°C with either 100µg/ml or 50µg/ml chymotrypsin and trypsin) protease resistance of translocated prochymosin could not be detected. This experiment also clearly demonstrates that the preprochymosin mRNA encodes two precursors of prochymosin (apparent molecular weight approx. M$_r$=43,000 and 42,500), each containing a signal sequence which enables their complete translocation into microsomal membranes where processing occurs, producing prochymosins (M$_r$=42,000 and 41,500) which are resistant to digestion by proteases outside the microsomes.

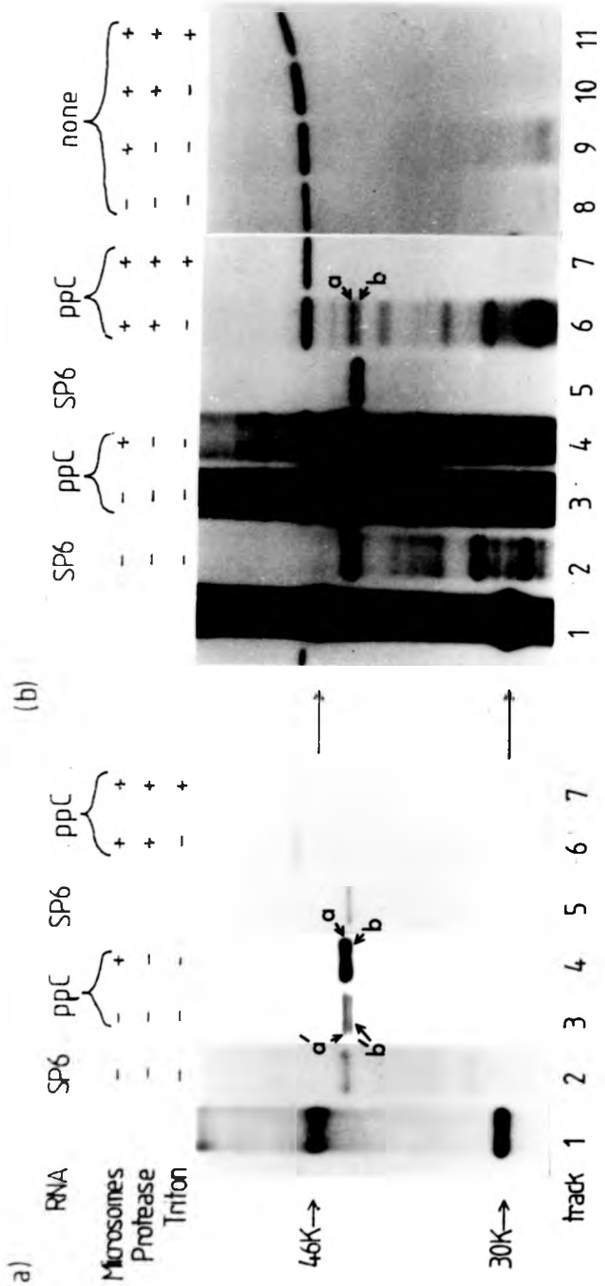


Figure IV.4 Protease resistance of proteins expressed from preprochymosin mRNA translated *in vitro* in the presence of microsomes

Preprochymosin mRNA (ppC) was translated *in vitro*, in the presence and absence of microsomes, in a reticulocyte lysate system in the same experiment as described in Fig.IV.3, (ppC mRNA concentration, 7ng/µl; 1.4µCi/µl [^{35}S]methionine). Aliquots from the translation containing microsomes were treated with protease as described in Fig.IV.3. Equivalents of the translations were analysed by SDS-PAGE on a 9% polyacrylamide (30:0.825 cross-linked) gel. Synthetic RNA SP6PPChy (SP6) (described in V.B) was expressed in a wheat germ cell-free system and in oocytes (O); samples of these translations were also included on the gel, tracks 2 & 5, to provide markers for the full-length and processed products, respectively, of cloned preprochymosin B cDNA. Track 1 is an aliquot of ^{14}C-labelled molecular weight marker proteins. In (b) tracks 8-11 are from the no RNA control reticulocyte lysate translation described in Fig.IV.3. The fluorographed gel was exposed to X-ray film for (a) 20h (b) 14days. In order to give clarity to the protein bands seen on the shorter exposure autoradiogram (a) different parts of the negative were exposed for different durations during the printing.

translation products. This autoradiogram showed that in the presence of microsomes (track 4) a triplet pattern of protein bands is observed, an intense middle band is flanked by bands of lower intensity. This indicates that both the primary translation products expressed from the mRNA in the absence of microsomes, designated and marked a' and b' (track 3, Fig.IV.4a), are partially processed when translation of the mRNA is carried out in the presence of microsomes; with b' being cleaved to the protein with the greatest mobility in track 4 (designated b), whilst the middle band of the triplet consists of the unprocessed b' migrating with the processed form (which I will call a) of the precursor species with the slower mobility a'. It is the unprocessed a' polypeptide which forms the slowest migrating band of the triplet in track 4 of Fig.IV.4a. If this interpretation of the radiolabelled protein bands seen on the autoradiogram of the gel is correct then it is expected that only bands representing the translocated and processed form of prochymosin will be resistant to protease digestion, ie. bands a and b. This was found to be the case when a longer exposure was made of the same gel (Fig.IV.4b). Track 6 is the sample of the translation in the presence of microsomes which was subsequently incubated with chymotrypsin and trypsin; this shows a doublet of protected proteins which migrate to the same postion as band a and b, described above. No radiolabelled proteins are seen in this part of the gel either in the sample treated with Triton and proteases (track 7) or the no RNA control translations (tracks 8-11). The radiolabelled band of approximately 50K which is present in the samples of *in vitro* translations in both Figs.IV.3 & 4, but is particularly prominent in Fig.IV.4b, has been noted to occur in reticulocyte lysate cell-free translation assays even in the absence of exogenous RNAs (Clements,1984). This experiment shows firstly that the proteolysis conditions used enable one to show the protection of prochymosin segregated within microsomes from exogenous proteases, although some digestion of the processed prochymosin has occurred. In other experiments using diffent proteolysis conditions (incubation at 15°C with either 100µg/ml or 50µg/ml chymotrypsin and trypsin) protease resistance of translocated prochymosin could not be detected. This experiment also clearly demonstrates that the preprochymosin mRNA encodes two precursors of prochymosin (apparent molecular weight approx. M_r=43,000 and 42,500), each containing a signal sequence which enables their complete translocation into microsomal membranes where processing occurs, producing prochymosins (M_r=42,000 and 41,500) which are resistant to digestion by proteases outside the microsomes.

Figure IV.4 also shows that the polypeptide (seen in track 2) produced by the *in vitro* translation of a synthetic RNA, SP6PPChy, which encodes the preprochymosin cDNA gene cloned by Harris *et al*(1982) (see III.E. and V.B), has the same mobility as the protein band b', one of the precursors encoded by preprochymosin mRNA (track 3). Similarly the protein expressed after injection of SP6PPChy RNA into oocytes (track 5) migrates to the same position on the gel as the lower band, b, of the translocated, processed prochymosins expressed from the mRNA (Fig.IVb track 6 and Fig.IVa track 4). This, together with the results shown in Fig.IV.2, suggests that the preprochymosin gene isolated on the 'PPChy' insert by Harris *et al*, represents the cDNA clone of the species in the natural mRNA which encodes a preprochymosin with a mobility on SDS-PAGE consistent with M_r=42,500; this preprochymosin is processed on translocation across membranes *in vitro* and *in vivo* to a prochymosin with an apparent molecular weight of Mr=41,500, judged by its migration on SDS-PAGE. This will be discussed further in D.2.

IV.D. Discussion

IV.D.1 **Preprochymosin is translocated and cleaved both in oocytes and *in vitro***

The results presented in this chapter show that preprochymosin, the primary translation product encoded by prochymosin mRNA, is cleaved and translocated both in *Xenopus* oocytes and *in vitro* systems supplemented with microsomes. Thus two types of assay system are characterised for use in the analysis of signal sequence mutants of preprochymosin: the oocyte is useful because of its high efficiency of translocation of foreign secretory proteins and also as an assay for secretion; while the *in vitro* system provides a means of dissecting the translocation process, as described earlier (I.E. and III.A).

As discussed fully below (D.2) the preprochymosin mRNA encoded not one but two precursors of prochymosin both of which are translocated and processed. These showed close migration on SDS-PAGE which caused difficulties in interpreting the *in vitro* translocation data, especially since the difference in mobilities of the full-length preprochymosins and the processed prochymosins was less than that expected from the amino acid sequence of the cleaved signal peptide. The necessity to use cloned cDNA encoding preprochymosin for the proposed signal sequence mutagenesis studies will remove part of these complications since it was found the cloned gene resulted in the expression of a single polypeptide

in vitro and in the oocyte. Further work on the expression of the preprochymosin cDNA is presented in the next chapter (V).

IV.D.2 **What are the two preprochymosin proteins encoded by the mRNA?**

It has been suggested that the translation product of the gene cloned by Harris *et al*(1982) accounts for one (designated b') of the prochymosin precursors encoded by the preprochymosin mRNA; if this is the case then what is the identity of the other (designated a')? The possibility that a' is a post-translational modification of b' which can take place both *in vitro* and in the oocyte can be excluded, since if this were the case then a protein doublet would also be seen on translation of the SP6PPChy RNA (transcribed from the cloned preprochymosin gene) in the same systems. An alternative is that a' is the product of a second preprochymosin gene represented in the mRNA preparation. As described in the Introduction (I.D.), to date two groups have cloned full-length preprochymosin genes from cDNA made from polyA RNA extracted from the mucosal layer of the abomasum of unweaned calves. The gene characterised by Harris *et al*(1982) was found to encode for a protein with an amino acid sequence corresponding to that of prochymosin B, as determined by Foltmann *et al*(1977). Moir *et al*(1982) cloned a gene with a slightly different DNA sequence resulting in a translation product with two amino acids different to prochymosin B, one of which corresponded with the known difference between prochymosin B and the partial amino acid sequence of prochymosin A (Foltmann *et al*,1979); hence it was suggested that this cDNA most probably represents the prochymosin A gene. The work by Moir *et al*(1982) also indicated that there is a single bovine chymosin locus, and they demonstrated that the A and B cDNAs are alleles of this locus. If this is the case then the simplest explanation of the results described in this chapter is that the mRNA used in the experiments was isolated from a calf heterozygotic at the chymosin locus, with both the A and B alleles; or from calves of different genotypes. Hence the slower mobility preprochymosin precursor, a', (seen on SDS-PAGE of the *in vitro* translation products of the mRNA) will be preprochymosin A. The cDNA sequence data showed that prochymosin A (encoded by the cDNA isolated by Moir *et al*) only differs from prochymosin B (Harris *et al* cDNA) at residues 214 and 286, which in prochymosin A are both aspartate whereas in prochymosin B they are asparagine and glycine respectively (see Table I.1 in the Introduction). It must therefore be concluded that despite the use of denaturing SDS-polyacrylamide gels in my experiments the two extra charged residues in preprochymosin A cause it to have a different

mobility on electrophoresis to the preprochymosin B gene product although their molecular weights are virtually identical.

Whilst this hypothesis of the nature of the preprochymosin 'doublet' encoded by the mRNA is the simplest consistent with the results I obtained and those of others, it can only be put forward tentatively at this stage and other interpretations of the data cannot be excluded. It is possible either that the a' polypeptide is the translation product of another allelic form of the single chymosin genetic locus, or that it is encoded by a chymosin gene non-allelic to A and B which is sufficiently divergent in its DNA sequence not to be detected by the preprochymosin A hybridization probes used by Moir *et al*(1982) in the genomic blots which detected the single chymosin locus. Initially it would be of interest to compare the migration on SDS-polyacrylamide gels of the *in vitro* and oocyte translation products of the putative preprochymosin A gene cloned by Moir's group with those of the preprochymosin B cDNA cloned by Harris *et al*(1982) and the preprochymosin mRNA preparation used in the experiments here. If the primary translation product of the preprochymosin A cDNA migrates with the precursor a' encoded by the mRNA, and there is also comigration of the corresponding processed and translocated prochymosins expressed in the oocyte, then this would strongly support the suggestion that the mRNA contains RNAs derived from the A and B preprochymosin genes. Should this equivalence of mobilities not be observed then this hypothesis cannot be valid and alternative explanations can be explored. However, only microsequencing of the two precursors, a' and b', encoded by the mRNA will give their absolute identity as preprochymosin A or B, or as other uncharacterized chymosins, as yet not cloned. Similarly only by determining the N-terminal sequences of the cleaved proteins secreted by oocytes and those translocated within microsomes *in vitro*, can one conclude that correct processing of the signal peptide has occurred in these systems.

A further complication to the interpretation arises from the noted heterogeneity which exists at the protein level *in vivo*, as described in the Introduction (I.D.1). Up to four different forms of prochymosin and chymosin isolated from calf stomachs have been distinguished by electrophoretic and chromatographic analysis (Foltmann,1970; Asato & Rand,1971; Asato & Rand,1972; Donnelly *et al*,1984), but the precise nature and relationship of all these species is not clear. In this context it is noted that examination of the amino acid sequence of the cloned preprochymosin cDNAs (Harris *et al*,1982; Moir *et al*,1982) reveals two potential sites for N-linked glycosylation, both Asn-X-Ser (see

Introduction I.A.2 & 3). It would be interesting to compare the translation products of the mRNA extracted from individual calf abomasums with the protein species isolated from the same animal. It is also noted that Donnelly *et al*(1984), who analysed the chymosin proteins from individual calves, comment that the two chymosins which they identified by their separation on DEAE-cellulose chromatography also displayed different electrophoretic mobilities on 6M-urea, 6%(w/v) polyacrylamide gels but they were not resolved on SDS-polyacrylamide gels, although they did not specify if this was also using a gel containing 6%(w/v) polyacrylamide. Therefore it is possible that the SDS-polyacrylamide gel system used in my translation studies may not have revealed the true complexity of chymosin proteins expressed from the preprochymosin mRNA, or indeed from the cloned chymosin genes.

IV.E. Summary

Messenger (polyA) RNA extracted from calf abomasum was translated *in vitro* in the absence and presence of pancreatic microsomes, and the products were analysed by electrophoresis on SDS-polyacrylamide gels (9%(w/v)polyacrylamide, with a ratio of acrylamide to bisacrylamide of 30:1.6 or 30:0.825). The major primary translation products resolved on the gels was a doublet of protein bands, M_r=43,000 and M_r=42,500, and these polypeptides were immunoprecipitated by antibodies raised against calf prochymosin. Both these proteins contain a signal sequence which directs their translocation within microsomes *in vitro*; translocation being accompanied by processing of the precursors to proteins of M_r=42,000 and 41,500. Prochymosin-specific peptides of the same size as those proteins translocated *in vitro* were also immunoprecipitated from oocytes injected with the preprochymosin mRNA. The two prochymosins expressed in oocytes were secreted into the medium surrounding the injected oocytes. One of the preprochymosins produced from translation of the mRNA *in vitro* displays the same migration on SDS-PAGE as the *in vitro* translation product of the cloned preprochymosin B gene; likewise comigration is observed of the corresponding processed prochymosin species secreted by injected oocytes. It is suggested that this preprochymosin mRNA (a gift from Celltech Ltd.) was isolated from a calf with a heterozygotic chymosin genotype, and encodes two prochymosin precursors representing the primary translation products of two allelic forms of the chymosin genetic locus, one of which is probably the preprochymosin B gene whilst the other may be preprochymosin A. However one cannot exclude the possibilities that the

two proteins, giving rise to the doublet of bands seen on gel electrophoresis, either represent other allelic forms of the same chymosin locus or are expressed from non-allelic chymosin genes.

V. *IN VIVO* & *IN VITRO* EXPRESSION OF SYNTHETIC RNAS FOR PREPROCHYMOSIN, PROCHYMOSIN AND CHYMOSIN

V.A. Introduction

At this stage we obtained in the laboratory the SP6 vectors developed by Melton's group (Harvard, USA) and the protocol for generating *in vitro* transcripts initiated from the bacteriophage SP6 promoter contained in these vectors (Melton *et al*,1984). By cloning cDNAs into these SP6 vectors it was then possible to transcribe large quantities of specific RNAs which had been shown to be biologically active and were translated following microinjection into *Xenopus* oocytes and also *in vitro* (Krieg & Melton,1984). It was decided to examine the expression of synthetic RNAs transcribed from wild-type preprochymosin cDNA (PPChy) and also from the methionyl-prochymosin (PChy) and methionyl-chymosin (Chy) constructs derived from this cDNA (see III.E), both in *Xenopus* oocytes and *in vitro*.

As noted in Chapter III the injection of pTK_2PChy+ into *Xenopus* oocytes only ever gave poor expression of prochymosin, and no chymosin-specific polypeptides were ever detected following injection of pTK_2Chy+. Would the cytoplasmic injection of synthetic RNAs result in better expression of the preprochymosin derivatives in oocytes, compared with their expression from cDNA injected into the nucleus? Also would the localisation of proteins expressed in the oocyte from cytoplasmically injected chymosin SP6 RNAs be the same as that seen following the nuclear injection of the corresponding pTK_2 plasmid DNAs, described in Chapter III? In the context of the proposed signal sequence mutagenesis studies (outlined in I.E) the question is whether miscompartmented derivatives of preprochymosin are sufficiently stable to be detected when expressed in *Xenopus* oocytes. It would therefore be reassuring to demonstrate that the signal-minus prochymosin polypeptide can be clearly detected in the cytosol of the oocyte; although it is recognised that mutant preprochymosins which are not translocated might contain a point mutation in the signal peptide which specifically destabilises the cytoplasmically localised preprochymosin protein. The SP6 RNAs can, in addition, be translated *in vitro* providing full-length primary translation products for comparison with the oocyte products. The SP6 RNAs also enable *in vitro* translocation studies to be carried out on the translation products of specific cDNA constructs. It is intended to use *in vitro* synthesized 'SP6 RNAs' to provide a means of testing signal sequence

mutants of preprochymosin for their ability to function in the process of translocation *in vitro*.

V.B. *In Vitro* Synthesis of SP6 RNAs

The *HindIII* inserts were excised from pTK$_2$PPChy+, pTK$_2$PChy+ and pTK$_2$Chy+ (described in III.E) and cloned into the pSP$_{64}$ vector (Melton *et al*,1984). Transformants were selected which contained the cDNA insert in the 'plus' orientation relative to the SP6 promoter (see Fig.III.2b); the plasmids were designated pSP$_{64}$PPChy+, pSP$_{64}$PChy+ and pSP$_{64}$Chy+ according to the insert they contained. *In vitro* transcription from these constructs would generate 'sense' RNA corresponding to the coding strand of the cDNA insert. The cloning was carried out by Dave Jackson. To provide a template for *in vitro* transcription the pSP$_{64}$ constructs were linearized at the unique *XbaI* site in the polylinker region downstream of the insert; the RNAs synthesized from these templates are referred to as SP6PPChy, SP6PChy and SP6Chy.

All eukaryotic mRNAs have a cap structure which consists of a 7 methylguanosine base joined by a triphosphate bridge to their 5' end. In addition the first (and sometimes second) nucleotide of the RNA is methylated (Banerjee,1980). Another feature of natural eukaryotic mRNAs is the polyadenylation at the 3' end of the message. Furuichi *et al*(1977) had shown that the 5' cap increased the stability of natural mRNAs in *Xenopus* oocytes. The work carried out by Krieg & Melton(1984) demonstrated that for the synthetic SP6 mRNA a 5' cap was essential for translation in oocytes, but that there was no absolute requirement for a 3' polyA tail for protein synthesis in oocytes. Kreig & Melton had added the 5' terminal cap to the SP6 RNAs after transcription using the enzyme guanyltransferase from vaccinia virus. However it is also possible to produce capped RNAs by including a 'capping dinucleotide' (m^7(5')Gppp(5')N or m^7(5')Gppp(5')Nm) in the transcription reaction; since the SP6 transcripts start with a guanosine residue the appropriate dinucleotides are m^7(5')Gppp(5')G and m^7(5')Gppp(5')Gm (abbreviated to m^{7}GG and m^{7}GGm respectively). At a 10× molar excess to GTP these dinucleotides are incorporated efficiently at the 5' terminus of the transcripts (Konarska *et al*,1984; Pelletier & Sonenberg,1985). Studies were initiated in this laboratory to determine if SP6 transcripts synthesized in the presence of capping dinucleotides were also translated efficiently in oocytes and *in vitro*. These preliminary studies, carried out by myself together with several others in the laboratory (Douglas Drummond, Dave Jackson, Alan Colman), examined the protein expression of 'dinucleotide capped' SP6 RNAs

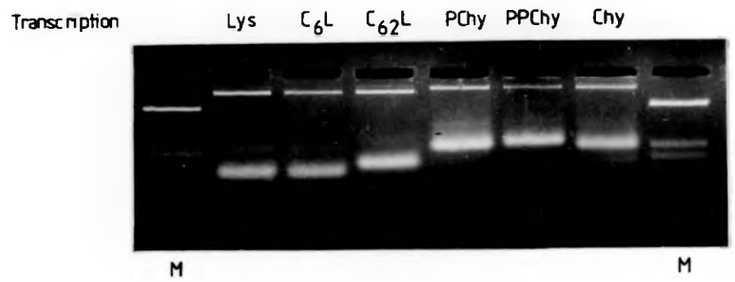
mutants of preprochymosin for their ability to function in the process of translocation *in vitro*.

V.B. *In Vitro* Synthesis of SP6 RNAs

The *HindIII* inserts were excised from pTK$_2$PPChy+, pTK$_2$PChy+ and pTK$_2$Chy+ (described in III.E) and cloned into the pSP$_{64}$ vector (Melton *et al*,1984). Transformants were selected which contained the cDNA insert in the 'plus' orientation relative to the SP6 promoter (see Fig.III.2b); the plasmids were designated pSP$_{64}$PPChy+, pSP$_{64}$PChy+ and pSP$_{64}$Chy+ according to the insert they contained. *In vitro* transcription from these constructs would generate 'sense' RNA corresponding to the coding strand of the cDNA insert. The cloning was carried out by Dave Jackson. To provide a template for *in vitro* transcription the pSP$_{64}$ constructs were linearized at the unique *XbaI* site in the polylinker region downstream of the insert; the RNAs synthesized from these templates are referred to as SP6PPChy, SP6PChy and SP6Chy.

All eukaryotic mRNAs have a cap structure which consists of a 7 methylguanosine base joined by a triphosphate bridge to their 5' end. In addition the first (and sometimes second) nucleotide of the RNA is methylated (Banerjee,1980). Another feature of natural eukaryotic mRNAs is the polyadenylation at the 3' end of the message. Furuichi *et al*(1977) had shown that the 5' cap increased the stability of natural mRNAs in *Xenopus* oocytes. The work carried out by Krieg & Melton(1984) demonstrated that for the synthetic SP6 mRNA a 5' cap was essential for translation in oocytes, but that there was no absolute requirement for a 3' polyA tail for protein synthesis in oocytes. Kreig & Melton had added the 5' terminal cap to the SP6 RNAs after transcription using the enzyme guanyltransferase from vaccinia virus. However it is also possible to produce capped RNAs by including a 'capping dinucleotide' (m^7(5')Gppp(5')N or m^7(5')Gppp(5')Nm) in the transcription reaction; since the SP6 transcripts start with a guanosine residue the appropriate dinucleotides are m^7(5')Gppp(5')G and m^7(5')Gppp(5')Gm (abbreviated to m^{7}GG and m^{7}GGm respectively). At a 10× molar excess to GTP these dinucleotides are incorporated efficiently at the 5' terminus of the transcripts (Konarska *et al*,1984; Pelletier & Sonenberg,1985). Studies were initiated in this laboratory to determine if SP6 transcripts synthesized in the presence of capping dinucleotides were also translated efficiently in oocytes and *in vitro*. These preliminary studies, carried out by myself together with several others in the laboratory (Douglas Drummond, Dave Jackson, Alan Colman), examined the protein expression of 'dinucleotide capped' SP6 RNAs

(a)



(b)

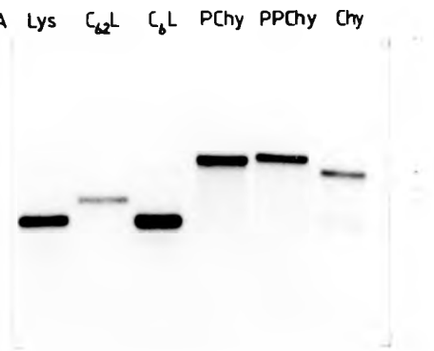
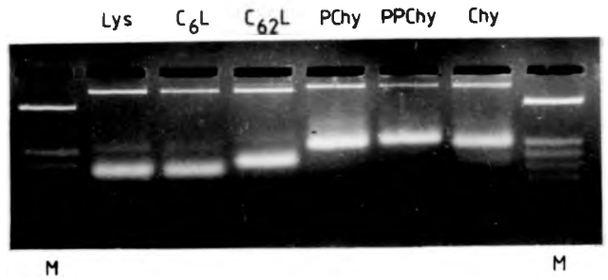


Figure V.1 *In vitro* synthesis of capped RNAs encoding the secretory proteins preprochymosin and prelysozyme and constructs derived from them
The linearised plasmids $pSP_{64}PPChy+$, $pSP_{64}PChy+$ and $pSP_{64}Chy+$ (described in V.B), and $pSP_{64}Lys+$, $pSP_{64}C_{62}L+$, $pSP_{64}C_6L+$ (described in VI.D), and were transcribed *in vitro*, at a concentration of 50µg/ml, in the presence of m^7G6 capping dinucleotide and 300units/ml SP6 RNA polymerase, as described in the Methods (II.D). At the end of the incubation period $^1/_{10}$ of each transcription reaction was analysed by electrophoresis on a 2% agarose/TEA gel which was stained with ethidium bromide, shown in (a) (see II.B.8 & II.D.2). The pBR322 *HinfI* digest also included on this gel (M) contains DNA fragments of the following sizes in base pairs: 1631, 516, 506, 396, 344, 298, 221, 220, 154, 75. The synthetic RNA was recovered from the transcription reactions as described in II.D. (b) shows an aliquot ($^1/_{10}$) of the recovered SP6 RNAs electrophoresed on a denaturing formaldehyde/MOPS agarose gel, which then was dried and exposed to X-ray sensitive film for 15min.

in oocytes, and also in the reticulocyte lysate and wheat germ *in vitro* translation systems. These experiments compared uncapped and the two differently capped SP6 RNAs, and showed that the use of the m^7GG capping dinucleotide resulted in a synthetic RNA which gave the most efficient translation in oocytes, and this m^7GG-capped RNA was also translated well *in vitro*. These initial findings were subsequently confirmed and extended in a detailed study made by Douglas Drummond which examined the effect of adding a 5' cap or 3' polyA tail on the translation, and also the stability and movement, of several SP6 transcripts following injection into *Xenopus* oocytes (see Drummond *et al*,1985). As a result of these early experiments with the SP6 system it was decided to generate capped SP6 RNAs encoding preprochymosin, prochymosin and chymosin by synthesizing the transcripts in the presence of $m^7(5')Gppp(5')G$, and then examine the expression of the chymosin proteins from these synthetic mRNAs in the oocyte and *in vitro*.

Capped SP6PPChy, SP6PChy and SP6Chy RNAs were transcribed from the linearized pSP_{64} chymosin plasmids, described above, in the presence of m^7GG capping dinucleotide (as described in the Methods, II.D.). When samples representing $^1/_{10}$ of each transcription were electrophoresed on an agarose/TEA gel (see II.B.8 and D.2) these gave strong, discrete ethidium stained bands (see Fig. V.1a), which indicated good yields of the RNAs. The incorporation of [^{32}-P]GTP was estimated by determining the material binding to DE81 paper, as described in II.D.2. This showed that for each transcription reaction >80% of the theoretical maximum yield of RNA had been obtained in this experiment, and the calculated yields of the RNAs were - SP6PPChy, 2.8µg; SP6PChy, 2.9µg; SP6Chy, 3µg. As a further check samples of the recovered SP6 RNAs were electrophoresed on a denaturing formaldehyde-agarose/MOPS gel which was then dried and autoradiographed; this was kindly carried out by D. Drummond. As Figure V.1b shows most of the radiolabelled GTP is incorporated into a single band in each of the transcription reactions, and the relative sizes of these are as expected from the lengths of the three inserts; this indicated that predominantly full-length transcripts had been synthesized. Although in the SP6Chy transcription (Fig.V.1b, track 6) there is a smaller, less intense radiolabelled band which could be a specific prematurely terminated transcript.

(a)



(b)

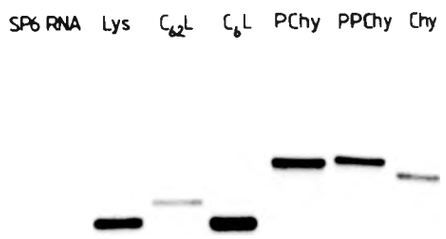


Figure V.1 *In vitro* synthesis of capped RNAs encoding the secretory proteins preprochymosin and prelysozyme and constructs derived from them
The linearised plasmids $pSP_{64}PPChy+$, $pSP_{64}PChy+$ and $pSP_{64}Chy+$ (described in V.B), and $pSP_{64}Lys+$, $pSP_{64}C_{62}L+$, $pSP_{64}C_6L+$ (described in VI.D), and were transcribed *in vitro*, at a concentration of 50µg/ml, in the presence of m^7GG capping dinucleotide and 300units/ml SP6 RNA polymerase, as described in the Methods (II.D). At the end of the incubation period $^1/_{10}$ of each transcription reaction was analysed by electrophoresis on a 2% agarose/TEA gel which was stained with ethidium bromide, shown in (a) (see II.B.8 & II.D.2). The pBR322 *HinfI* digest also included on this gel (M) contains DNA fragments of the following sizes in base pairs: 1631, 516, 506, 396, 344, 298, 221, 220, 154, 75. The synthetic RNA was recovered from the transcription reactions as described in II.D. (b) shows an aliquot ($^1/_{10}$) of the recovered SP6 RNAs electrophoresed on a denaturing formaldehyde/MOPS agarose gel, which then was dried and exposed to X-ray sensitive film for 15min.

in oocytes, and also in the reticulocyte lysate and wheat germ *in vitro* translation systems. These experiments compared uncapped and the two differently capped SP6 RNAs, and showed that the use of the m^7GG capping dinucleotide resulted in a synthetic RNA which gave the most efficient translation in oocytes, and this m^7GG-capped RNA was also translated well *in vitro*. These initial findings were subsequently confirmed and extended in a detailed study made by Douglas Drummond which examined the effect of adding a 5' cap or 3' polyA tail on the translation, and also the stability and movement, of several SP6 transcripts following injection into *Xenopus* oocytes (see Drummond *et al*,1985). As a result of these early experiments with the SP6 system it was decided to generate capped SP6 RNAs encoding preprochymosin, prochymosin and chymosin by synthesizing the transcripts in the presence of $m^7(5')Gppp(5')G$, and then examine the expression of the chymosin proteins from these synthetic mRNAs in the oocyte and *in vitro*.

Capped SP6PPChy, SP6PChy and SP6Chy RNAs were transcribed from the linearized pSP_{64} chymosin plasmids, described above, in the presence of m^7GG capping dinucleotide (as described in the Methods, II.D.). When samples representing $^1/_{10}$ of each transcription were electrophoresed on an agarose/TEA gel (see II.B.8 and D.2) these gave strong, discrete ethidium stained bands (see Fig. V.1a), which indicated good yields of the RNAs. The incorporation of [^{32}-P]GTP was estimated by determining the material binding to DE81 paper, as described in II.D.2. This showed that for each transcription reaction >80% of the theoretical maximum yield of RNA had been obtained in this experiment, and the calculated yields of the RNAs were - SP6PPChy, 2.8µg; SP6PChy, 2.9µg; SP6Chy, 3µg. As a further check samples of the recovered SP6 RNAs were electrophoresed on a denaturing formaldehyde-agarose/MOPS gel which was then dried and autoradiographed; this was kindly carried out by D. Drummond. As Figure V.1b shows most of the radiolabelled GTP is incorporated into a single band in each of the transcription reactions, and the relative sizes of these are as expected from the lengths of the three inserts; this indicated that predominantly full-length transcripts had been synthesized. Although in the SP6Chy transcription (Fig.V.1b, track 6) there is a smaller, less intense radiolabelled band which could be a specific prematurely terminated transcript.

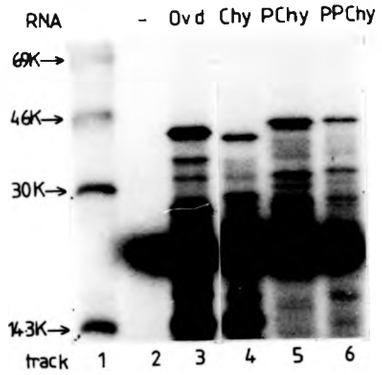


Figure V.2 *In vitro* translation of SP6 RNAs encoding preprochymosin, prochymosin and chymosin
m⁷GG-capped SP6PPChy, SP6PChy and SP6Chy RNAs were synthesized *in vitro* as described in Fig.V.1, and translated in a wheat germ cell-free system (at 8ng.µl RNA) as described in II.E.2. Aliquots (2µl) of the *in vitro* translations were separated by SDS-PAGE on a 12.5% polyacrylamide gel (cross linking ratio 30:0.825) (see II.F.2). Fluorography of the gel was carried out using 'Amplify' (Amersham), and the dried gel was then exposed to preflashed X-ray sensitive film for 30h. The samples shown are from translations of the following RNAs: track 2, no exogenous RNA; track 3 oviduct mRNA (Ovd); track 4, SP6Chy; track 5, SP6PChy; and track 6, SP6PPChy. Track 1 contains radiolabelled marker proteins (II.F.2).

V.C. Translation *In Vitro* and in *Xenopus* Oocytes of SP6 RNAs Encoding Preprochymosin, Prochymosin and Chymosin

The m⁷GG-capped SP6Chy, SP6PChy and SP6PPChy were translated in a wheat germ cell-free system (as described in II.E.2) and the translation products were separated by SDS-PAGE without immunoprecipitation. It has been shown that the initiation methionine is removed from polypeptides synthesized *in vitro* in cell-free systems (Housman *et al*,1970). Therefore, calculated from amino acid composition the expected sizes of the SP6PPChy, SP6PChy and SP6Chy translation products are 42.4K, 40.8K and 35.7K respectively. Judged by migration on the gel shown in Figure V.2 the approximate size of the largest polypeptide band seen in the *in vitro* translation of each of the synthetic RNAs was as follows; SP6PPChy 42.5K (track 6), SP6PChy 42K (track 5), SP6Chy 39.5K (track 4), but a number of other smaller bands were also seen. The diffuse smear across the gel around 20K is thought to be due to the use of the commercial fluorography agent 'Amplify'; it was only ever seen when a commercial single step fluorography agent was used with samples of *in vitro* translations which were electrophoresed without immunoprecipitation. It seems likely that the largest translation products of the chymosin SP6 RNAs are full-length preprochymosin, prochymosin and chymosin; the smaller species may be products of premature termination. It has been noted that a disadvantage of the wheat germ *in vitro* translation system is the tendancy for incomplete translation products to be produced (Clemens,1984). This experiment demonstrated that each of the chymosin SP6 RNAs did produce a translation product *in vitro*, including the Chy construct which did not give a detectable protein product in the oocyte from the injected cDNA.

Xenopus oocytes were injected with capped SP6PPChy, SP6PChy and SP6Chy, then cultured and fractionated as described previously (II.H.). Aliquots from the *in vitro* translations described above and samples of the injected oocyte fractions were immunoprecipitated with antiprochymosin antibodies (αpC_{PR}, see II.G), and the immunoprecipitated proteins were analysed by SDS-PAGE (Fig.V.3). The apparent molecular weight of the major polypeptide precipitated by αpC_{PR} from the wheat germ translations was 43.5K from SP6PPChy (track 5), 43K from SP6PChy (track 6) and 39.5K from SP6Chy (track 7), but some other smaller polypeptides were also precipitated. In contrast a single polypeptide species was precipitated by αpC_{PR} from the oocytes injected with the chymosin SP6 RNAs, except in the secreted (S) fraction where several other bands were seen, but these were also present in the immunoprecipitated S sample of uninjected oocytes (track 10). The injection of SP6PPChy resulted in the

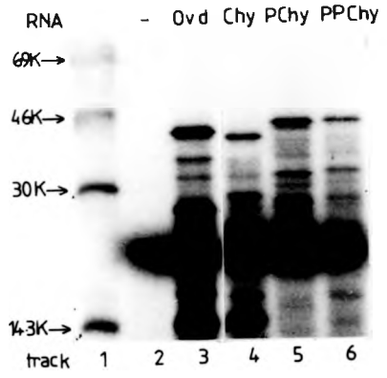


Figure V.2 *In vitro* translation of SP6 RNAs encoding preprochymosin, prochymosin and chymosin

m^{7}GG-capped SP6PPChy, SP6PChy and SP6Chy RNAs were synthesized *in vitro* as described in Fig.V.1, and translated in a wheat germ cell-free system (at 8ng.µl RNA) as described in II.E.2. Aliquots (2µl) of the *in vitro* translations were separated by SDS-PAGE on a 12.5% polyacrylamide gel (cross linking ratio 30:0.825) (see II.F.2). Fluorography of the gel was carried out using 'Amplify' (Amersham), and the dried gel was then exposed to preflashed X-ray sensitive film for 30h. The samples shown are from translations of the following RNAs: track 2, no exogenous RNA; track 3 oviduct mRNA (Ovd); track 4, SP6Chy; track 5, SP6PChy; and track 6, SP6PPChy. Track 1 contains radiolabelled marker proteins (II.F.2).

V.C. Translation *In Vitro* and in *Xenopus* Oocytes of SP6 RNAs Encoding Preprochymosin, Prochymosin and Chymosin

The m^{7}GG-capped SP6Chy, SP6PChy and SP6PPChy were translated in a wheat germ cell-free system (as described in II.E.2) and the translation products were separated by SDS-PAGE without immunoprecipitation. It has been shown that the initiation methionine is removed from polypeptides synthesized *in vitro* in cell-free systems (Housman *et al*,1970). Therefore, calculated from amino acid composition the expected sizes of the SP6PPChy, SP6PChy and SP6Chy translation products are 42.4K, 40.8K and 35.7K respectively. Judged by migration on the gel shown in Figure V.2 the approximate size of the largest polypeptide band seen in the *in vitro* translation of each of the synthetic RNAs was as follows; SP6PPChy 42.5K (track 6), SP6PChy 42K (track 5), SP6Chy 39.5K (track 4), but a number of other smaller bands were also seen. The diffuse smear across the gel around 20K is thought to be due to the use of the commercial fluorography agent 'Amplify'; it was only ever seen when a commercial single step fluorography agent was used with samples of *in vitro* translations which were electrophoresed without immunoprecipitation. It seems likely that the largest translation products of the chymosin SP6 RNAs are full-length preprochymosin, prochymosin and chymosin; the smaller species may be products of premature termination. It has been noted that a disadvantage of the wheat germ *in vitro* translation system is the tendancy for incomplete translation products to be produced (Clemens,1984). This experiment demonstrated that each of the chymosin SP6 RNAs did produce a translation product *in vitro*, including the Chy construct which did not give a detectable protein product in the oocyte from the injected cDNA.

Xenopus oocytes were injected with capped SP6PPChy, SP6PChy and SP6Chy, then cultured and fractionated as described previously (II.H.). Aliquots from the *in vitro* translations described above and samples of the injected oocyte fractions were immunoprecipitated with antiprochymosin antibodies (αpC_{ms}, see II.G), and the immunoprecipitated proteins were analysed by SDS-PAGE (Fig.V.3). The apparent molecular weight of the major polypeptide precipitated by αpC_{ms} from the wheat germ translations was 43.5K from SP6PPChy (track 5), 43K from SP6PChy (track 6) and 39.5K from SP6Chy (track 7), but some other smaller polypeptides were also precipitated. In contrast a single polypeptide species was precipitated by αpC_{ms} from the oocytes injected with the chymosin SP6 RNAs, except in the secreted (S) fraction where several other bands were seen, but these were also present in the immunoprecipitated S sample of uninjected oocytes (track 10). The injection of SP6PPChy resulted in the

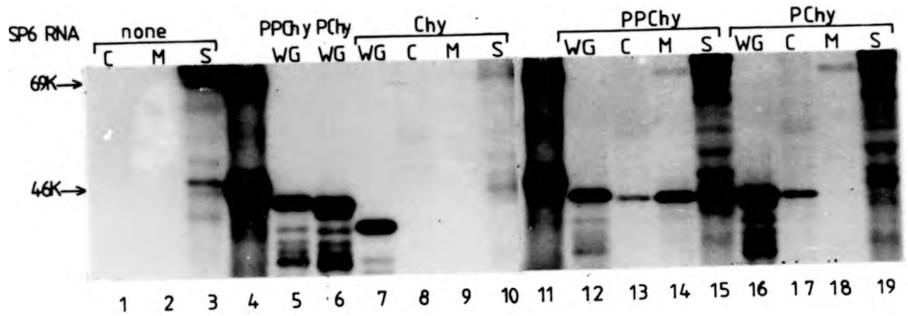


Figure V.3 Compartmentation in the oocyte of preprochymosin, prochymosin and chymosin expressed from synthetic RNAs

Xenopus oocytes were injected with m^{7}GG-capped SP6PPChy (tracks 13-15), SP6PChy (tracks 14-19) and SP6Chy (tracks 8-10), or left uninjected (tracks 1-3); the oocytes were cultured with [^{35}S]methionine and fractionated as described previously (II.H). Aliquots equivalent to 2 oocytes were taken from each cytosol (C) and membrane (M) fraction, and a sample representing the products secreted by 4 oocytes was taken from the S fractions; these samples were immunoprecipitated with antiprochymosin (αpC_{RB}, see II.G) together with samples from the wheat germ translations (WG) of the same RNAs (tracks 5-7, 12 & 16), described in Fig.V.2. The immunoprecipitated proteins were electrophoresed on a SDS-polyacrylamide gel (9%^{w}/v acrylamide with a cross linking ratio of 30:1.6); the fixed gel was fluorographed by the method of Bonner & Laskey(1974). The fluorographed gel was exposed to X-ray sensitive film for 7d.

expression of a chymosin-specific product which is secreted from the oocyte (track 15) and within the oocyte is predominately associated with the membrane and vesicles (track 14), with only a little detected in the cytosol fraction (track 13). The protein detected in, and secreted by, oocytes expressing SP6PPChy is the same size as the *in vitro* and oocyte translation product of SP6PChy (track 16 & 17), and is processed relative to the *in vitro* translation product of SP6PPChy (track 12). These results indicate that the preprochymosin expressed in the oocyte from the SP6PPChy RNA is translocated, processed and secreted by the oocyte, in agreement with the results obtained when preprochymosin was expressed in the oocyte from mRNA or the cloned PPChy cDNA (described and discussed in chapters III & IV). In addition results presented in chapter IV (IV.C & Fig.IV.4) showed that the product expressed in oocytes from SP6PPChy (which is derived from the cloned preprochymosin B gene) showed the same mobility on SDS-PAGE as the faster migrating species of the translocated prochymosin doublet expressed from the 'natural' preprochymosin mRNA; likewise the precursor seen on translation of SP6PPChy *in vitro* displayed the same migration as the mRNA primary *in vitro* translation product with the greater mobility on SDS-PAGE. The experiment shown in Fig.V.3 also clearly demonstrated that the signal-minus prochymosin expressed from SP6PChy in the oocyte is only present in the cytosol (track 17) and shows no association with membranes (track 18); hence even when the protein encoded by the PChy construct is well expressed in oocytes it is neither secreted (track 19) nor translocates the ER, thus confirming the earlier results discussed in chapter III. The injection of synthetic RNA transcribed from the Chy cDNA still gave no detectable chymosin translation product in oocytes (tracks 8-10).

It appears, therefore, that when *Xenopus* oocytes are injected with synthetic RNAs encoding preprochymosin, prochymosin and chymosin, the localisation of these polypeptides in *Xenopus* oocytes reflects that seen following injection of the corresponding cDNAs; and in the case of the synthetic SP6PPChy RNA, the preprochymosin expressed is compartmented as the preprochymosin expressed from the authentic mRNA. The expression and compartmentalisation of preprochymosin and the two signal-minus derivatives in the oocyte was discussed fully in Chapters III and IV, firstly in relation to results published by others, and also in the context of the proposed study of the translocation of signal sequence mutants of preprochymosin. The results presented here are discussed further in V.E.

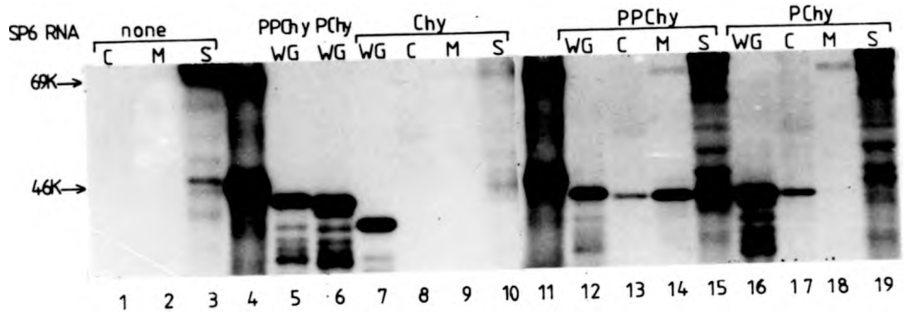


Figure V.3 Compartmentation in the oocyte of preprochymosin, prochymosin and chymosin expressed from synthetic RNAs

Xenopus oocytes were injected with m^{7}GG-capped SP6PPChy (tracks 13-15), SP6PChy (tracks 14-19) and SP6Chy (tracks 8-10), or left uninjected (tracks 1-3); the oocytes were cultured with [^{35}S]methionine and fractionated as described previously (II,H). Aliquots equivalent to 2 oocytes were taken from each cytosol (C) and membrane (M) fraction, and a sample representing the products secreted by 4 oocytes was taken from the S fractions; these samples were immunoprecipitated with antiprochymosin (αpC_{RS}, see II,G) together with samples from the wheat germ translations (WG) of the same RNAs (tracks 5-7, 12 & 16), described in Fig.V.2. The immunoprecipitated proteins were electrophoresed on a SDS-polyacrylamide gel (9%^{w}/v acrylamide with a cross linking ratio of 30:1,6); the fixed gel was fluorographed by the method of Bonner & Laskey(1974). The fluorographed gel was exposed to X-ray sensitive film for 7d.

expression of a chymosin-specific product which is secreted from the oocyte (track 15) and within the oocyte is predominately associated with the membrane and vesicles (track 14), with only a little detected in the cytosol fraction (track 13). The protein detected in, and secreted by, oocytes expressing SP6PPChy is the same size as the *in vitro* and oocyte translation product of SP6PChy (track 16 & 17), and is processed relative to the *in vitro* translation product of SP6PPChy (track 12). These results indicate that the preprochymosin expressed in the oocyte from the SP6PPChy RNA is translocated, processed and secreted by the oocyte, in agreement with the results obtained when preprochymosin was expressed in the oocyte from mRNA or the cloned PPChy cDNA (described and discussed in chapters III & IV). In addition results presented in chapter IV (IV.C & Fig.IV.4) showed that the product expressed in oocytes from SP6PPChy (which is derived from the cloned preprochymosin B gene) showed the same mobility on SDS-PAGE as the faster migrating species of the translocated prochymosin doublet expressed from the 'natural' preprochymosin mRNA; likewise the precursor seen on translation of SP6PPChy *in vitro* displayed the same migration as the mRNA primary *in vitro* translation product with the greater mobility on SDS-PAGE. The experiment shown in Fig.V.3 also clearly demonstrated that the signal-minus prochymosin expressed from SP6PChy in the oocyte is only present in the cytosol (track 17) and shows no association with membranes (track 18); hence even when the protein encoded by the PChy construct is well expressed in oocytes it is neither secreted (track 19) nor translocates the ER, thus confirming the earlier results discussed in chapter III. The injection of synthetic RNA transcribed from the Chy cDNA still gave no detectable chymosin translation product in oocytes (tracks 8-10).

It appears, therefore, that when *Xenopus* oocytes are injected with synthetic RNAs encoding preprochymosin, prochymosin and chymosin, the localisation of these polypeptides in *Xenopus* oocytes reflects that seen following injection of the corresponding cDNAs; and in the case of the synthetic SP6PPChy RNA, the preprochymosin expressed is compartmented as the preprochymosin expressed from the authentic mRNA. The expression and compartmentalisation of preprochymosin and the two signal-minus derivatives in the oocyte was discussed fully in Chapters III and IV, firstly in relation to results published by others, and also in the context of the proposed study of the translocation of signal sequence mutants of preprochymosin. The results presented here are discussed further in V.E.

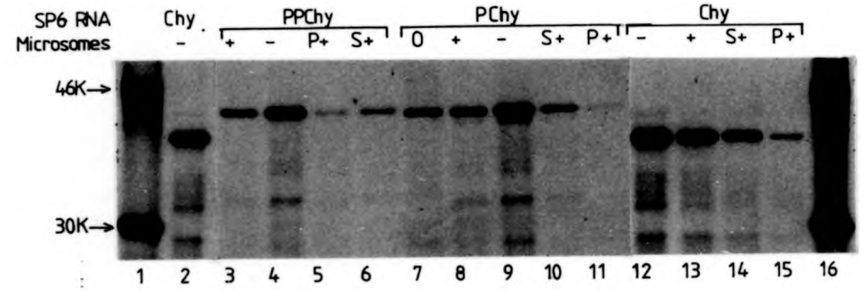


Figure V.4 Translation of chymosin SP6 RNAs in the presence of microsomal vesicles
Capped SP6PPChy (tracks 3-6), SP6PChy (tracks 8-11) and SP6Chy (tracks 2 & 12-15) transcripts (described in Fig.V.1) were translated in a wheat germ cell-free system in the absence (-) and presence (+) of canine pancreatic microsomes, as described in the Methods (II.E). Following the translation period a portion was removed from the '+' microsomes translations and the microsomes were pelleted by centrifugation, as described in II.E.3. Aliquots (6µl) from the untreated '-' and '+' microsomes translations and equivalent samples of the supernatant (S) and pelleted (P) material were analysed by immunoprecipitation with antiprochymosin (αpC_{ms}, see II.G) followed by SDS-PAGE on a 9% polyacrylamide gel (cross linking ratio 30:1.6). Track 7 (0) is a sample of the αpC_{ms}-precipitated proteins from the cytosolic fraction of oocytes injected with SP6PChy (see Fig.V.3). Tracks 1 & 16 contain molecular weight marker proteins.

V.D. *In Vitro* Translation of Chymosin SP6 RNAs in the Presence of Microsomes

Further aliquots of the m^{7}GG-capped chymosin RNAs were translated in a wheat germ system in the presence and absence of pancreatic microsomes, as described in the Materials & Methods (II.E). SDS-PAGE analysis of the products expressed from oviduct and lactogen mRNAs translated under the same conditions showed that in the presence of microsomes processing of the prelysozyme and prelactogen precursors, encoded by these mRNAs, had taken place (not shown). It is expected that in the presence of the microsomal vesicles only the PPChy translation product will be translocated and processed; both the PChy and Chy polypeptides will remain (uncleaved) outside the microsomes. After the translation period it was hoped to be able to show the specific translocation of the PPChy product into microsomes by separating the microsomes and associated proteins from the nontranslocated proteins in the translation mix. Part of the *in vitro* translations carried out in the presence of microsomes were layered on to a sucrose solution cushion, and these samples were centrifuged in an airfuge (see II.E.3); by this procedure proteins which are not translocated remain in the supernatant whilst the microsomal vesicles (and their contents) are pelleted. Antiprochymosin antibodies were used to immunoprecipitate chymosin polypeptides in the pelleted and supernatant material from the centrifuged samples, and also from equivalent portions of the untreated '+' and '-' microsomes translation assay mixes; the immunoprecipitated proteins were analysed by SDS-PAGE. In the case of the SP6PPChy translation products it is expected that unprocessed preprochymosin will be detected in the supernatant fraction whilst translocated, signal-processed prochymosin will be pelleted with the microsomes; in the unfractionated '+' microsomes sample both preprochymosin and prochymosin should be seen. Figure V.4 shows that in the centrifuged samples of the signal-sequence minus PChy and Chy products, which do not translocate the ER membrane, most of the prochymosin or chymosin is found in the supernatant (tracks 10 and 14), but a small proportion of these polypeptides were also detected in the pelleted microsomal fraction (tracks 11 and 15). Compared with these nontranslocated proteins, there was no pronounced association of the product expressed from SP6PPChy with the microsomes (see tracks 3, 5 & 6). This could reflect the observations that when prochymosin and chymosin polypeptides were synthesized in *E.coli* and yeast from specific cDNA constructs they, like the preprochymosin cDNA product, were found to be associated with the

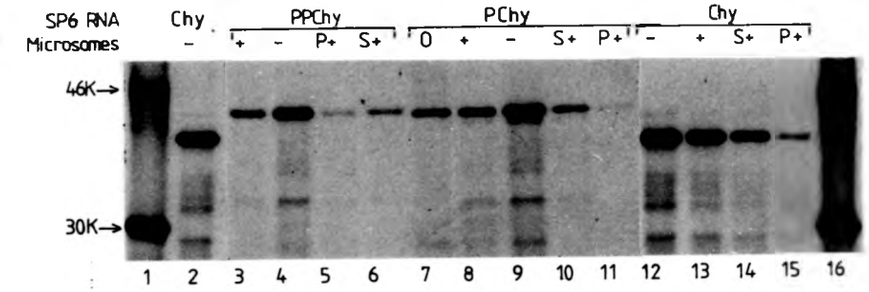


Figure V.4 Translation of chymosin SP6 RNAs in the presence of microsomal vesicles
Capped SP6PPChy (tracks 3-6), SP6PChy (tracks 8-11) and SP6Chy (tracks 2 & 12-15) transcripts (described in Fig.V.1) were translated in a wheat germ cell-free system in the absence (-) and presence (+) of canine pancreatic microsomes, as described in the Methods (II.E). Following the translation period a portion was removed from the '+' microsomes translations and the microsomes were pelleted by centrifugation, as described in II.E.3. Aliquots (6µl) from the untreated '-' and '+' microsomes translations and equivalent samples of the supernatant (S) and pelleted (P) material were analysed by immunoprecipitation with antiprochymosin (αpC_{ms}, see II.G) followed by SDS-PAGE on a 9% polyacrylamide gel (cross linking ratio 30:1.6). Track 7 (0) is a sample of the αpC_{ms}-precipitated proteins from the cytosolic fraction of oocytes injected with SP6PChy (see Fig.V.3). Tracks 1 & 16 contain molecular weight marker proteins.

V.D. *In Vitro* Translation of Chymosin SP6 RNAs in the Presence of Microsomes

Further aliquots of the m^{7}GG-capped chymosin RNAs were translated in a wheat germ system in the presence and absence of pancreatic microsomes, as described in the Materials & Methods (II.E). SDS-PAGE analysis of the products expressed from oviduct and lactogen mRNAs translated under the same conditions showed that in the presence of microsomes processing of the prelysozyme and prelactogen precursors, encoded by these mRNAs, had taken place (not shown). It is expected that in the presence of the microsomal vesicles only the PPChy translation product will be translocated and processed; both the PChy and Chy polypeptides will remain (uncleaved) outside the microsomes After the translation period it was hoped to be able to show the specific translocation of the PPChy product into microsomes by separating the microsomes and associated proteins from the nontranslocated proteins in the translation mix. Part of the *in vitro* translations carried out in the presence of microsomes were layered on to a sucrose solution cushion, and these samples were centrifuged in an airfuge (see II.E.3); by this procedure proteins which are not translocated remain in the supernatant whilst the microsomal vesicles (and their contents) are pelleted. Antiprochymosin antibodies were used to immunoprecipitate chymosin polypeptides in the pelleted and supernatant material from the centrifuged samples, and also from equivalent portions of the untreated '+' and '-' microsomes translation assay mixes; the immunoprecipitated proteins were analysed by SDS-PAGE. In the case of the SP6PPChy translation products it is expected that unprocessed preprochymosin will be detected in the supernatant fraction whilst translocated, signal-processed prochymosin will be pelleted with the microsomes; in the unfractionated '+' microsomes sample both preprochymosin and prochymosin should be seen. Figure V.4 shows that in the centrifuged samples of the signal-sequence minus PChy and Chy products, which do not translocate the ER membrane, most of the prochymosin or chymosin is found in the supernatant (tracks 10 and 14), but a small proportion of these polypeptides were also detected in the pelleted microsomal fraction (tracks 11 and 15). Compared with these nontranslocated proteins, there was no pronounced association of the product expressed from SP6PPChy with the microsomes (see tracks 3, 5 & 6). This could reflect the observations that when prochymosin and chymosin polypeptides were synthesized in *E.coli* and yeast from specific cDNA constructs they, like the preprochymosin cDNA product, were found to be associated with the

membranes and cell wall and did not remain free in the cytosol (Emtage *et al*,1983; Nishimori *et al*,1984; Mellor *et al*,1983; Goff *et al*,1984; discussed in I.D.3 & 4); although my own work has shown the PChy prochymosin product is not associated with membranes in oocytes. However it is not at all clear from the gel in Fig.V.4 that any cleavage of preprochymosin occurred in the '+' microsomes translation of SP6PPChy (track 3), to indicate that translocation of the precursor has taken place; although as noted above both prelysozyme and prelactogen were seen to be processed by microsomes in this experiment. Therefore further experiments are required tc characterise the translocation and processing of preprochymosin expressed from *in vitro* synthesized SP6 RNA (see V.E).

V.E. Discussion

The results described in this chapter showed the *in vitro* synthesized 'SP6 RNAs' provide an improved means of expressing foreign proteins encoded by cloned genes in *Xenopus* oocytes, and studying the compartmentation of such polypeptides. In the oocyte the localisation of prochymosin expressed from the PPChy and PChy capped synthetic RNAs was the same as that seen following injection of the corresponding cDNAs contained in the pTK_2 vector (described in Chapter III). As anticipated the translation product of the SP6PPChy RNA was segregated, cleaved and secreted in a manner analagous to the products expressed in oocytes from preprochymosin cDNA and mRNA (see Chapters III & IV). In contrast to the poor and variable expression seen from the PChy cDNA, the injection of PChy SP6 RNA resulted in the expression of readily detectable amounts of prochymosin in the oocyte, but the prochymosin was still only found in the cytosol and showed no association with the membrane and vesicle fraction of the oocyte; this was the expected localisation of this signal-minus prochymosin construct.

The results of this chapter also showed that the Chy construct was capable of generating a chymosin product when the SP6Chy RNA was translated *in vitro*; but injection of the Chy RNA into oocytes gave no detectable chymosin protein. Unless the SP6Chy transcript is considerably less stable in the oocyte than the SP6PChy or SP6PPChy RNAs (which could be examined), it is likely that the failure to detect the chymosin translation product *in vivo* is a consequence of the instability of chymosin in the cytoplasm of the oocyte, possibly due to the instability of chymosin at a pH>5.0 (as discussed in Chapter III).

In terms of the proposed mutagenesis of the signal sequence of preprochymosin these results reveal a serious limitation to the oocyte as

membranes and cell wall and did not remain free in the cytosol (Emtage *et al*,1983; Nishimori *et al*,1984; Mellor *et al*,1983; Goff *et al*,1984; discussed in I.D.3 & 4); although my own work has shown the PChy prochymosin product is not associated with membranes in oocytes. However it is not at all clear from the gel in Fig.V.4 that any cleavage of preprochymosin occurred in the '+' microsomes translation of SP6PPChy (track 3), to indicate that translocation of the precursor has taken place; although as noted above both prelysozyme and prelactogen were seen to be processed by microsomes in this experiment. Therefore further experiments are required to characterise the translocation and processing of preprochymosin expressed from *in vitro* synthesized SP6 RNA (see V.E).

V.E. Discussion

The results described in this chapter showed the *in vitro* synthesized 'SP6 RNAs' provide an improved means of expressing foreign proteins encoded by cloned genes in *Xenopus* oocytes, and studying the compartmentation of such polypeptides. In the oocyte the localisation of prochymosin expressed from the PPChy and PChy capped synthetic RNAs was the same as that seen following injection of the corresponding cDNAs contained in the pTK_2 vector (described in Chapter III). As anticipated the translation product of the SP6PPChy RNA was segregated, cleaved and secreted in a manner analagous to the products expressed in oocytes from preprochymosin cDNA and mRNA (see Chapters III & IV). In contrast to the poor and variable expression seen from the PChy cDNA, the injection of PChy SP6 RNA resulted in the expression of readily detectable amounts of prochymosin in the oocyte, but the prochymosin was still only found in the cytosol and showed no association with the membrane and vesicle fraction of the oocyte; this was the expected localisation of this signal-minus prochymosin construct.

The results of this chapter also showed that the Chy construct was capable of generating a chymosin product when the SP6Chy RNA was translated *in vitro*, but injection of the Chy RNA into oocytes gave no detectable chymosin protein. Unless the SP6Chy transcript is considerably less stable in the oocyte than the SP6PChy or SP6PPChy RNAs (which could be examined), it is likely that the failure to detect the chymosin translation product *in vivo* is a consequence of the instability of chymosin in the cytoplasm of the oocyte, possibly due to the instability of chymosin at a pH>5.0 (as discussed in Chapter III).

In terms of the proposed mutagenesis of the signal sequence of preprochymosin these results reveal a serious limitation to the oocyte as

an *in vivo* system in which to study the compartmentation of mutant polypeptides derived from secretory proteins. It is difficult to predict the effect of even the most innocuous amino acid change on the likely stability of the protein *in vivo*. It is expected that certain signal sequence mutants will be unable to initiate translocation across the ER membrane and these will remain in the cytoplasm when expressed in the oocyte. It has been reported that if secretory protein precursors are injected into *Xenopus* oocytes the presence of a cleavable signal sequence appears to make the miscompartmentalised protein unstable in the oocyte cytosol (Lane *et al*,1983). Although the results with the SP6PChy RNA suggest that the cytoplasmically localised prochymosin can be detected, it remains to be seen whether mutant preprochymosin precursors with functionally defective signal sequences will be detected when expressed in *Xenopus* oocytes from SP6 RNAs, or whether the presence of the uncleaved mutant signal peptide leads to their rapid degredation. It would be possible to carry out experiments to test the stability in the oocyte cytoplasm of the wild-type preprochymosin and prochymosin proteins by injecting the radiolabelled *in vitro* translation products of SP6PPChy and SP6PChy. Similarly once signal sequence mutants of preprochymosin were generated mutant precursors could be translated *in vitro* from SP6 transcripts and injected into oocytes to study their rate of degredation, particularly if it was found that, like the Chy construct, no chymosin-specific product could be detected from these mutants when expressed in the oocyte from SP6 RNAs. Thus we may anticipate that signal mutagenesis will generate a number of nontranslocating preprochymosins only some of which would be stable enough in the oocyte to reveal their *in vivo* phenotype. However, as noted previously (I.E) it is proposed to complement studies on the translocation of preprochymosin signal sequence mutants in the oocyte with *in vitro* translation translocation assays using reconstituted systems.

Preprochymosin (without its initiation methionine) has a molecluar weight 1,600 larger than than prochymosin, which has a molecular weight of 40,800; and it had been expected that the precursor and signal-processed proteins would be easily resolved on SDS-PAGE. However it had been found in the experiments on the expression of preprochymosin mRNA (described in Chapter IV) that judging by migration on SDS-PAGE both precursors encoded by the mRNA showed an apparent size reduction of only 1,000 on cleavage following translocation both in the oocyte and *in vitro*. Similarly in work with two preprochymosin/lysozyme fusion proteins which contained the signal peptide region of preprochymosin, the precursor and

cleaved forms displayed mobilities on SDS-PAGE consistent with a molecular weight difference of approx. 1,000 (described in the next chapter, VI). In the experiments carried out in Chapter IV the preprochymosin and prochymosin proteins expressed from the mRNA were found to be most clearly resolved on a 9% polyacrylamide gel containing 30% acrylamide:1.6% bisacrylamide. Although gels of this composition were also used for the experiments shown in Figures V.2 and V.3 the difference in migration of preprochymosin and prochymosin was less marked than had been observed previously (ie see Figs.IV.2 & IV.4); perhaps reflecting differences in the running conditions, although it was tried to keep these constant. It may be that through further tests a different gel composition could be found which would provide a more marked resolution of preprochymosin and prochymosin.

It had been hoped to be able to use signal peptide cleavage as an indicator for determining whether signal sequence mutants of preprochymosin were capable of initiating translocation either in the oocyte or *in vitro*, but as noted above this is not so readily assayed by electrophoresis of the chymosin proteins on SDS-polyacrylamide gels of the type used. It is also noted that preprochymosin signal sequence mutants may have a different migration on SDS-PAGE from wild-type preprochymosin and this could reduce further the mobility difference between the precursor and signal-processed forms of the protein. As mentioned earlier, certain mutations towards the carboxyl end of the signal peptide may destroy the signal peptidase recognition and cleavage site although the signal sequence function is not abolished. Hence the translocation of these mutants will initiated but the preprochymosin precursors will not be processed; it is possible also that in the absence of signal processing translocation may not be completed (eg Schauer *et al*,1985; discussed in I.B.3 and VI.F.2). Owing to the poor resolution of preprochymosin and prochymosin by SDS-PAGE and the possibility of signal sequence mutations affecting signal processing without abolishing translocation, other means are required to assay for translocation of preprochymosin.

It was thought that the separation of the microsomal vesicles from *in vitro* translation translocation assays, as described in V.C, could provide a means of distinguishing nascent proteins which were translocated *in vitro*. However in the experiment shown in Fig.V.4 the pelleted microsomes showed no enrichment of the translocated SP6PPChy product. It appears therefore that the resistance of translocated proteins to digestion by exogenous proteases will need to be used as the main

criterion for discriminating which products of wild-type and mutant preprochymosin SP6 RNAs are translocated when microsomes are present during translation *in vitro*. Experiments described in Chapter IV involved protease resistance assays to determine which proteins encoded by preprochymosin mRNA were translocated *in vitro*. It was found in these experiments that the conditions required to show protease protection varied according to the nature and amount of protein synthesized, and visualising the protected protein bands on a gel required a long exposure autoradiograph. Whilst the experiments described in IV.C did show the protease resistance of the two translocated prochymosins it is hoped that further modifications of the experimental conditions could improve the procedure for detecting translocated prochymosin; firstly to confirm the translocation of wild-type preprochymosin encoded by SP6PPChy, and to provide a convenient and clear means of determining the translocation phenotype of signal sequence mutants of preprochymosin *in vitro*. However, no further experiments were carried out at this stage towards optimising the protease resistance assay for preprochymosin, using the PPChy RNA and the PChy signal-minus derivative.

V.F. Summary

When *Xenopus* oocytes were injected with capped SP6PPChy RNA, transcribed *in vitro* from cloned preprochymosin B cDNA, the immunospecific product expressed was translocated and secreted by the oocytes in a manner similar to preprochymosin expressed in the oocyte from preprochymosin cDNA or from 'natural' preprochymosin mRNA. Consistent also with previous results, the protein secreted from oocytes injected with synthetic PPChy RNA was processed relative to the full-length primary translation product expressed when SP6PPChy was translated *in vitro*; and the processed PPChy product showed the same mobility on SDS-PAGE as the translation product of SP6PChy, which encodes prochymosin without the signal sequence. These results indicate that preprochymosin contains a signal sequence which directs its translocation across ER membranes in *Xenopus* oocytes, with cleavage of the signal peptide occurring; irrespective of whether the precursor is expressed from synthetic or authentic RNA, or cDNA.

Whereas nuclear injection of the PChy cDNA had only ever given very poor expression of a prochymosin protein in oocytes, the cytoplasmic injection of the corresponding PChy SP6 RNA resulted in the clear cut detection of a prochymosin polypeptide which was localised exclusively in the cytoplasm of the oocyte. This confirmed that in the absence of a

signal sequence prochymosin will not translocate the ER. As expected there was no difference in the size of the *in vitro* and *in vivo* translation product of SP6PChy, and either can provide a marker for the protein produced when the signal sequence is cleaved from preprochymosin. Whilst *in vitro* translation of SP6Chy encoding methionyl-chymosin produced a polypeptide of approx. 39K, which was immunoprecipitated by antibodies raised against prochymosin, no chymosin-specific polypeptide was detected following injection of capped SP6Chy RNA into *Xenopus* oocytes.

VI. THE EXPRESSION OF CHYMOSIN FUSION PROTEINS - ARE EUKARYOTIC SIGNAL SEQUENCES AUTONOMOUS?

VI.A. Introduction

When the work of this thesis was initiated the question still remained open as to whether all the information for translocation of a nascent protein across the ER membrane was contained within the signal sequence of eukaryotic secretory proteins. I felt it was important to establish whether eukaryotic signal sequences function autonomously in their role of acting as the signal for translocation across the ER membrane. It would be pertinent to the interpretation of results obtained with signal sequence mutants to know if nontranslocated mutant precursors could be regarded as being solely a consequence of altering the conformation of the signal peptide as a self-contained unit, or whether the effect of disruption of interactions between the signal peptide and the mature part of the protein were also to be considered.

In 1981 Engelman & Steitz had proposed the 'Helical Hairpin Hypothesis' for the insertion of proteins into and across membranes. A key postulate of this hypothesis was that, as the initial event in the translocation of a protein, the hydrophobic signal peptide specifically interacts with polar portions of the protein to form a helical hairpin structure. Thus according to this hypothesis the signal peptide region of a secretory protein precursor does not function as a self-contained unit in the process of translocation. Engelman & Steitz also envisaged the formation of the proposed hairpin conformation enabled the spontaneous partitioning of the nascent polypeptide chain into the membrane, without the participation of membrane receptors or transport proteins. On the other hand the concept of cotranslational translocation included in the Signal Hypothesis (Blobel & Dobberstein,1975a; see I.A.2) tended to support the idea of the functional autonomy of eukaryotic signal sequences (see below). Although it is noted that a modification of the original Signal Hypothesis introduced the idea that the N-terminal or internal signal sequence of secretory and membrane proteins inserted into the ER membrane as part of a looped structure (ie Blobel,1980, see Fig.I.2). These concepts were included in the Theory of Topogenic Sequences, proposed by Blobel in 1980, which postulated that protein sorting signals in general would be discrete, positively-acting units (see Introduction B.1.).

The experimental data concerning the signal sequence for translocation was not in complete agreement with the above hypotheses. At the outset of the work presented here there was accumulating experimental evidence for the involvement of specific ER membrane proteins in the translocation of nascent polypeptides (ie Walter & Blobel,1980; Meyer & Dobberstein, 1980a and b; see Introduction I.A.2). As outlined in the Introduction (I.A.2.) the current data on the translocation of eukaryotic secretory proteins had been largely derived from *in vitro* systems and this supported the idea of an obligate coupling of translocation to translation. These *in vitro* experiments led then to an updating of the Signal Hypothesis in which it was proposed that when the signal sequence of the nascent secretory protein emerged from the ribosome it interacted with SRP in the cytoplasm and translation was arrested until the SRP 'docked' with the SRP Receptor in the RER membrane, resulting in a release of the elongation arrest and completion of translation concomitant with translocation of the nascent polypeptide chain across the ER membrane (see Figure I.2). Signal processing would take place as the signal cleavage site was exposed to the signal peptidase on the lumenal side of the ER membrane. In such a scenario no interaction can occur between the signal sequence and other regions of the nascent polypeptide prior to translocation. However more of the published material at that time concerning the structure and function of the signal sequence related to prokaryotic secretory proteins. Whilst there were reports that the transfer of the leader peptide region of a prokaryotic secretory protein to a nonsegregating protein resulted in translocation of the hybrid protein (eg Michaelis & Beckwith,1982), there were also several results which indicated that the leader peptide alone was not sufficient to direct the translocation of all prokaryotic proteins (ie Moreno *et al*,1980; reviewed in Oliver,1985a; and see I.A.5 and I.B.2). In addition there was evidence that in prokaryotes translocation was not obligatorily coupled to translation, but that the leader peptide and part or all of the mature protein sequence could be synthesized before translocation commenced (Josefsson & Randall,1981); such a situation enables interactions to be set up between the signal sequence and other regions of the protein, and these could be important for the process of translocation. However, it was not clear to what extent the situation in prokaryotes reflected the process of translocation across the ER in eukaryotes. Furthermore it was also not known whether the tight coupling of translocation to translation seen in eukaryotic *in vitro* systems provided an accurate model for the mechanism *in vivo*.

To determine whether a eukaryotic signal peptide was a self contained unit which could function when transferred to another eukaryotic protein, it was decided to construct a hybrid cDNA in which the region coding for the signal sequence of one secretory precursor was fused to cDNA encoding the mature, signal-processed region of another secretory protein; the translocation of this chimaeric protein would then be examined, initially using the oocyte *in vivo* system described in Chapter III, and later *in vitro* using synthetic 'SP6 RNAs'. This strategy of creating a hybrid from two secretory proteins was adopted to avoid problems arising from a region of the hybrid molecule having a structure which could be nonpermissive for translocation across a membrane. Such problems had been encountered with prokaryotic signal sequence fusions involving the cytosolic protein β-galactosidase (encoded by the *lac Z* gene) as the 'recipient' part of the hybrid to which a signal sequence region was fused (see Michaelis & Beckwith,1982; Silhavy *et al*,1983).

VI.B. Construction of C_cL cDNA Encoding a Fusion Protein in Which the Signal Sequence of Preprochymosin Replaces the Signal Peptide of Prelysozyme

An examination of the DNA and amino acid sequences of the preprochymosin and prelysozyme clones we had in the laboratory (Harris *et al*,1982; Jung *et al*,1980) showed that it was possible to construct a fusion gene which encoded a hybrid protein in which the signal sequence and first 7 amino acids of lysozyme were replaced by the signal peptide region plus the first 6 residues of prochymosin. The fusion gene was accordingly termed C_cL and the primary translation product it encodes, preC_cL, has an expected molecular weight of M_r=15,900 (including the initiation methionine) which on correct cleavage of the chymosin signal peptide would result in the processed protein C_cL, M_r=14,200. It is anticipated that in the chimaeric polypeptide the preprochymosin sequences involved in the signal peptidase recognition site would be intact and therefore processing of the signal peptide on translocation is possible (Perlman & Halvorson,1983; Von Heijne,1983 and 1984a; see Introduction I.A.2). It has been shown that signal-processed secretory proteins, including lysozyme, are not secreted or sequestered into the ER when injected into *Xenopus* oocytes (Lane *et al*,1979; Lane *et al*,1980).

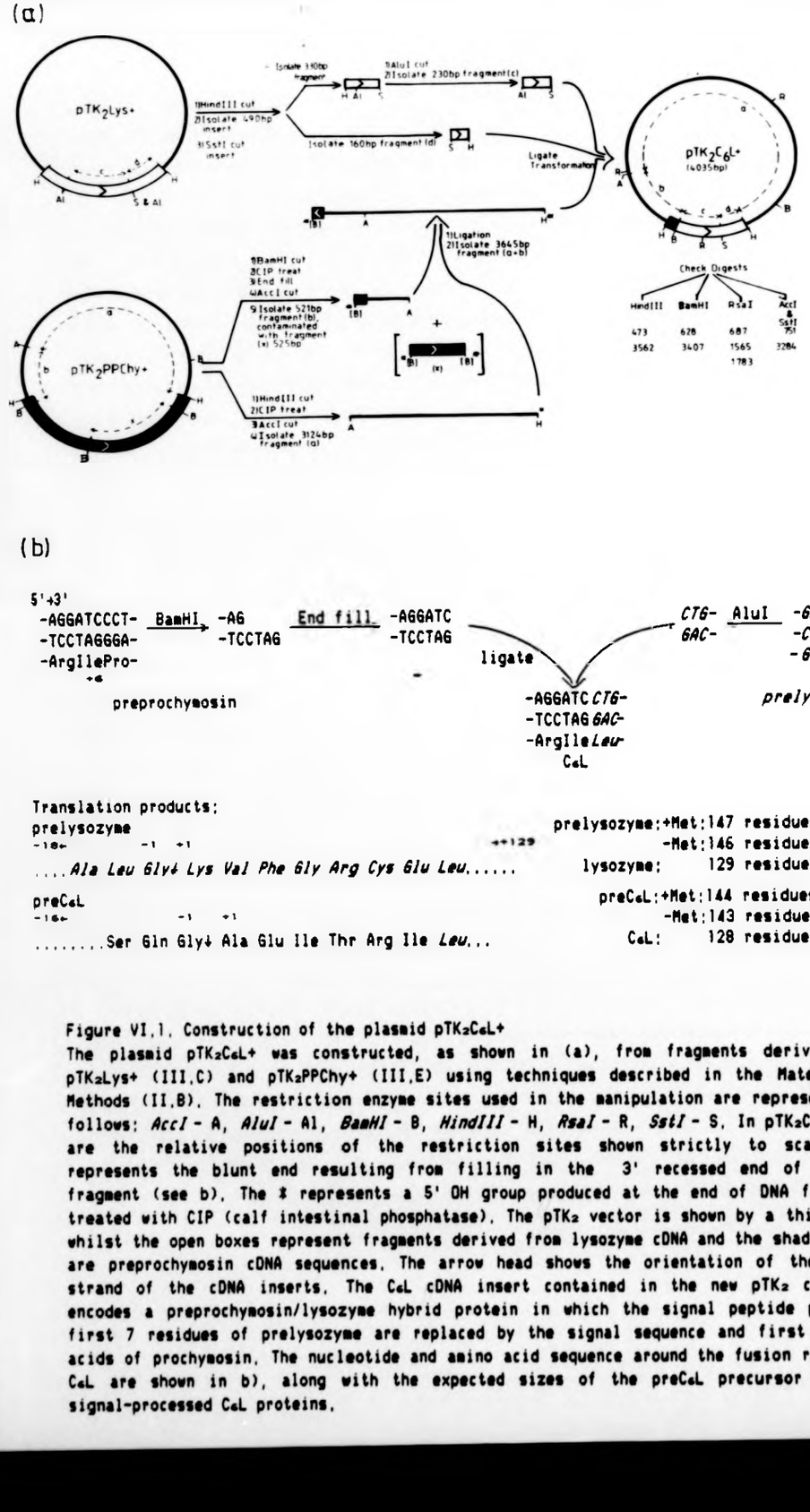


Figure VI.1. Construction of the plasmid pTK2C6L+
The plasmid pTK2C6L+ was constructed, as shown in (a), from fragments derived from pTK2Lys+ (III.C) and pTK2PPChy+ (III.E) using techniques described in the Materials & Methods (II.B). The restriction enzyme sites used in the manipulation are represented as follows: *AccI* - A, *AluI* - Al, *BamHI* - B, *HindIII* - H, *RsaI* - R, *SstI* - S. In pTK2C6L+ only are the relative positions of the restriction sites shown strictly to scale. [B] represents the blunt end resulting from filling in the 3' recessed end of a *BamHI* fragment (see b). The ↓ represents a 5' OH group produced at the end of DNA fragments treated with CIP (calf intestinal phosphatase). The pTK2 vector is shown by a thick line, whilst the open boxes represent fragments derived from lysozyme cDNA and the shaded boxes are preprochymosin cDNA sequences. The arrow head shows the orientation of the coding strand of the cDNA inserts. The C6L cDNA insert contained in the new pTK2 construct encodes a preprochymosin/lysozyme hybrid protein in which the signal peptide plus the first 7 residues of prelysozyme are replaced by the signal sequence and first 6 amino acids of prochymosin. The nucleotide and amino acid sequence around the fusion region in C6L are shown in b), along with the expected sizes of the preC6L precursor and the signal-processed C6L proteins.

VI.B.1 Construction of pTK2C6L+

Figure VI.1. shows the strategy employed to generate the plasmid pTK2C6L+, in which the C6L chimaeric cDNA is inserted in the expression vector pTK2; this figure also gives the nucleotide and amino acid structure of the preprochymosin/lysozyme fusion region. When the work described in this chapter was started the SP6 vectors were not yet available in the laboratory, therefore the fusion gene was constructed in the pTK2 vector for expression in oocytes. The sequence of the prelysozyme and preprochymosin cDNAs meant that there was no need to use oligonucleotide linkers to ensure the two coding regions were in the same reading frame; this was advantageous since it avoided any problems in introducing 'foreign' amino acids into the fusion, giving rise to a sequence of residues not present in either of the constituent secretory proteins. As shown in Fig.VI.1 transformants were selected which displayed the expected restriction fragments of pTK2C6L+ on digestion of plasmid DNA with the following enzymes: i)*HindIII* ii)*BamHI* iii)*AccI* & *SstI* iv)*RsaI*. The first three digests check the restriction sites used in the construction of the plasmid. In particular the *BamHI* site is maintained in the C6L insert if the DNA manipulations have occurred as shown in Fig.VI.1; the presence of this restriction site is a good indication that the fusion join between the preprochymosin and lysozyme cDNAs is correct and, therefore, the lysozyme coding sequence will be in the same reading frame as the preprochymosin signal peptide.

VI.B.2 DNA sequencing of the signal sequence and fusion regions of C6L

The nucleotide sequence at the junction between the preprochymosin and lysozyme cDNAs in C6L was later checked by DNA sequencing using the M13 vectors and the dideoxy chain termination method of Sanger *et al*(1977) as described in the Materials & Methods (II.C.). The C6L insert was excised from pTK2C6L+ using the restriction endonuclease *HindIII*, and this fragment was then cloned into the *HindIII* site of M13mp10. Figure VI.2 shows the partial sequence of a recombinant in which the chymosin sequences in the fusion were nearest to the sequencing primer site, and the sequence of bases read from the gel corresponded to the coding strand of C6L. This sequence data confirmed that the fusion between the preprochymosin and lysozyme coding regions was correct (Fig.VI.2b). The sequencing of the 5' part of the coding strand of C6L also provided an opportunity to verify the nucleotide sequence of the preprochymosin signal peptide; however one nucleotide in this region (marked with an asterix in Fig.VI2.a & b) did not agree with the

Figure VI.2 Partial DNA sequence of the fusion gene CεL

DNA encoding the hybrid protein CεL (described in Fig.VI.1) was cloned into the *HindIII* site of the M13 vector mp10, and the nucleotide sequence of the CεL insert was determined using the 'dideoxy' chain termination method of Sanger *et al*(1977). The products of the 4 sequencing reactions were separated, in the order shown (TCAG), by electrophoresis on an ultra-thin denaturing urea/polyacrylamide gel for (a) 90min (b) 195min. The fixed gel was dried onto 3MM paper then exposed at room temperature to X-ray film for 16h. The M13 cloning and sequencing techniques were carried out as described in the Materials & Methods (II.C).

The sequence of bases read upwards from the bottom of the gel corresponds to the 5'→3' nucleotide sequence of the CεL coding strand. The *HindIII* cloning site, and the *BamHI* site at the junction of the preprochymosin and lysozyme DNAs are marked. The asterix * marks the nucleotide which differs to the published preprochymosin B cDNA sequence (Harris *et al*,1982), and the arrow ,↓, indicates the first nucleotide in the fusion derived from the lysozyme cDNA. To give clarity to the bands seen on the autoradiograph different areas of the photograph were exposed for different durations during printing.

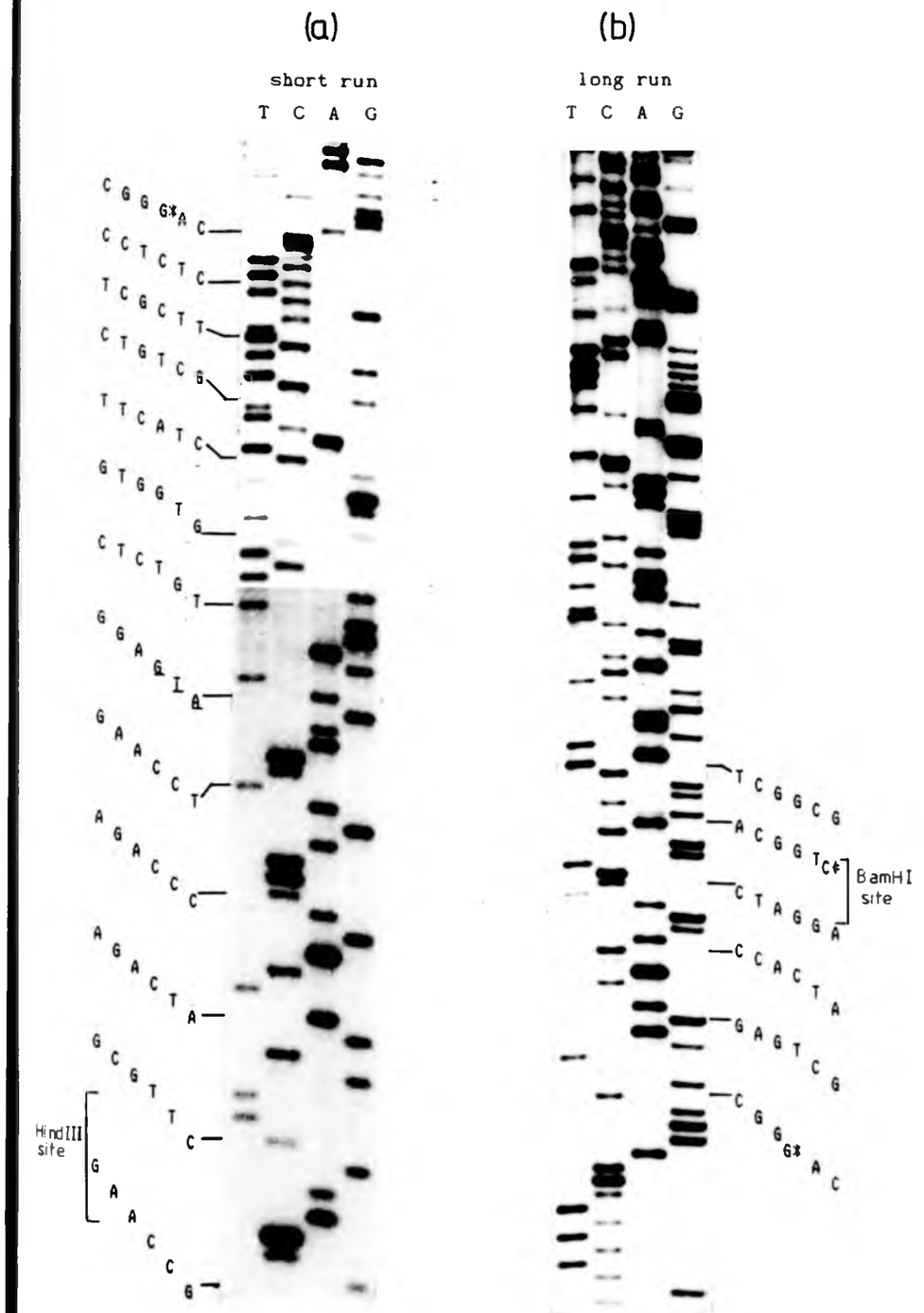
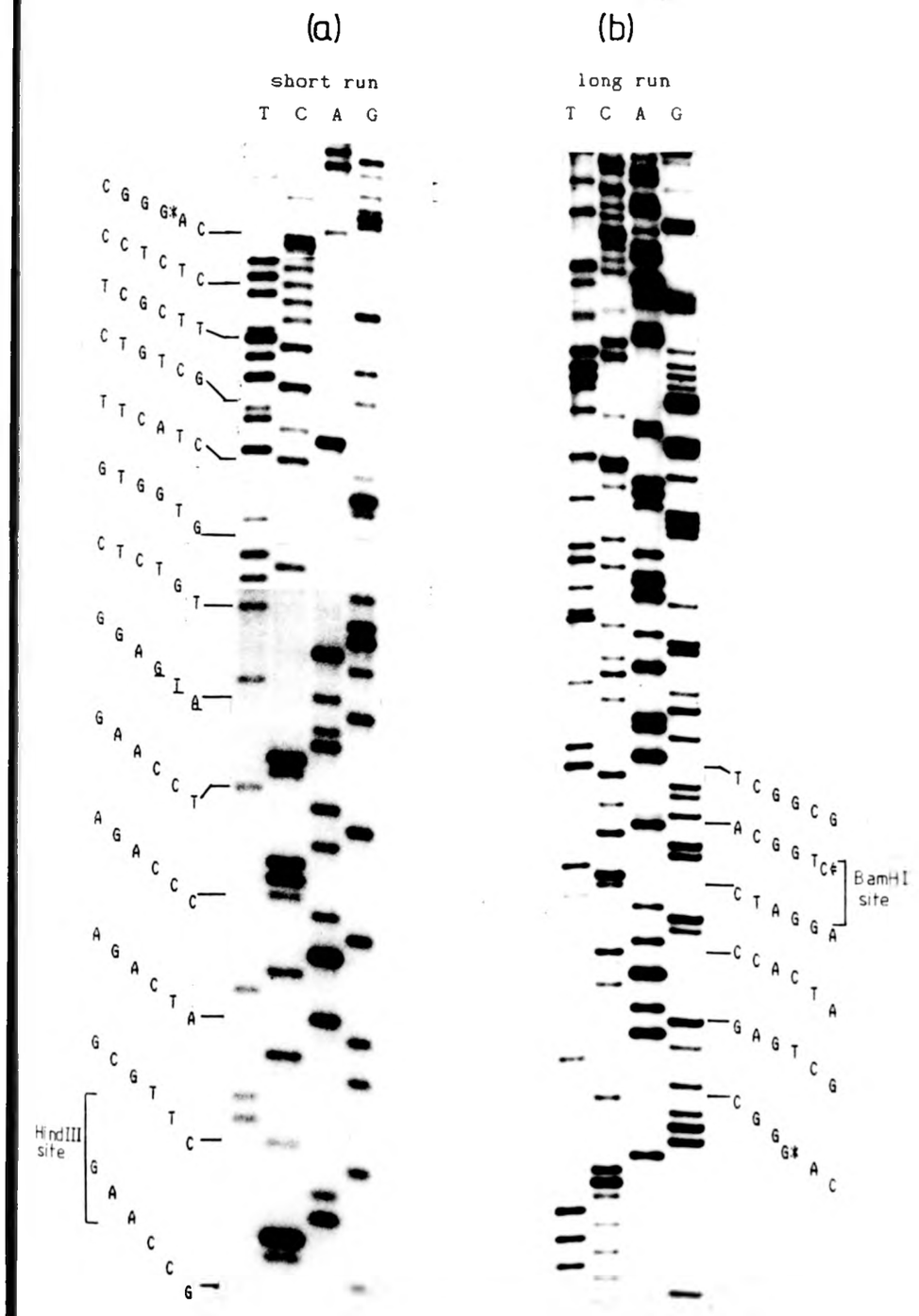


Figure VI.2 Partial DNA sequence of the fusion gene CεL

DNA encoding the hybrid protein CεL (described in Fig.VI.1) was cloned into the *HindIII* site of the M13 vector mp10, and the nucleotide sequence of the CεL insert was determined using the 'dideoxy' chain termination method of Sanger *et al*(1977). The products of the 4 sequencing reactions were separated, in the order shown (TCAG), by electrophoresis on an ultra-thin denaturing urea/polyacrylamide gel for (a) 90min (b) 195min. The fixed gel was dried onto 3MM paper then exposed at room temperature to X-ray film for 16h. The M13 cloning and sequencing techniques were carried out as described in the Materials & Methods (II.C).

The sequence of bases read upwards from the bottom of the gel corresponds to the 5'→3' nucleotide sequence of the CεL coding strand. The *HindIII* cloning site, and the *BamHI* site at the junction of the preprochymosin and lysozyme DNAs are marked. The asterix * marks the nucleotide which differs to the published preprochymosin B cDNA sequence (Harris *et al*,1982), and the arrow ,↓, indicates the first nucleotide in the fusion derived from the lysozyme cDNA. To give clarity to the bands seen on the autoradiograph different areas of the photograph were exposed for different durations during printing.



published sequence of the preprochymosin cDNA which we were using, cloned at Celltech (Harris *et al*,1982). According to Harris *et al* the penultimate codon of the signal sequence is CAA, whilst from my own sequence data it is CAG. Although both are codons for glutamine this apparent difference in the nucleotide sequence would be important when considering the site specific mutagenesis of the preprochymosin signal sequence (see Chapter VII). On all occasions when determining the DNA sequence of fusion constructs containing the prochymosin signal sequence (C_6L and $C_{\delta 2}L$) and the preprochymosin cDNA (the 'PPChy' insert, described in III.E.), the nucleotide sequence of the signal peptide of preprochymosin was consistently found to be the same as that shown in Fig.VI.2.

The position of the above discrepancy in the preprochymosin signal sequence corresponds to one of the six differences noted by Moir *et al*(1982) between the nucleotide sequence of the preprochymosin cDNA these workers had isolated and that reported for the gene cloned by Harris *et al*(1982) (see Introduction Table I.1 and section I.D.2); both Moir *et al* and Harris *et al* used the chemical cleavage method of Maxam & Gilbert(1980) to determine the DNA sequence of their clones. Like myself Moir *et al* found this codon of the signal peptide is CAG. Moir *et al* comment that this glutamine codon lies in one of the regions of the DNA in which only one strand was sequenced by Harris *et al*, and also involves nucleotides which would be modified as part of *EcoRII* recognition sequences. They suggest that the nucleotide discrepancy at this position is more likely to represent a modified cytosine residue which was misidentified by Harris *et al*, than a true allelic difference between the two cDNAs - the preprochymosin B gene cloned by Harris *et al*(1982) and the preprochymosin A gene which Moir *et al*(1982) had characterised. My own results with the Harris *et al* preprochymosin B cDNA support this suggestion made by Moir *et al*(1982).

VI.B.3 Expression in *Xenopus* oocytes of pTK_2C_6L+ encoding the chimaeric protein C_6L

The plasmid pTK_2C_6L+ was injected into the nucleus of oocytes which were then cultured and analysed as described previously (II.H. and III.C.), using anti-lysozyme sera to immunoprecipitate any C_6L present in the oocyte fractions. Several early experiments showed that injection of pTK_2C_6L+ into oocytes only ever resulted in very poor expression of a detectable translation product. Using S_1 analysis Dr. Linda Tabe carried out an analysis of the transcripts initiating at the vector TK promoter in oocytes injected with pTK_2 constructs, including pTK_2C_6L+. This showed

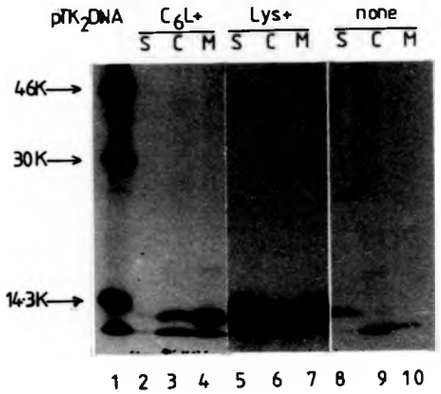


Figure VI.3 Expression of the fusion protein C₆L following injection of DNA into *Xenopus* oocytes

Xenopus oocytes were injected with pTK₂C₆L+ (tracks 2-4) or pTK₂Lys+ (tracks 5-7) plasmid DNA, and some were left uninjected as controls (tracks 8-10). The oocytes were incubated with [³⁵S]methionine for 24h, then the media containing secreted proteins was retained (S fraction) and the oocytes were fractionated into cytosolic (C) and membrane (M) associated proteins, as described in the Methods (II.H). Samples of the fractions were immunoprecipitated with antibodies to lysozyme, and the precipitated proteins were analysed by SDS-PAGE on a gel containing 12.5% polyacrylamide (30:0.825 cross-linked) as described in II.F. The C and M samples represent the products immunoprecipitated from 1 oocyte and the S sample 2.5 oocytes. The figure shows an autoradiograph of the fluorographed gel exposed to preflashed X-ray film for 7d. Track 1 shows the 30,000 and 14,300 molecular weight marker proteins (see II.F.2).

that large amounts of transcripts initiated at the TK promoter were present in the cytosol of oocytes injected with pTK₂C₆L+ (see Fig.7, Krieg *et al*,1984 in Appendix). Hence it was known that in the oocyte this construct was capable of generating specific transcripts from the TK promoter; however the probe used in these experiments to detect the transcripts was a general one and did not extend beyond the *HindIII* cloning site into the insert, so no information was obtained on the length of these transcripts. However it appeared that pTK₂C₆L+ was expressed in oocytes at the level of RNA, in a similar manner to other plasmids from which protein products had been detected (ie pTK₂PPChy+ (pTK₂82+), see Fig.7 Krieg *et al*,1984 and III.E & F). Furthermore, the possibility that an error had occurred in the construction of pTK₂C₆L+, with the effect that the preprochymosin and lysozyme reading frames were not in phase, had been eliminated by the DNA sequence analysis of the preprochymosin/lysozyme junction, described in the previous section (B.2).

The initial difficulty in demonstrating unequivocally either the expression of the hybrid C₆L protein or its localisation in oocytes raised the possibility that the signal sequence of preprochymosin was not sufficient to direct the translocation of C₆L and therefore the hybrid protein remained in the oocyte cytosol where it was rapidly degraded; in a situation similar to that encountered with the expression in oocytes of pTK₂PChy+ and pTK₂Chy+, which encode truncated chymosins lacking the signal peptide region (see III.E & G). Lane *et al*(1983) have reported that miscompartmentalized secretory proteins with cleavable signal sequences are very unstable in the cytosol of *Xenopus* oocytes. This prompted the decision to construct a cDNA for another hybrid protein containing not only the chymosin signal sequence but also part of the mature prochymosin amino acid sequence, to determine if this larger domain of preprochymosin carried the information required to achieve translocation in oocytes. The construction and expression of this second preprochymosin/lysozyme fusion gene is described in the next section.

In a few later experiments, however, it was found that pTK₂C₆L+ did clearly express a protein product in *Xenopus* oocytes. Figure VI.3 is from an experiment in which the expression of a product from pTK₂C₆L+ was relatively good, and a protein with a migration consistent with the expected molecular weight of the C₆L hybrid protein (M_r=14,200) was immunoprecipitated by anti-lysozyme sera from the injected oocytes. This protein was detected in the cytosol and membrane fractions of the oocyte (tracks 3 & 4, respectively). A faint band of the same size is also seen in track 2 which is from media surrounding the oocytes injected with

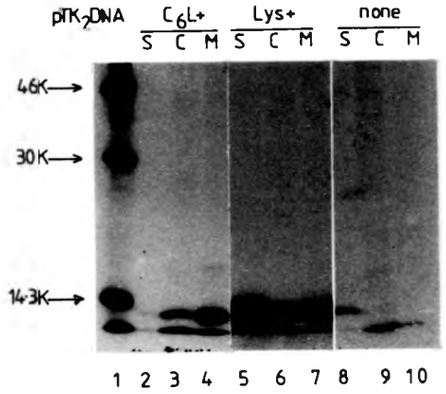


Figure VI.3 Expression of the fusion protein C₆L following injection of DNA into *Xenopus* oocytes

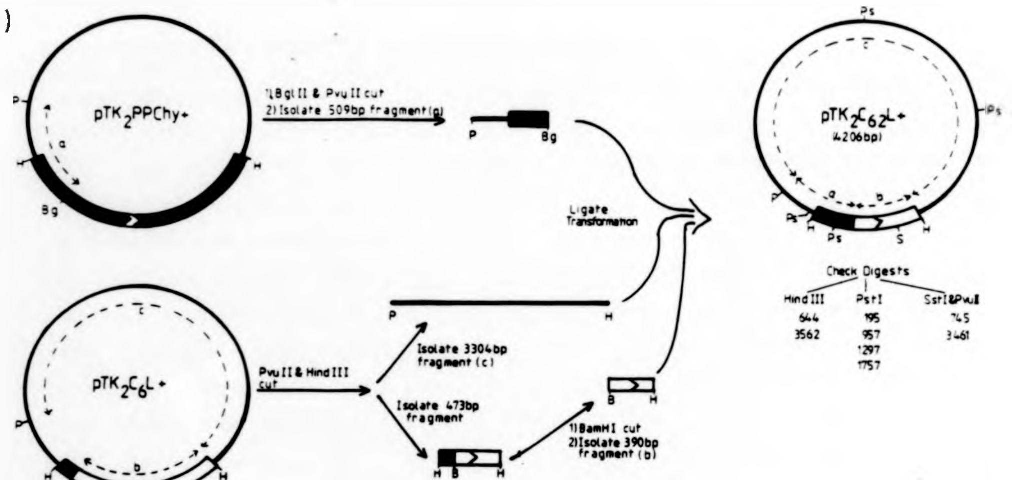
Xenopus oocytes were injected with pTK₂C₆L+ (tracks 2-4) or pTK₂Lys+ (tracks 5-7) plasmid DNA, and some were left uninjected as controls (tracks 8-10). The oocytes were incubated with [^{35}S]methionine for 24h, then the media containing secreted proteins was retained (S fraction) and the oocytes were fractionated into cytosolic (C) and membrane (M) associated proteins, as described in the Methods (II.H). Samples of the fractions were immunoprecipitated with antibodies to lysozyme, and the precipitated proteins were analysed by SDS-PAGE on a gel containing 12.5% polyacrylamide (30:0.825 cross-linked) as described in II.F. The C and M samples represent the products immunoprecipitated from 1 oocyte and the S sample 2.5 oocytes. The figure shows an autoradiograph of the fluorographed gel exposed to preflashed X-ray film for 7d. Track 1 shows the 30,000 and 14,300 molecular weight marker proteins (see II.F.2).

that large amounts of transcripts initiated at the TK promoter were present in the cytosol of oocytes injected with pTK₂C₆L+ (see Fig.7, Krieg *et al*,1984 in Appendix). Hence it was known that in the oocyte this construct was capable of generating specific transcripts from the TK promoter; however the probe used in these experiments to detect the transcripts was a general one and did not extend beyond the *HindIII* cloning site into the insert, so no information was obtained on the length of these transcripts. However it appeared that pTK₂C₆L+ was expressed in oocytes at the level of RNA, in a similar manner to other plasmids from which protein products had been detected (ie pTK₂PPChy+ (pTK₂82+), see Fig.7 Krieg *et al*,1984 and III.E & F). Furthermore, the possibility that an error had occurred in the construction of pTK₂C₆L+, with the effect that the preprochymosin and lysozyme reading frames were not in phase, had been eliminated by the DNA sequence analysis of the preprochymosin/lysozyme junction, described in the previous section (B.2).

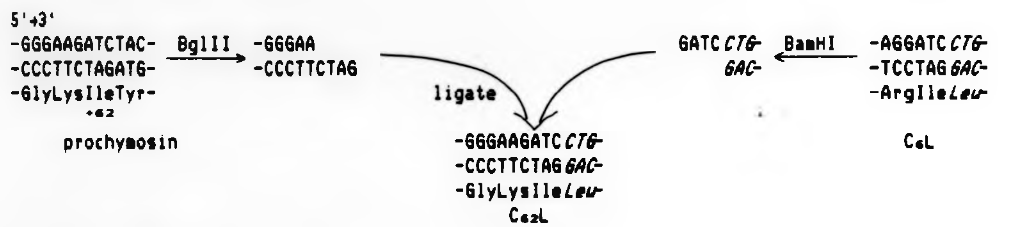
The initial difficulty in demonstrating unequivocally either the expression of the hybrid C₆L protein or its localisation in oocytes raised the possibility that the signal sequence of preprochymosin was not sufficient to direct the translocation of C₆L and therefore the hybrid protein remained in the oocyte cytosol where it was rapidly degraded; in a situation similar to that encountered with the expression in oocytes of pTK₂PChy+ and pTK₂Chy+, which encode truncated chymosins lacking the signal peptide region (see III.E & G). Lane *et al*(1983) have reported that miscompartmentalized secretory proteins with cleavable signal sequences are very unstable in the cytosol of *Xenopus* oocytes. This prompted the decision to construct a cDNA for another hybrid protein containing not only the chymosin signal sequence but also part of the mature prochymosin amino acid sequence, to determine if this larger domain of preprochymosin carried the information required to achieve translocation in oocytes. The construction and expression of this second preprochymosin/lysozyme fusion gene is described in the next section.

In a few later experiments, however, it was found that pTK₂C₆L+ did clearly express a protein product in *Xenopus* oocytes. Figure VI.3 is from an experiment in which the expression of a product from pTK₂C₆L+ was relatively good, and a protein with a migration consistent with the expected molecular weight of the C₆L hybrid protein (M_r=14,200) was immunoprecipitated by anti-lysozyme sera from the injected oocytes. This protein was detected in the cytosol and membrane fractions of the oocyte (tracks 3 & 4, respectively). A faint band of the same size is also seen in track 2 which is from media surrounding the oocytes injected with

(a)



(b)



$C_{62}L$ translation products:

preC62L +Met: 200 residues (22.2K)
-Met: 199 residues (22.1K)
C62L 184 residues (20.5K)

Figure VI.4. Construction of $pTK_2C_{62}L+$

a) The plasmid $pTK_2C_{62}L+$ was constructed from fragments derived from pTK_2C_6L+ (described in Fig.VI.1) and $pTK_2PPChy+$ (see III.E), using methods described in II.B. The restriction enzyme sites used in the manipulations are shown as follows: *BamHI* - B, *BglII* - Bg, *HindIII* - H, *PvuII* - P, *PstI* - Ps, *SstI* - S. Only in $pTK_2C_{62}L+$ are the relative positions of the restriction sites strictly to scale. The vector, lysozyme and preprochymosin sequences are represented as in Fig.VI.1.

(b) The nucleotide and amino acid sequence at the preprochymosin/lysozyme junction in $C_{62}L$ are shown. The primary translation product of the $C_{62}L$ cDNA, pre$C_{62}L$, is a hybrid polypeptide in which the signal sequence and first 62 residues of prochymosin are joined to the same C-terminal signal-minus lysozyme domain present in C_6L (see Fig.VI.1). The cleavage of the preprochymosin signal peptide from pre$C_{62}L$ gives rise to the processed $C_{62}L$ protein.

pTK_2C_6L+, but it is not clear if this is an artefactual band resulting from overspill of ^{14}C-labelled lysozyme from the adjacent molecular weight markers (track 1). A protein band of similar mobility was also present in track 8 which is the S sample of uninjected control oocytes, but this too is probably overspill from the adjacent track which contained lysozyme expressed from pTK_2Lys+ DNA (described in III.C). Clearly further experiments are required to investigate the compartmentation of the C_6L chimaeric protein in *Xenopus* oocytes (see VI.D).

VI.C. Construction and Expression in Oocytes of $pTK_2C_{62}L+$, Encoding A Second Preprochymosin/Lysozyme Fusion Protein

VI.C.1. Construction of a cDNA encoding $C_{62}L$, comprising the signal peptide and part of mature prochymosin fused to mature lysozyme

As outlined above, the negative results from early experiments on the expression of pTK_2C_6L+ gave rise to the idea that the signal peptide alone does not contain all the information necessary for the efficient translocation of secretory proteins, at least in the oocyte; and possibly part of the mature prochymosin protein sequence was also needed for the translocation of signal-minus lysozyme (more than the 6 amino acids present in C_6L). As noted earlier (VI.A.) work with prokaryotic fusion proteins had shown that the signal sequence region of a secretory protein could be insufficient to direct the translocation of a hybrid protein. For example work by Moreno *et al*(1980) showed that the leader peptide and 15 amino acids of the mature part of the LamB secretory protein, when fused to β-galactosidase did not result in the export of the hybrid protein. In the course of the work of this thesis Benson *et al*(1984) reported that LamB/β-galactosidase fusions containing, in addition to the signal sequence, less than 27 amino acids of mature LamB were not exported; and only those with more than 49 residues of mature LamB resulted in export of the hybrid protein to the outer membrane of *Escherichia coli*.

It was decided to construct a larger preprochymosin/lysozyme fusion protein which contained part of the mature prochymosin sequence as well as the signal peptide. Examination of the cDNAs available showed that it was possible to generate a second fusion gene in which the signal sequence of prelysozyme was replaced by a fragment from the preprochymosin cDNA which encoded the signal peptide and the first 62 amino acids of prochymosin. Figure VI.4 shows the strategy used to construct this fusion, termed $C_{62}L$, contained in the vector pTK_2; the

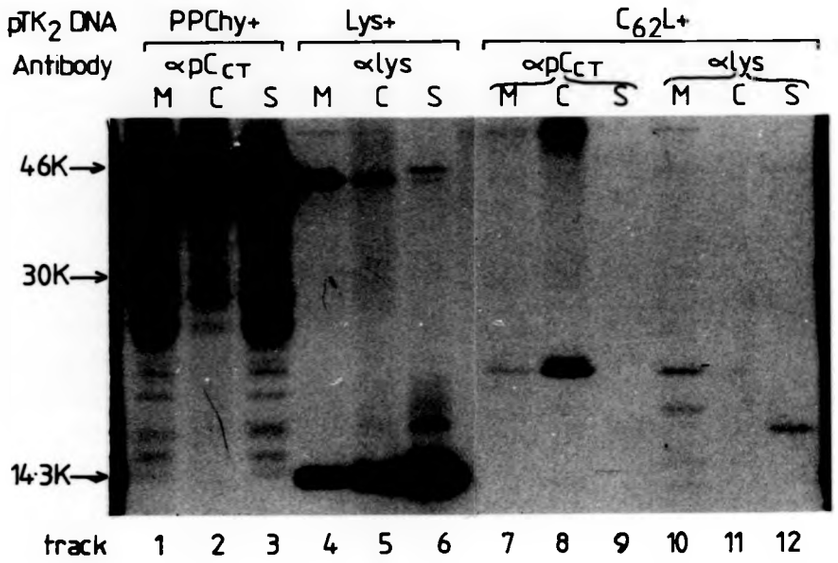


Figure VI.5 Expression in oocytes of the preprochymosin/lysozyme hybrid protein $C_{62}L$, following injection of $pTK_2C_{62}L+$ DNA

Xenopus oocytes were injected with the plasmids $pTK_2C_{62}L+$ (tracks 7-12), $pTK_2PPChy+$ (tracks 1-3) and pTK_2Lys+ (tracks 4-6), described previously (Fig.VI.4, Fig.III.5, and section III.C, respectively). The microinjection and subsequent analysis of the oocytes was carried out as described before (Fig.III.3 and in the Methods II.H). Duplicate samples of fractions from oocytes injected with $pTK_2C_{62}L+$ were immunoprecipitated with antiprochymosin sera (αpC_{CT}, II.G) (tracks 7-9) or with antibodies to lysozyme (αlys) (tracks 10-12). αlys was used for immunoprecipitating the products following injection of pTK_2Lys+, and αpC_{CT} for fractions from $pTK_2PPChy+$ injected oocytes. The immunoprecipitated proteins were analysed by SDS-PAGE on a 12.5% polyacrylamide gel with a cross linking ratio of 30:0.825 (see II.F.2). The cytosol (C) and membrane (M) fractions represent immunoprecipitated proteins from 1 oocyte, but 2.5 oocyte equivalents were used for the secreted (S) samples.

translation product it encodes, referred to as $preC_{62}L$, has an expected molecular weight M_r=22,200 (including the initiation methionine) which on cleavage of the signal peptide is reduced to M_r=20,500. Transformants were selected which gave the correct sized restriction fragments for $pTK_2C_{62}L+$ on digestion of the plasmid DNA with the following enzymes i)*HindIII* ii)*PstI* iii)*SstI & PvuII* (see Fig.VI.4).

VI.C.2. **Expression in oocytes of the fusion protein $C_{62}L$ from injected DNA**

The construct $pTK_2C_{62}L+$, like pTK_2C_6L+, gave poor and variable expression in *Xenopus* oocytes. Figure VI.5 shows an experiment in which oocytes were injected with various pTK_2 constructs, including $pTK_2C_{62}L+$. The fractions derived from oocytes injected with $pTK_2C_{62}L+$ were immunoprecipitated separately with antiprochymosin sera, αpC_{CT} (see II.G), and lysozyme antibodies (tracks 7-9 and 10-12 respectively). Injection of $pTK_2C_{62}L+$ into oocytes resulted in the expression of a product, precipitated by αpC_{CT}, which was detected in the cytosol (track 8) and vesicle (track 7) fractions of the oocyte but not in the secreted (S) fraction (track 9). The amount of product detected from $pTK_2C_{62}L+$ (tracks 7-12) was considerably less than that expressed from either $pTK_2PPChy+$ (tracks 1-3) or pTK_2Lys+ (tracks 4-6) encoding the full-length secretory proteins used to generate the $C_{62}L$ chimaera (see III.E & C). The size of the $pTK_2C_{62}L+$ product, judged by its mobility is approx. M_r=20,500, which corresponds to the expected size of the signal-processed hybrid protein; but in the absence of comparison with the full-length primary translation product it is not possible to draw any conclusions whether or not the signal peptide has been cleaved from the $C_{62}L$ oocyte product. A slightly larger minor product immunoprecipitated by αpC_{CT} was also detected in the cytosol fraction (track 8), this has a mobility consistent with a molecular weight of approx. M_r=21,000; the significance of this protein band is uncertain.

A curious feature displayed in Fig.VI.5 is the differential immunoprecipitation of $C_{62}L$ by antiprochymosin and antilysozyme (compare tracks 7-9 with 10-12). Whilst αpC_{CT} precipitated more product from the cytosol (track 8) than the membrane fraction (track 9), when antilysozyme was used as the antibody, although the material precipitated from the membrane sample (track 10) is roughly comparable to that precipitated by αpC_{CT}, virtually no product is detected in the cytosol (track 11). Neither of the antibodies immunoprecipitated any $C_{62}L$ from the medium surrounding the injected oocytes (track 9 & 12). The low levels of $C_{62}L$ protein detected, combined with this differential immunoprecipitation, makes it

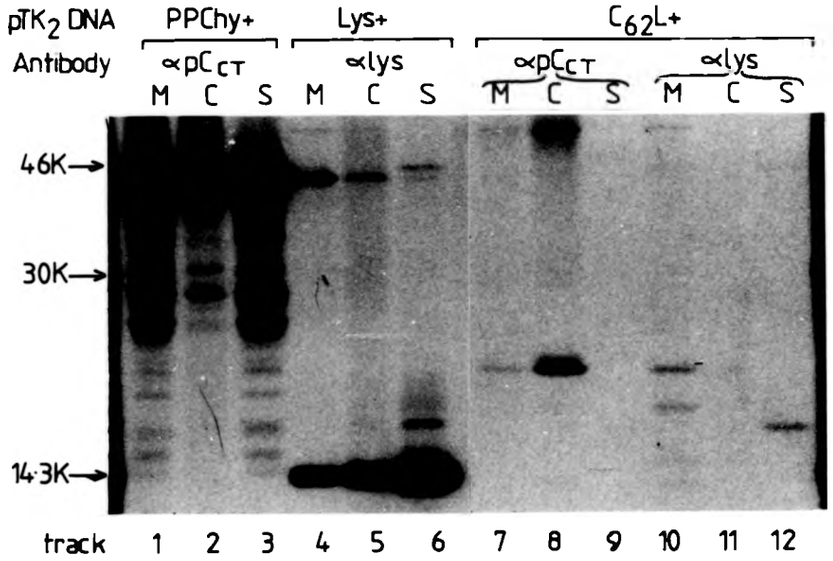


Figure VI.5 Expression in oocytes of the preprochymosin/lysozyme hybrid protein $C_{62}L$, following injection of $pTK_2C_{62}L+$ DNA

Xenopus oocytes were injected with the plasmids $pTK_2C_{62}L+$ (tracks 7-12), $pTK_2PPChy+$ (tracks 1-3) and pTK_2Lys+ (tracks 4-6), described previously (Fig.VI.4, Fig.III.5, and section III.C, respectively). The microinjection and subsequent analysis of the oocytes was carried out as described before (Fig.III.3 and in the Methods II.H). Duplicate samples of fractions from oocytes injected with $pTK_2C_{62}L+$ were immunoprecipitated with antiprochymosin sera (αpC_{CT}, II.G) (tracks 7-9) or with antibodies to lysozyme (αlys) (tracks 10-12). αlys was used for immunoprecipitating the products following injection of pTK_2Lys+, and αpC_{CT} for fractions from $pTK_2PPChy+$ injected oocytes. The immunoprecipitated proteins were analysed by SDS-PAGE on a 12.5% polyacrylamide gel with a cross linking ratio of 30:0.825 (see II.F.2). The cytosol (C) and membrane (M) fractions represent immunoprecipitated proteins from 1 oocyte, but 2.5 oocyte equivalents were used for the secreted (S) samples.

translation product it encodes, referred to as $preC_{62}L$, has an expected molecular weight M_r=22,200 (including the initiation methionine) which on cleavage of the signal peptide is reduced to M_r=20,500. Transformants were selected which gave the correct sized restriction fragments for $pTK_2C_{62}L+$ on digestion of the plasmid DNA with the following enzymes i)*HindIII* ii)*PstI* iii)*SstI* & *PvuII* (see Fig.VI.4).

VI.C.2. **Expression in oocytes of the fusion protein $C_{62}L$ from injected DNA**

The construct $pTK_2C_{62}L+$, like pTK_2C_6L+, gave poor and variable expression in *Xenopus* oocytes. Figure VI.5 shows an experiment in which oocytes were injected with various pTK_2 constructs, including $pTK_2C_{62}L+$. The fractions derived from oocytes injected with $pTK_2C_{62}L+$ were immunoprecipitated separately with antiprochymosin sera, αpC_{CT} (see II.G), and lysozyme antibodies (tracks 7-9 and 10-12 respectively). Injection of $pTK_2C_{62}L+$ into oocytes resulted in the expression of a product, precipitated by αpC_{CT}, which was detected in the cytosol (track 8) and vesicle (track 7) fractions of the oocyte but not in the secreted (S) fraction (track 9). The amount of product detected from $pTK_2C_{62}L+$ (tracks 7-12) was considerably less than that expressed from either $pTK_2PPChy+$ (tracks 1-3) or pTK_2Lys+ (tracks 4-6) encoding the full-length secretory proteins used to generate the $C_{62}L$ chimaera (see III.E & C). The size of the $pTK_2C_{62}L+$ product, judged by its mobility is approx. M_r=20,500, which corresponds to the expected size of the signal-processed hybrid protein; but in the absence of comparison with the full-length primary translation product it is not possible to draw any conclusions whether or not the signal peptide has been cleaved from the $C_{62}L$ oocyte product. A slightly larger minor product immunoprecipitated by αpC_{CT} was also detected in the cytosol fraction (track 8), this has a mobility consistent with a molecular weight of approx. M_r=21,000; the significance of this protein band is uncertain.

A curious feature displayed in Fig.VI.5 is the differential immunoprecipitation of $C_{62}L$ by antiprochymosin and antilysozyme (compare tracks 7-9 with 10-12). Whilst αpC_{CT} precipitated more product from the cytosol (track 8) than the membrane fraction (track 9), when antilysozyme was used as the antibody, although the material precipitated from the membrane sample (track 10) is roughly comparable to that precipitated by αpC_{CT}, virtually no product is detected in the cytosol (track 11). Neither of the antibodies immunoprecipitated any $C_{62}L$ from the medium surrounding the injected oocytes (track 9 & 12). The low levels of $C_{62}L$ protein detected, combined with this differential immunoprecipitation, makes it

difficult to interpret the results of the expression of $pTK_2C_{e2}L+$ in the oocyte, particularly with respect to the compartmentation of the hybrid protein in the oocyte. The phenomenum of the differential immunoprecipitation of $C_{e2}L$ and its localisation in the oocyte will be discussed later (VI.F.3).

VI.D. The Fusion Proteins, C_eL and $C_{e2}L$, are Segregated and Processed in Oocytes but are not Secreted

The results of Chapter V showed that the *in vitro* synthesized 'SP6 RNAs' could give good expression of foreign proteins in *Xenopus* oocytes, even when the corresponding cDNA gave poor expression. It was therefore decided to determine whether the cytoplasmic injection of synthetic SP6 RNAs would result in better expression of the fusion proteins C_eL and $C_{e2}L$ in oocytes, compared with their expression from cDNA injected into the nucleus. The SP6 RNAs could also be translated *in vitro* providing the full-length precursors for comparison with the oocyte products, to determine whether signal cleavage has occurred. In addition the SP6 RNAs enable a study to be made of the translocation *in vitro* of the hybrid proteins (see VI.E).

The *HindIII* inserts were excised from pTK_2Lys+, pTK_2C_eL+ and $pTK_2C_{e2}L+$ and cloned into the *HindIII* site of the polylinker of the vector pSP_{64} (Melton *et al*,1984). Transformants were selected which would generate 'sense' RNA corresponding to the coding strand of the cDNA. This cloning was carried out by Dave Jackson. To provide a template for *in vitro* transcription each of the pSP_{64} constructs were linearized by cutting the plasmid DNA at the unique *EcoRI* site in the polylinker region downstream of the insert, as described in the Methods (II.D). The RNAs synthesized from these templates are referred to as SP6Lys, $SP6C_eL$ and $SP6C_{e2}L$.

To investigate the expression in oocytes of the preprochymosin/lysozyme hybrid proteins from injected synthetic RNA, m^7GG-capped SP6 RNAs encoding lysozyme and the fusions C_eL and $C_{e2}L$ were prepared by *in vitro* transcription of the linearized DNA templates, as shown in Fig.V.1. The calculated yields of the RNAs were - SP6Lys, 2.8µg; $SP6C_eL$, 3.0µg; $SP6C_{e2}L$, 2.9µg, and the relative sizes of these transcripts was as expected from the lengths of the three inserts (see Fig.V.1b). *Xenopus* oocytes were microinjected with these capped SP6 RNAs and then cultured with [^{35}S]methionine, fractionated and analysed as described in II.H. Further samples of SP6Lys, $SP6C_eL$ and $SP6C_{e2}L$ were translated in a wheat germ cell-free system as described in the Methods (II.E.1). The fractions

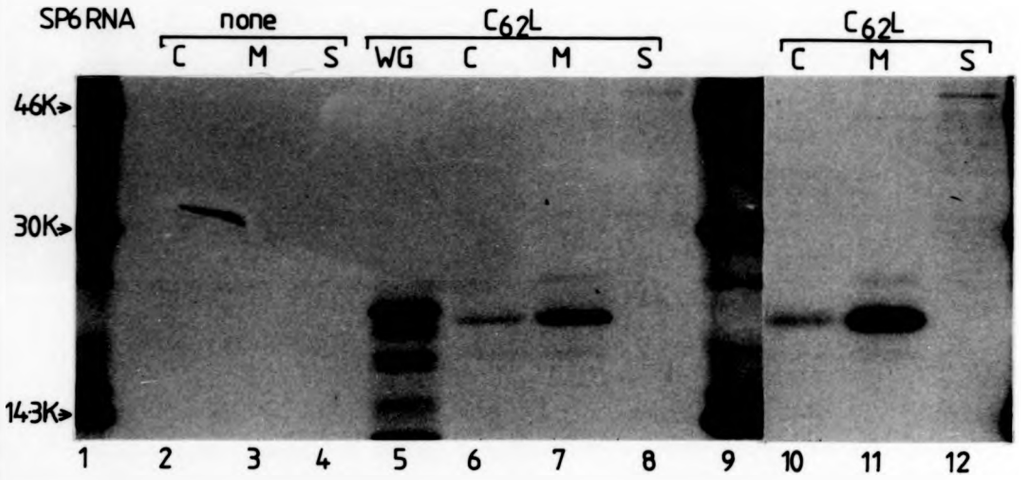


Figure VI.6 Segregation and processing of the $C_{62}L$ fusion protein expressed in oocytes from synthetic RNA.
m^7GG-capped $SP6C_{62}L$ RNA was synthesized as described in Fig.V.1. These synthetic transcripts were microinjected into *Xenopus* oocytes which were then labelled with [^{35}S]methionine for 24h, and analysed as described in the Methods (II.H). The $SP6C_{62}L$ RNA was also translated *in vitro* in a wheat germ system (at 8ng/µl, see II.E.1). Aliquots from the *in vitro* translation (WG) and from oocyte fractions representing secreted (S), cytosolic (C), and membrane and vesicle (M) associated proteins were immunoprecipitated with antibodies raised against prochymosin (αpC_{RB}, II.G) and lysozyme; the immunoprecipitated proteins were then separated by SDS-PAGE on a gel containing 15% polyacrylamide (40% acrylamide:0.2% bisacrylamide), as described in II.F. The dried fluorographed gel was exposed to X-ray sensitive film for 2 weeks (tracks 1-9) and 5 weeks (tracks 10-12). The C & M samples are equivalent to 2 oocytes, whilst the S samples represent proteins secreted by 4 oocytes. Tracks 6-8 and 10-12 are samples from oocytes injected with m^7GG-capped $SP6C_{62}L$ RNA, whilst tracks 2-4 are from control uninjected oocytes treated in the same way, outlined above. Tracks 1 & 9 show the 14,300, 30,000 and 46,000 ^{14}C-labelled molecular weight marker proteins (see II.F.2).

from oocytes injected with SP6Lys and $SP6C_6L$, and the corresponding *in vitro* translations, were immunoprecipitated with lysozyme antibodies (see II.F.1); whilst the samples from the oocyte and wheat germ translation of $SP6C_{62}L$ were precipitated with a mixture of antilysozyme and antiprochymosin (αpC_{RB}, see II.G). The immunoprecipitated proteins were then analysed by SDS-PAGE and fluorography. The results obtained from these experiments are described below, but they are discussed in detail later in section VI.F.

Figure VI.6 shows the proteins, immunoprecipitated by antilysozyme and antiprochymosin, expressed from $SP6C_{62}L$ in oocytes and *in vitro*. Translation of $SP6C_{62}L$ in the wheat germ cell-free system should produce the full-length $preC_{62}L$ polypeptide which from its amino acid sequence is calculated to have a molecular weight of 22,100, without its initiation methionine (see Fig.VI.4). The major immunoprecipitated protein produced from translation of m^7GG-capped $SP6C_{62}L$ in the wheat germ (track 5) has a mobility consistent with a molecular weight of approx. 22,000, and is likely to be $preC_{62}L$. In the oocyte the $C_{62}L$ synthetic RNA results in the expression of a protein which judged by its migration on SDS-PAGE is smaller than the $preC_{62}L$ *in vitro* product and has a molecular weight of approx. M_r=21,000 (tracks 6-8). Since this 21K protein is also associated with the vesicle fraction of the oocyte (track 7) this indicates that in the oocyte the hybrid protein $preC_{62}L$ encoded by $SP6C_{62}L$ translocates into the ER membrane where the signal sequence is processed by signal peptidase. Only a small amount of the processed $C_{62}L$ is detected in the cytosol fraction of the oocyte (track 6) whilst the majority is found to be associated with the membrane fraction; as discussed earlier (III.F), this distribution is typical of translocated proteins in the oocyte. Hence the information to direct the translocation in oocytes of signal-minus lysozyme is apparently contained within the signal peptide plus the first 62 residues of prochymosin. However none of the hybrid $C_{62}L$ protein was detected in the medium surrounding the oocytes injected with $SP6C_{62}L$ (track 8), not even on a longer exposure autoradiograph of the same gel (track 12). Yet oocytes from the same batch efficiently secreted lysozyme following injection of SP6Lys (see Fig.VI.7). The failure to detect secreted $C_{62}L$ suggests that sorting signals, required after translocation to direct a protein along the secretory pathway for export from the oocyte, are defective in the $C_{62}L$ chimaeric protein.

In Figure VI.7 the proteins immunoprecipitated by antilysozyme from oocytes injected with capped $SP6C_6L$ and SP6Lys are compared with the *in vitro* translation products of the same RNAs. These proteins were

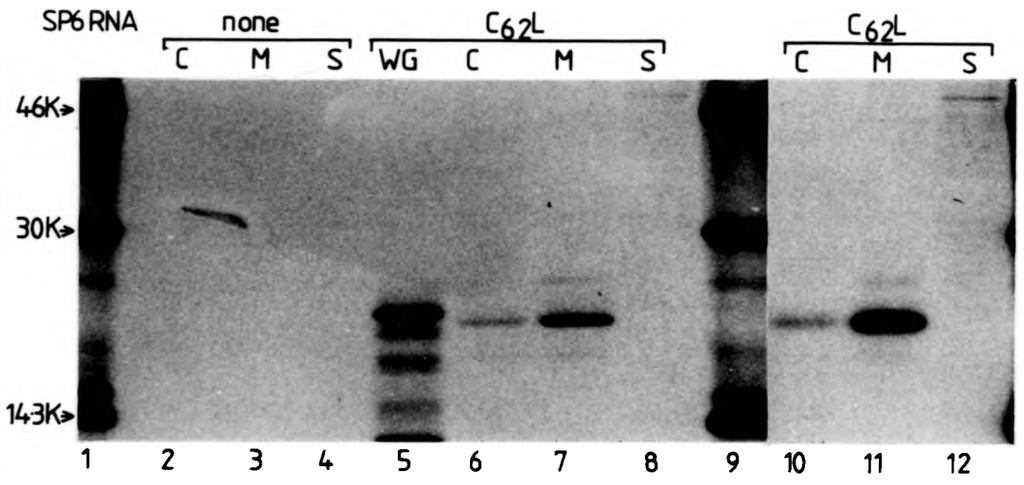


Figure VI.6 Segregation and processing of the $C_{62}L$ fusion protein expressed in oocytes from synthetic RNA.
m^7GG-capped $SP6C_{62}L$ RNA was synthesized as described in Fig.V.1. These synthetic transcripts were microinjected into *Xenopus* oocytes which were then labelled with [^{35}S]methionine for 24h, and analysed as described in the Methods (II.H). The $SP6C_{62}L$ RNA was also translated *in vitro* in a wheat germ system (at 8ng/µl, see II.E.1). Aliquots from the *in vitro* translation (WG) and from oocyte fractions representing secreted (S), cytosolic (C), and membrane and vesicle (M) associated proteins were immunoprecipitated with antibodies raised against prochymosin (αpC_{ms}, II.6) and lysozyme; the immunoprecipitated proteins were then separated by SDS-PAGE on a gel containing 15% polyacrylamide (40% acrylamide:0.2% bisacrylamide), as described in II.F. The dried fluorographed gel was exposed to X-ray sensitive film for 2 weeks (tracks 1-9) and 5 weeks (tracks 10-12). The C & M samples are equivalent to 2 oocytes, whilst the S samples represent proteins secreted by 4 oocytes. Tracks 6-8 and 10-12 are samples from oocytes injected with m^7GG-capped $SP6C_{62}L$ RNA, whilst tracks 2-4 are from control uninjected oocytes treated in the same way, outlined above. Tracks 1 & 9 show the 14,300, 30,000 and 46,000 ^{14}C-labelled molecular weight marker proteins (see II.F.2).

from oocytes injected with SP6Lys and $SP6C_{6}L$, and the corresponding *in vitro* translations, were immunoprecipitated with lysozyme antibodies (see II.F.1); whilst the samples from the oocyte and wheat germ translation of $SP6C_{62}L$ were precipitated with a mixture of antilysozyme and antiprochymosin (αpC_{ms}, see II.G). The immunoprecipitated proteins were then analysed by SDS-PAGE and fluorography. The results obtained from these experiments are described below, but they are discussed in detail later in section VI.F.

Figure VI.6 shows the proteins, immunoprecipitated by antilysozyme and antiprochymosin, expressed from $SP6C_{62}L$ in oocytes and *in vitro*. Translation of $SP6C_{62}L$ in the wheat germ cell-free system should produce the full-length $preC_{62}L$ polypeptide which from its amino acid sequence is calculated to have a molecular weight of 22,100, without its initiation methionine (see Fig.VI.4). The major immunoprecipitated protein produced from translation of m^7GG-capped $SP6C_{62}L$ in the wheat germ (track 5) has a mobility consistent with a molecular weight of approx. 22,000, and is likely to be $preC_{62}L$. In the oocyte the $C_{62}L$ synthetic RNA results in the expression of a protein which judged by its migration on SDS-PAGE is smaller than the $preC_{62}L$ *in vitro* product and has a molecular weight of approx. M_r=21,000 (tracks 6-8). Since this 21K protein is also associated with the vesicle fraction of the oocyte (track 7) this indicates that in the oocyte the hybrid protein $preC_{62}L$ encoded by $SP6C_{62}L$ translocates into the ER membrane where the signal sequence is processed by signal peptidase. Only a small amount of the processed $C_{62}L$ is detected in the cytosol fraction of the oocyte (track 6) whilst the majority is found to be associated with the membrane fraction; as discussed earlier (III.F), this distribution is typical of translocated proteins in the oocyte. Hence the information to direct the translocation in oocytes of signal-minus lysozyme is apparently contained within the signal peptide plus the first 62 residues of prochymosin. However none of the hybrid $C_{62}L$ protein was detected in the medium surrounding the oocytes injected with $SP6C_{62}L$ (track 8), not even on a longer exposure autoradiograph of the same gel (track 12). Yet oocytes from the same batch efficiently secreted lysozyme following injection of SP6Lys (see Fig.VI.7). The failure to detect secreted $C_{62}L$ suggests that sorting signals, required after translocation to direct a protein along the secretory pathway for export from the oocyte, are defective in the $C_{62}L$ chimaeric protein.

In Figure VI.7 the proteins immunoprecipitated by antilysozyme from oocytes injected with capped $SP6C_{6}L$ and SP6Lys are compared with the *in vitro* translation products of the same RNAs. These proteins were

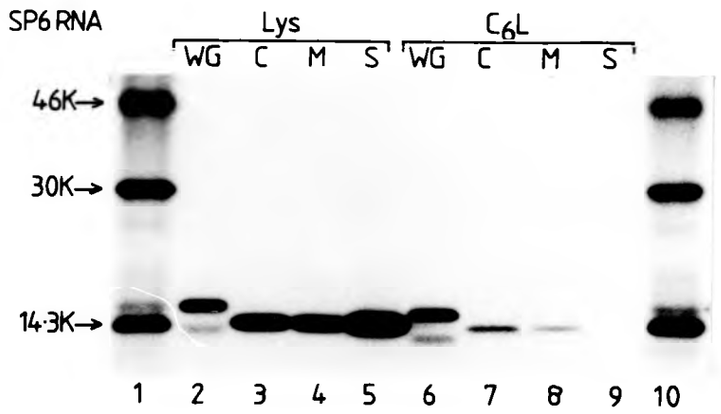


Figure VI.7 Compartmentation and cleavage in oocytes of C₆L expressed from synthetic RNA
Capped SP6C₆L and SP6Lys transcripts were synthesized as described in Fig.V.1. The analysis of the proteins expressed from these synthetic RNAs following injection into *Xenopus* oocytes and on translation *in vitro* was carried out as described in Fig.VI.6, except that lysozyme antibodies only were employed for immunoprecipitation of the SP6C₆L and SP6Lys products. In addition only 1 oocyte equivalent of the C, M & S fractions of oocytes injected with SP6Lys were electrophoresed (tracks 3-5, respectively); but as in Fig.VI.6, for the SP6C₆L injected oocytes the C & M samples represent 2 oocytes (tracks 7 & 8, respectively), and the S sample 4 oocytes (track 9). The *in vitro* translation products immunoprecipitated by antilysozyme expressed from SP6C₆L and SP6Lys are shown in track 5 and track 2, respectively. Tracks 1 & 10 are ^{14}C-labelled molecular weight marker proteins (see II.F.2), including lysozyme (14.3K). The figure is of a 7d exposure autoradiograph of the fluorographed gel (15% polyacrylamide with a cross-linking ratio of 40:0.2).

separated on a gel containing 15% polyacrylamide with a ratio of acrylamide to bisacrylamide of 40%:0.2%; this gel composition was used as it was hoped to provide resolution of polypeptides the size of prelysozyme and preC₆L. In contrast to the experiments which examined the expression of the C₆L fusion cDNA in oocytes (VI.B.3), the cytoplasmic injection of the synthetic SP6C₆L transcript clearly resulted in the expression of a protein, recognised by lysozyme antibodies, which migrated to the same position on the gel as the lysozyme in the molecular weight markers, M_r=14,300 (tracks 1 & 7); this is consistent with the expected size of the signal processed C₆L hybrid protein which contains only one less amino acid than lysozyme (see Fig.VI.1). This implies that the signal sequence has been cleaved from the SP6C₆L product in the oocyte. This suggestion is further supported by the fact that the polypeptide expressed from SP6C₆L on translation *in vitro* (track 6) is larger than the protein detected in the oocyte; this precursor, preC₆L, displays a mobility on SDS-PAGE consistent with a molecular weight of approx. M_r=15,500, which compares with a molecular weight of M_r=15,800 calculated from the amino acid sequence of preC₆L excluding the initiation methionine residue (see Fig.VI.1). Like the C₆₂L hybrid protein, the C₆L product is detected in the cytosol (Fig.VI.7 track 7) and vesicle (track 8) fractions of the oocyte but is not secreted from the oocytes (track 9). As the processed C₆L protein is associated with the vesicle fraction of the oocyte this indicates that the signal sequence plus the first six amino acids of prochymosin is sufficient to initiate the translocation of the signal-minus lysozyme *in vivo* in the oocyte. Furthermore the observed processing of the oocyte product relative to the full-length *in vitro* product strongly suggests that the preC₆L fusion protein encoded by SP6C₆L also contains a functional signal peptidase recognition site at which cleavage has occurred when the hybrid protein gains access to the ER in oocytes. Therefore the additional residues from the mature prochymosin protein contained in the larger preC₆₂L fusion are not essential for translocation.

Compared with the RNAs for both the hybrid proteins, injection of SP6Lys resulted in the detection of large amounts of immunospecific protein in the oocyte (Fig.VI.7 tracks 3-5), producing a product which had the same mobility as the ^{14}C-labelled lysozyme, M_r=14,300 (track 1), and was processed relative to the full-length precursor expressed *in vitro*, M_r=16,000 (track 2). In agreement with results obtained with lysozyme mRNA and lysozyme cDNA (see III.C), the lysozyme expressed from the synthetic RNA is secreted efficiently by the oocyte. In addition, as

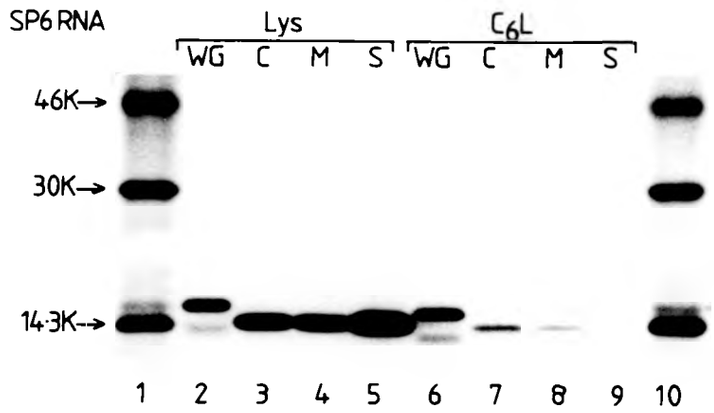


Figure VI.7 Compartmentation and cleavage in oocytes of C₆L expressed from synthetic RNA
Capped SP6C₆L and SP6Lys transcripts were synthesized as described in Fig.V.1. The analysis of the proteins expressed from these synthetic RNAs following injection into *Xenopus* oocytes and on translation *in vitro* was carried out as described in Fig.VI.6, except that lysozyme antibodies only were employed for immunoprecipitation of the SP6C₆L and SP6Lys products. In addition only 1 oocyte equivalent of the C, M & S fractions of oocytes injected with SP6Lys were electrophoresed (tracks 3-5, respectively); but as in Fig.VI.6, for the SP6C₆L injected oocytes the C & M samples represent 2 oocytes (tracks 7 & 8, respectively), and the S sample 4 oocytes (track 9). The *in vitro* translation products immunoprecipitated by antilysozyme expressed from SP6C₆L and SP6Lys are shown in track 5 and track 2, respectively. Tracks 1 & 10 are ^{14}C-labelled molecular weight marker proteins (see II.F.2), including lysozyme (14.3K). The figure is of a 7d exposure autoradiograph of the fluorographed gel (15% polyacrylamide with a cross-linking ratio of 40:0.2).

separated on a gel containing 15% polyacrylamide with a ratio of acrylamide to bisacrylamide of 40%:0.2%; this gel composition was used as it was hoped to provide resolution of polypeptides the size of prelysozyme and preC₆L. In contrast to the experiments which examined the expression of the C₆L fusion cDNA in oocytes (VI.B.3), the cytoplasmic injection of the synthetic SP6C₆L transcript clearly resulted in the expression of a protein, recognised by lysozyme antibodies, which migrated to the same position on the gel as the lysozyme in the molecular weight markers, M_r=14,300 (tracks 1 & 7); this is consistent with the expected size of the signal processed C₆L hybrid protein which contains only one less amino acid than lysozyme (see Fig.VI.1). This implies that the signal sequence has been cleaved from the SP6C₆L product in the oocyte. This suggestion is further supported by the fact that the polypeptide expressed from SP6C₆L on translation *in vitro* (track 6) is larger than the protein detected in the oocyte; this precursor, preC₆L, displays a mobility on SDS-PAGE consistent with a molecular weight of approx. M_r=15,500, which compares with a molecular weight of M_r=15,800 calculated from the amino acid sequence of preC₆L excluding the initiation methionine residue (see Fig.VI.1). Like the C₆₂L hybrid protein, the C₆L product is detected in the cytosol (Fig.VI.7 track 7) and vesicle (track 8) fractions of the oocyte but is not secreted from the oocytes (track 9). As the processed C₆L protein is associated with the vesicle fraction of the oocyte this indicates that the signal sequence plus the first six amino acids of prochymosin is sufficient to initiate the translocation of the signal-minus lysozyme *in vivo* in the oocyte. Furthermore the observed processing of the oocyte product relative to the full-length *in vitro* product strongly suggests that the preC₆L fusion protein encoded by SP6C₆L also contains a functional signal peptidase recognition site at which cleavage has occurred when the hybrid protein gains access to the ER in oocytes. Therefore the additional residues from the mature prochymosin protein contained in the larger preC₆₂L fusion are not essential for translocation.

Compared with the RNAs for both the hybrid proteins, injection of SP6Lys resulted in the detection of large amounts of immunospecific protein in the oocyte (Fig.VI.7 tracks 3-5), producing a product which had the same mobility as the ^{14}C-labelled lysozyme, M_r=14,300 (track 1), and was processed relative to the full-length precursor expressed *in vitro*, M_r=16,000 (track 2). In agreement with results obtained with lysozyme mRNA and lysozyme cDNA (see III.C), the lysozyme expressed from the synthetic RNA is secreted efficiently by the oocyte. In addition, as

before, the (processed) lysozyme detected within the oocyte is distributed about equally between the cytosolic and membrane fractions of the injected oocytes, and is not predominantly associated with the vesicles as with most foreign secretory proteins. However it is interesting to note that the C_6L expressed in the *Xenopus* oocyte also displays the anomalous fractionation pattern, previously only observed with lysozyme; in the experiment shown in Fig.VI.7 more of the C_6L detected in the oocyte is in the cytosolic fraction than in the membrane and vesicle fraction. Oocytes from the same batch were used to study the expression of $SPC_{62}L$ (Fig.VI.6) and these were processed alongside those injected with SPC_6L and SP6Lys; yet $C_{62}L$, which contains a larger region of prochymosin but the same lysozyme domain as C_6L, was detected primarily in the membrane and vesicles fraction, as described earlier. The compartmentation of $C_{62}L$ predominantly with the membrane fraction in the oocyte reflects the distribution within the oocyte of preprochymosin expressed from SP6PPChy RNA (see V.C and Fig.V.3)

VI.E. The Fusion Proteins C_6L and $C_{62}L$ are Processed on Translocation *In Vitro*

The experiments described in the previous section suggest that the signal sequence of preprochymosin functions as an autonomous unit in initiating translocation across the ER, and is not a 'protein-specific' signal capable only of directing the translocation of the native prochymosin protein. It was of interest to determine whether similar conclusions could be drawn from studies of the translocation of C_6L and $C_{62}L$ *in vitro*; would *in vitro* systems reflect the *in vivo* situation seen in the *Xenopus* oocyte? As noted earlier (VI.A.) it was not clear to what extent the translocation of proteins in reconstituted *in vitro* systems provided an accurate model for the translocation process *in vivo*; during the course of the work described in this chapter several reported results obtained with cell-free systems suggested that translocation *in vitro* may not necessarily mimic the mechanism of translocation of nascent polypeptides across the ER *in vivo* (see Introduction I.A.2). It was also proposed ultimately to characterise the translocation phenotype of signal sequence mutants of a secretory protein primarily by *in vitro* translation translocation assays to complement studies using the oocyte as an *in vivo* system. The advent of the SP6 *in vitro* transcription system to generate synthetic mRNAs which could be translated in cell-free systems made it possible to study the translocation *in vitro* of polypeptides encoded by specific cDNA constructs. It was therefore decided to determine if the

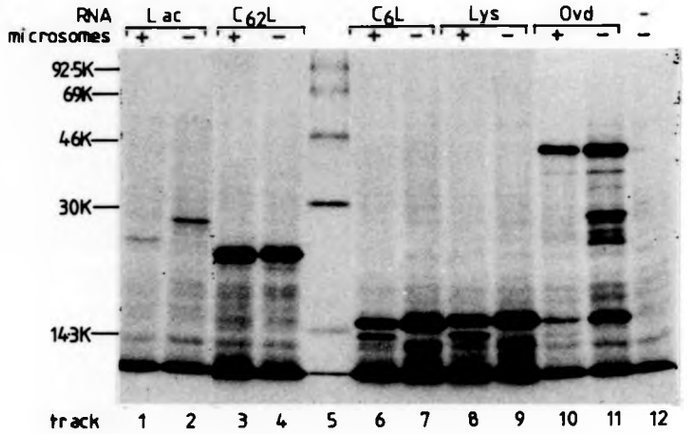


Figure VI.8 *In vitro* translation and translocation of precursor proteins expressed from SP6 RNAs

Capped SP6 RNAs encoding prelysozyme and the hybrid proteins, preC_6L & pre$C_{62}L$, (described in Fig.V.1) were translated (at 4ng/µl) in a wheat germ cell-free system in the absence (-) and presence (+) of dog pancreatic microsomes, as described in the Methods (II.E.). Two natural mRNAs, oviduct and lactogen, were also translated under the same conditions, except that the concentration of the lactogen mRNA was unknown. Aliquots (2µl) from each *in vitro* translation were analysed by SDS-PAGE on a 12.5% polyacrylamide gel (30% acrylamide:0.825% bisacrylamide) as follows: lactogen mRNA (tracks 1 & 2), SP6$C_{62}L$ (tracks 3 & 4), SP6C_6L (tracks 6 & 7), SP6Lys (tracks 8 & 9), oviduct (tracks 10 & 11). Track 12 is an aliquot from the translation in which no exogenous RNA was added and track 5 contains molecular weight marker proteins (II.F.2).

hybrid proteins preC_6L and pre$C_{62}L$ were also translocated and processed *in vitro* by translating the SP6C_6L and SP6$C_{62}L$ synthetic RNAs in a cell-free system in the presence of dog pancreatic microsomes. These experiments are described in this section and their results are discussed in VI.F.

The wheat germ cell-free translation system was chosen to assess the translocation of the fusion proteins *in vitro*. It was known that the m^{7}GG-capped SP6 C_6L and $C_{62}L$ RNAs were translated well in this system (see Fig.VI.6 & 7) and that the translation products of exogenous RNAs could be analysed by SDS-PAGE without immunoprecipitation. Furthermore it had been observed that when products from a reticulocyte lysate translation were electrophoresed without immunoprecipitation the migration of lysozyme and prelysozyme, expressed from oviduct mRNA, was distorted by the large amounts of globin present in the lysate (see Fig.IV.3).

Further aliquots of the m^{7}GG-capped SP6C_6L, SP6$C_{62}L$ and SP6Lys transcripts, prepared as described in Fig.V.1, were translated in a wheat germ cell-free system in the absence and presence of canine pancreatic microsomes (Methods II.E). Lactogen and oviduct mRNAs were also translated in the same system to provide controls for the translocation and processing activity of the dog microsomes in the wheat germ assay; translation products from these mRNAs had been shown to be translocated and processed by microsomes from the same batch in a reticulocyte lysate system (Fig.IV.3). Analysis of the proteins by SDS-PAGE showed that translation of both the natural and synthetic mRNAs had resulted in the synthesis of specific products (Fig.VI.8). In addition, when microsomes were included in the translations the expected processing of the lactogen and oviduct mRNA products was observed (tracks 1 & 10, respectively, see IV.C.); also in the presence of microsomes some processing of the precursors expressed by SP6$C_{62}L$ (track 3 & 4), SP6C_6L (tracks 6 & 7) and SP6Lys (tracks 8 & 9) had apparently occurred. Judged by the migration relative to the molecular weight marker proteins (track 5) the major proteins detected from the synthetic RNAs translated in the presence of microsomes are as follows: SP6$C_{62}L$, approx 22.4K and 21.5K; SP6C_6L, approx 14.8K and 13.8K; and from SP6Lys approx 15.3K and 14.0K. The prelysozyme and lysozyme expressed from the oviduct mRNA (tracks 10 & 11) were the same size as the the SP6Lys precursor and processed products; it is not clear why on this gel the processed lysozyme detected from the SP6Lys and oviduct RNAs do not display the same mobility as the ^{14}C-labelled lysozyme (14.3K) in the molecular weight markers (track 5). The lactogen

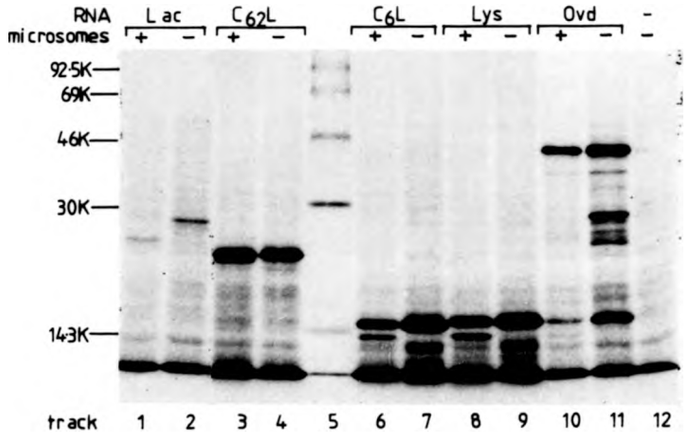


Figure VI.8 *In vitro* translation and translocation of precursor proteins expressed from SP6 RNAs

Capped SP6 RNAs encoding prelysozyme and the hybrid proteins, preC$_6$L & preC$_{62}$L, (described in Fig.V.1) were translated (at 4ng/μl) in a wheat germ cell-free system in the absence (-) and presence (+) of dog pancreatic microsomes, as described in the Methods (II.E.). Two natural mRNAs, oviduct and lactogen, were also translated under the same conditions, except that the concentration of the lactogen mRNA was unknown. Aliquots (2μl) from each *in vitro* translation were analysed by SDS-PAGE on a 12.5% polyacrylamide gel (30% acrylamide:0.825% bisacrylamide) as follows: lactogen mRNA (tracks 1 & 2), SP6C$_{62}$L (tracks 3 & 4), SP6C$_6$L (tracks 6 & 7), SP6Lys (tracks 8 & 9), oviduct (tracks 10 & 11). Track 12 is an aliquot from the translation in which no exogenous RNA was added and track 5 contains molecular weight marker proteins (II.F.2).

hybrid proteins preC$_6$L and preC$_{62}$L were also translocated and processed *in vitro* by translating the SP6C$_6$L and SP6C$_{62}$L synthetic RNAs in a cell-free system in the presence of dog pancreatic microsomes. These experiments are described in this section and their results are discussed in VI.F.

The wheat germ cell-free translation system was chosen to assess the translocation of the fusion proteins *in vitro*. It was known that the m^{7}GG-capped SP6 C$_6$L and C$_{62}$L RNAs were translated well in this system (see Fig.VI.6 & 7) and that the translation products of exogenous RNAs could be analysed by SDS-PAGE without immunoprecipitation. Furthermore it had been observed that when products from a reticulocyte lysate translation were electrophoresed without immunoprecipitation the migration of lysozyme and prelysozyme, expressed from oviduct mRNA, was distorted by the large amounts of globin present in the lysate (see Fig.IV.3).

Further aliquots of the m^{7}GG-capped SP6C$_6$L, SP6C$_{62}$L and SP6Lys transcripts, prepared as described in Fig.V.1, were translated in a wheat germ cell-free system in the absence and presence of canine pancreatic microsomes (Methods II.E). Lactogen and oviduct mRNAs were also translated in the same system to provide controls for the translocation and processing activity of the dog microsomes in the wheat germ assay; translation products from these mRNAs had been shown to be translocated and processed by microsomes from the same batch in a reticulocyte lysate system (Fig.IV.3). Analysis of the proteins by SDS-PAGE showed that translation of both the natural and synthetic mRNAs had resulted in the synthesis of specific products (Fig.VI.8). In addition, when microsomes were included in the translations the expected processing of the lactogen and oviduct mRNA products was observed (tracks 1 & 10, respectively, see IV.C.); also in the presence of microsomes some processing of the precursors expressed by SP6C$_{62}$L (track 3 & 4), SP6C$_6$L (tracks 6 & 7) and SP6Lys (tracks 8 & 9) had apparently occurred. Judged by the migration relative to the molecular weight marker proteins (track 5) the major proteins detected from the synthetic RNAs translated in the presence of microsomes are as follows: SP6C$_{62}$L, approx 22.4K and 21.5K; SP6C$_6$L, approx 14.8K and 13.8K; and from SP6Lys approx 15.3K and 14.0K. The prelysozyme and lysozyme expressed from the oviduct mRNA (tracks 10 & 11) were the same size as the the SP6Lys precursor and processed products; it is not clear why on this gel the processed lysozyme detected from the SP6Lys and oviduct RNAs do not display the same mobility as the ^{14}C-labelled lysozyme (14.3K) in the molecular weight markers (track 5). The lactogen

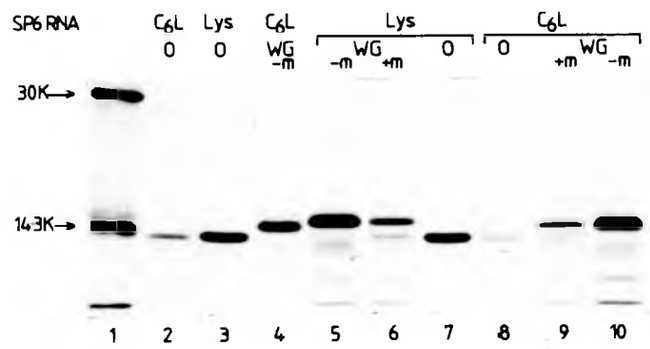


Figure VI.9 Comparison of the oocyte and *in vitro* translation products of SP6C$_6$L RNA
Capped SP6C$_6$L and SP6Lys RNA (see Fig.V.1) were expressed *in vivo* following injection into the cytoplasm of *Xenopus* oocytes (as described in Fig.VI.7) and were translated *in vitro* in the absence (-m) and presence (+m) of microsomes (see Fig.VI.8). Samples from the membrane fraction of the injected oocytes (O) and from the *in vitro* translations (WG) were immunoprecipitated with antilysozyme antibodies, and the precipitated proteins were analysed by SDS-PAGE on a 15% polyacrylamide (40:0.2 cross-linked) gel (as described in the Methods, II.F).

mRNA translation resulted in a polypeptide of approx. 26K (track 2), and when microsomes were included in the translation a 23K product is detected (track 1). Since the translation and translocation of the translation products of the synthetic RNAs had been effective in this experiment a closer examination was made of the processed and precursor proteins expressed from these RNAs *in vitro* in comparison to the proteins detected in oocytes following injection of the same transcripts.

VI.E.1. Comparison of the *in vitro* and *in vivo* translation products of SP6C$_6$L and SP6Lys RNA

It was of interest to compare the relative sizes of the translation products of the C$_6$L construct when expressed in the oocyte and *in vitro* in the presence of microsomes. Judging by migration on SDS-PAGE does the processing of the preC$_6$L species seen on translocation of the precursor *in vitro* result in a product of the same size as the C$_6$L protein detected in the oocyte? The enzyme signal peptidase does not show species specificity (see I.A.2) and it is anticipated that the site at which recognition and cleavage of the signal peptide takes place in a particular protein will be the same with the *Xenopus* and canine ER membranes. A comparison of the SP6C$_6$L products with those of SP6Lys *in vitro* and in the oocyte will provide further information on the processing of the C$_6$L fusion protein. To compare the *in vitro* and *in vivo* products of the lysozyme cDNA gene and the C$_6$L fusion gene equivalent aliquots from the wheat germ translations of capped SP6Lys and SP6C$_6$L were immunoprecipitated with antilysozyme antibodies, together with samples from oocytes injected with the same RNAs (described in VI.D). The immunoprecipitated proteins were separated by electrophoresis on a SDS-polyacrylamide gel (15% acrylamide with a cross linking ratio of 40:0.2), shown in Figure VI.9.

From the sequence of the C$_6$L and Lys cDNAs the primary translation product of SP6Lys, prelysozyme, is 146 residues excluding the initiation methionine (16,100 molecular weight), which is 3 amino acids longer than the corresponding preC$_6$L precursor (15.8K) (see Fig.VI.1). The removal of the signal peptide (17 residues, excluding the Met) from prelysozyme results in mature lysozyme which is 129 amino acids (14.3K); however, the signal sequence of prochymosin present in preC$_6$L is only 15 amino acids and hence cleavage of the prochymosin signal peptide from preC$_6$L results in the 128 residue C$_6$L protein (14.2K). The pattern of protein bands seen in Fig.VI.9 are consistent with this picture. Firstly it appears that the cleaved C$_6$L protein detected in the oocyte is the same size as the

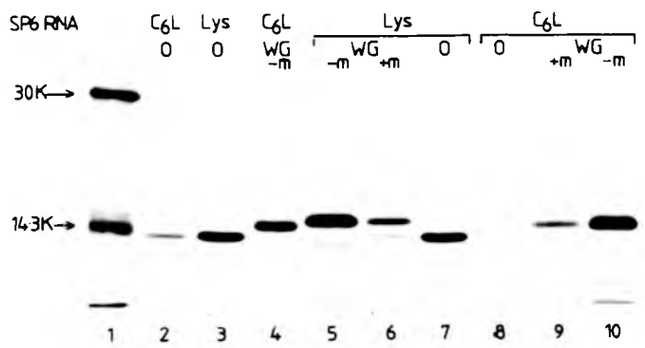


Figure VI.9 Comparison of the oocyte and *in vitro* translation products of SP6C$_6$L RNA
Capped SP6C$_6$L and SP6Lys RNA (see Fig.V.1) were expressed *in vivo* following injection into the cytoplasm of *Xenopus* oocytes (as described in Fig.VI.7) and were translated *in vitro* in the absence (-m) and presence (+m) of microsomes (see Fig.VI.8). Samples from the membrane fraction of the injected oocytes (O) and from the *in vitro* translations (WG) were immunoprecipitated with antilysozyme antibodies, and the precipitated proteins were analysed by SDS-PAGE on a 15% polyacrylamide (40:0.2 cross-linked) gel (as described in the Methods, II.F).

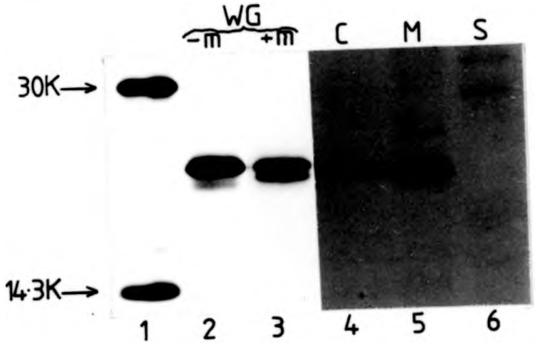
mRNA translation resulted in a polypeptide of approx. 26K (track 2), and when microsomes were included in the translation a 23K product is detected (track 1). Since the translation and translocation of the translation products of the synthetic RNAs had been effective in this experiment a closer examination was made of the processed and precursor proteins expressed from these RNAs *in vitro* in comparison to the proteins detected in oocytes following injection of the same transcripts.

VI.E.1. **Comparison of the *in vitro* and *in vivo* translation products of SP6C$_6$L and SP6Lys RNA**

It was of interest to compare the relative sizes of the translation products of the C$_6$L construct when expressed in the oocyte and *in vitro* in the presence of microsomes. Judging by migration on SDS-PAGE does the processing of the preC$_6$L species seen on translocation of the precursor *in vitro* result in a product of the same size as the C$_6$L protein detected in the oocyte? The enzyme signal peptidase does not show species specificity (see I.A.2) and it is anticipated that the site at which recognition and cleavage of the signal peptide takes place in a particular protein will be the same with the *Xenopus* and canine ER membranes. A comparison of the SP6C$_6$L products with those of SP6Lys *in vitro* and in the oocyte will provide further information on the processing of the C$_6$L fusion protein. To compare the *in vitro* and *in vivo* products of the lysozyme cDNA gene and the C$_6$L fusion gene equivalent aliquots from the wheat germ translations of capped SP6Lys and SP6C$_6$L were immunoprecipitated with antilysozyme antibodies, together with samples from oocytes injected with the same RNAs (described in VI.D). The immunoprecipitated proteins were separated by electrophoresis on a SDS-polyacrylamide gel (15% acrylamide with a cross linking ratio of 40:0.2), shown in Figure VI.9.

From the sequence of the C$_6$L and Lys cDNAs the primary translation product of SP6Lys, prelysozyme, is 146 residues excluding the initiation methionine (16,100 molecular weight), which is 3 amino acids longer than the corresponding preC$_6$L precursor (15.8K) (see Fig.VI.1). The removal of the signal peptide (17 residues, excluding the Met) from prelysozyme results in mature lysozyme which is 129 amino acids (14.3K); however, the signal sequence of prochymosin present in preC$_6$L is only 15 amino acids and hence cleavage of the prochymosin signal peptide from preC$_6$L results in the 128 residue C$_6$L protein (14.2K). The pattern of protein bands seen in Fig.VI.9 are consistent with this picture. Firstly it appears that the cleaved C$_6$L protein detected in the oocyte is the same size as the

(a)



(b)

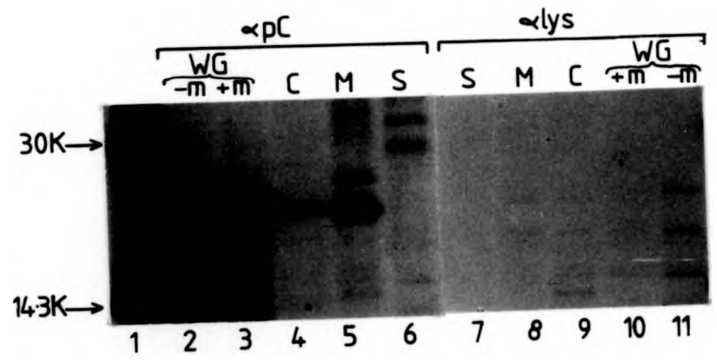


Figure VI.10 Comparison of the products of SP6$C_{\alpha 2}$L RNA in the oocyte and *in vitro*
Capped SP6$C_{\alpha 2}$L RNA (see Fig.V.1) was expressed in the oocyte (see Fig.VI.6) and translated in a wheat germ cell-free system (WG) in the absence (-m) and presence (+m) of microsomes (see Fig.VI.8). Samples from the oocyte fractions (C, M & S as described in Fig.VI.6) and from the *in vitro* translations were immunoprecipitated either with antiprochymosin (αpC - tracks 2-6) or with antilysozyme (αlys - tracks 7-11). The immunoprecipitated proteins were separated by SDS-PAGE on a 12.5% polyacrylamide gel. The dried, fluorographed gel was exposed to X-ray sensitive film for (a) 14days (b) 28 days. For clarity the tracks in (a) were exposed for different durations during printing of the negative. Track 1 contains molecular weight marker proteins (II.F.2).

processed product resulting from translocation of the preC_{α}L precursor by microsomes *in vitro* (compare tracks 8 & 9). Similarly the lysozyme expressed from SP6Lys in the oocyte displays the same mobility as the processed *in vitro* SP6Lys product (tracks 7 & 6). Tracks 2-5 of Fig.VI.9 show the relative sizes of the precursor and processed proteins expressed from SP6C_{α}L and SP6Lys; as anticipated the preC_{α}L product (track 4) is clearly smaller than prelysozyme (track 5), and the processed C_{α}L (track 2) appears have a marginally greater mobility than lysozyme (track 3). The apparent size reduction resulting from signal-processing of the precursors is greater with prelysozyme (track 6) than with preC_{α}L (track 9); this is in agreement with the cleavage of the 17 residue lysozyme signal peptide, compared with the loss of 15 amino acids on removal of the prochymosin signal sequence of C_{α}L. From this gel the mobilities of the different polypeptides, relative to the molecular weight markers (track 1), are consistent with approx. sizes of - preC_{α}L, 14.5K; C_{α}L, 13.5K; prelysozyme, 14.8K and lysozyme, 13.6K. However if the precursor and processed products expressed by SP6Lys are used as 16.1K and 14.3K markers; then judged by comparison with the migration of these prelysozyme and lysozyme polypeptides the size of preC_{α}L is 15.5K and C_{α}L 14.1K.

It is interesting to note that similar amounts of prelysozyme and preC_{α}L are expressed from the synthetic RNAs *in vitro* (see Fig.VI.8, tracks 7 & 9 and Fig.VI.9, tracks 4 & 5); in contrast when the same RNAs were expressed in the oocyte, antilysozyme immunoprecipitated significantly greater amounts of protein following injection of SP6Lys than when oocytes were injected with SP6C_{α}L (Fig.VI.7, but see VI.F.3). The proportion of the prelysozyme and preC_{α}L precursor polypeptides which are processed in the presence of the microsomes is also similar.

VI.E.2. Immunoprecipitation of the $C_{\alpha 2}$L fusion protein expressed *in vitro* and in the oocyte

The proteins expressed from the capped SP6$C_{\alpha 2}$L RNA *in vitro* and in the oocyte were immunoprecipitated with antiprochymosin antibodies (αpC) and analysed by SDS-PAGE. Fig.VI.10a shows that, as with the SP6C_{α}L and SP6Lys products, when SP6$C_{\alpha 2}$L is translated *in vitro* the pre$C_{\alpha 2}$L precursor (track 2, estimated molecular weight 22.3K) is processed in the presence of microsomes to a protein (approx. 21.2K, track 3) which displayed the same mobility as the products detected in the oocyte from the same RNA (tracks 4 & 5). However when duplicate samples were immunoprecipitated with antilysozyme (αlys) very little protein was

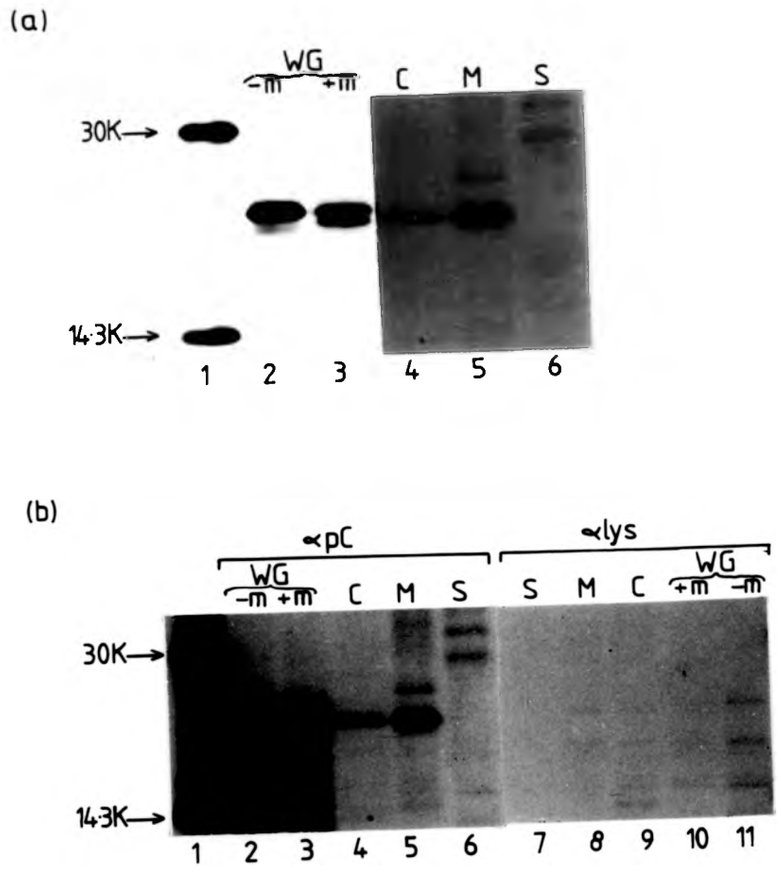


Figure VI.10 Comparison of the products of SP6$C_{\alpha 2}L$ RNA in the oocyte and *in vitro*
Capped SP6$C_{\alpha 2}L$ RNA (see Fig.V.1) was expressed in the oocyte (see Fig.VI.6) and translated in a wheat germ cell-free system (WG) in the absence (-m) and presence (+m) of microsomes (see Fig.VI.8). Samples from the oocyte fractions (C, M & S as described in Fig.VI.6) and from the *in vitro* translations were immunoprecipitated either with antiprochymosin (αpC - tracks 2-6) or with antilysozyme (αlys - tracks 7-11). The immunoprecipitated proteins were separated by SDS-PAGE on a 12.5% polyacrylamide gel. The dried, fluorographed gel was exposed to X-ray sensitive film for (a) 14days (b) 28 days. For clarity the tracks in (a) were exposed for different durations during printing of the negative. Track 1 contains molecular weight marker proteins (II.F.2).

processed product resulting from translocation of the pre$C_{\alpha}L$ precursor by microsomes *in vitro* (compare tracks 8 & 9). Similarly the lysozyme expressed from SP6Lys in the oocyte displays the same mobility as the processed *in vitro* SP6Lys product (tracks 7 & 6). Tracks 2-5 of Fig.VI.9 show the relative sizes of the precursor and processed proteins expressed from SP6$C_{\alpha}L$ and SP6Lys; as anticipated the pre$C_{\alpha}L$ product (track 4) is clearly smaller than prelysozyme (track 5), and the processed $C_{\alpha}L$ (track 2) appears have a marginally greater mobility than lysozyme (track 3). The apparent size reduction resulting from signal-processing of the precursors is greater with prelysozyme (track 6) than with pre$C_{\alpha}L$ (track 9); this is in agreement with the cleavage of the 17 residue lysozyme signal peptide, compared with the loss of 15 amino acids on removal of the prochymosin signal sequence of $C_{\alpha}L$. From this gel the mobilities of the different polypeptides, relative to the molecular weight markers (track 1), are consistent with approx. sizes of - pre$C_{\alpha}L$, 14.5K; $C_{\alpha}L$, 13.5K; prelysozyme, 14.8K and lysozyme, 13.6K. However if the precursor and processed products expressed by SP6Lys are used as 16.1K and 14.3K markers; then judged by comparison with the migration of these prelysozyme and lysozyme polypeptides the size of pre$C_{\alpha}L$ is 15.5K and $C_{\alpha}L$ 14.1K.

It is interesting to note that similar amounts of prelysozyme and pre$C_{\alpha}L$ are expressed from the synthetic RNAs *in vitro* (see Fig.VI.8, tracks 7 & 9 and Fig.VI.9, tracks 4 & 5); in contrast when the same RNAs were expressed in the oocyte, antilysozyme immunoprecipitated significantly greater amounts of protein following injection of SP6Lys than when oocytes were injected with SP6$C_{\alpha}L$ (Fig.VI.7, but see VI.F.3). The proportion of the prelysozyme and pre$C_{\alpha}L$ precursor polypeptides which are processed in the presence of the microsomes is also similar.

VI.E.2. Immunoprecipitation of the $C_{\alpha 2}L$ fusion protein expressed *in vitro* and in the oocyte

The proteins expressed from the capped SP6$C_{\alpha 2}L$ RNA *in vitro* and in the oocyte were immunoprecipitated with antiprochymosin antibodies (αpC) and analysed by SDS-PAGE. Fig.VI.10a shows that, as with the SP6$C_{\alpha}L$ and SP6Lys products, when SP6$C_{\alpha 2}L$ is translated *in vitro* the pre$C_{\alpha 2}L$ precursor (track 2, estimated molecular weight 22.3K) is processed in the presence of microsomes to a protein (approx. 21.2K, track 3) which displayed the same mobility as the products detected in the oocyte from the same RNA (tracks 4 & 5). However when duplicate samples were immunoprecipitated with antilysozyme (αlys) very little protein was

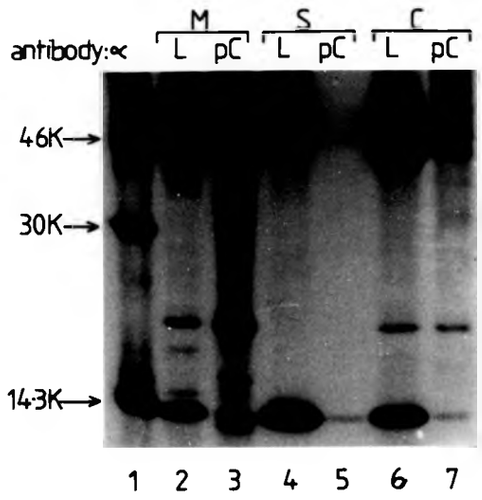


Figure VI.11 Immunoprecipitation of proteins expressed in the oocyte from SP6C$_{\alpha 2}$L and oviduct mRNA

Xenopus oocytes were coinjected with SP6C$_{\alpha 2}$L and oviduct mRNA, then cultured and fractionated described previously (Fig.VI.6 & II.H). Duplicate samples representing 1 oocyte from the M (tracks 2 & 3) and C (tracks 6 & 7) fractions and 2½ oocytes of the S (tracks 4 & 5) fraction were immunoprecipitated either with antilysozyme (αL) or antiprochymosin (αpC) as shown; and the immunoprecipitated proteins were analysed by SDS-PAGE (as described in the Methods).

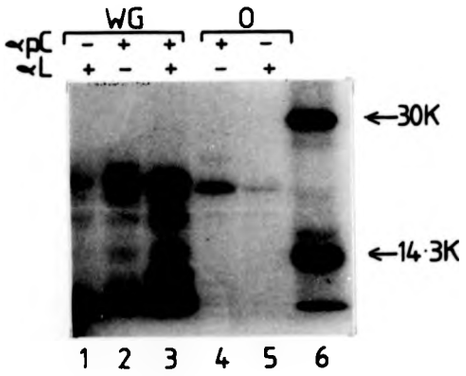


Figure VI.12 Comparison of SP6C$_{\alpha 2}$L translation products immunoprecipitated by antilysozyme and antiprochymosin

Replicate samples from an *in vitro* translation of SP6C$_{\alpha 2}$L (WG) (II.E.2) and from oocytes (O, membrane fraction) injected with the same RNA (Fig.VI.6) were immunoprecipitated with antilysozyme (αL) and/or antiprochymosin (αpC) as shown. The immunoprecipitated proteins were separated on a 12.5% polyacrylamide gel, as described in II.F.2.

detected (Fig.VI.10b) from either the *in vitro* translation (tracks 10 & 11) or oocyte fractions (track 7-9). As noted earlier, when C$_{\alpha 2}$L was expressed in oocytes from nuclear injected pTK$_2$C$_{\alpha 2}$L+ DNA there was a difference in the amount of protein detected in the oocyte fractions when αlys and αpC were used for immunoprecipitation (see VI.C). Further investigations were made of the immunoprecipitation by αlys and αpC of the translation products of SP6C$_{\alpha 2}$L *in vitro* and in the oocyte.

In the experiment shown in Figure VI.11 oocytes were coinjected with oviduct mRNA and capped SP6C$_{\alpha 2}$L RNA, and duplicate samples from the fractionated oocytes were immunoprecipitated separately with αlys and αpC. This experiment established that there was little recognition of wild type lysozyme, expressed from the oviduct mRNA, by the αpC antibodies; since, compared with the lysozyme precipitated by αlys (Figure VI.11 tracks 2, 4 & 6), very little protein was detected at the position of lysozyme in samples immunoprecipitated with αpC (tracks 3, 5 & 7). Approximately equal amounts of C$_{\alpha 2}$L are immunoprecipitated by αlys and αpC from the cytosol fraction of the oocytes (tracks 6 & 7, respectively), but more C$_{\alpha 2}$L was precipitated in the vesicle fraction by αpC (track 3) than by αlys (track 2). As in other experiments no C$_{\alpha 2}$L was detected in medium surrounding the injected oocytes by either antibody (tracks 4 & 5).

Figure VI.12 shows replicate aliquots from a wheat germ translation of capped SP6C$_{\alpha 2}$L immunoprecipitated with αlys (track 1) and αpC (track 2) separately, and together (track 3). Whilst it appears that αpC is almost as efficient as the two antibodies together at precipitating full-length preC$_{\alpha 2}$L (apparent molecular weight 22K), this precursor was poorly recognised by αlys alone. Curiously, when both antibodies were used together another much smaller major product (approx. 13.5K) is also detected, but this protein is not seen when either of the antibodies are used alone.

The experiments described here will be discussed in the next section (VI.F) together with other results concerning the expression of the C$_{\alpha}$L and C$_{\alpha 2}$L hybrid proteins which have been described in this chapter.

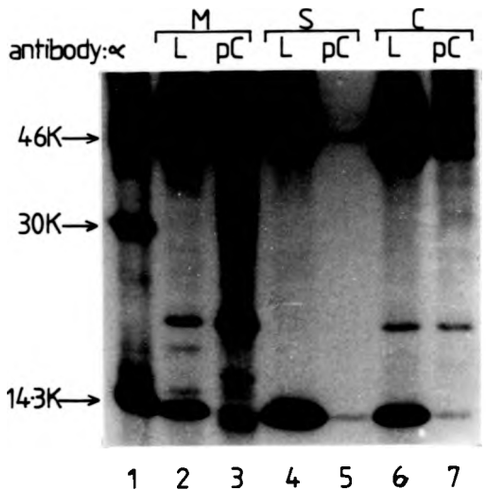


Figure VI.11 Immunoprecipitation of proteins expressed in the oocyte from SP6C$_{\alpha 2}$L and oviduct mRNA

Xenopus oocytes were coinjected with SP6C$_{\alpha 2}$L and oviduct mRNA, then cultured and fractionated described previously (Fig.VI.6 & II.H). Duplicate samples representing 1 oocyte from the M (tracks 2 & 3) and C (tracks 6 & 7) fractions and 2M oocytes of the S (tracks 4 & 5) fraction were immunoprecipitated either with antilysozyme (αL) or antiprochymosin (αpC) as shown; and the immunoprecipitated proteins were analysed by SDS-PAGE (as described in the Methods).

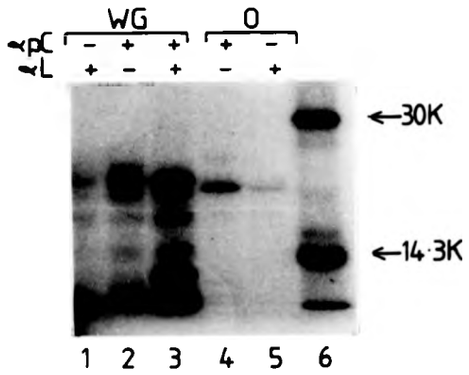


Figure VI.12 Comparison of SP6C$_{\alpha 2}$L translation products immunoprecipitated by antilysozyme and antiprochymosin

Replicate samples from an *in vitro* translation of SP6C$_{\alpha 2}$L (WG) (II.E.2) and from oocytes (O, membrane fraction) injected with the same RNA (Fig.VI.6) were immunoprecipitated with antilysozyme (αL) and/or antiprochymosin (αpC) as shown. The immunoprecipitated proteins were separated on a 12.5% polyacrylamide gel, as described in II.F.2.

detected (Fig.VI.10b) from either the *in vitro* translation (tracks 10 & 11) or oocyte fractions (track 7-9). As noted earlier, when C$_{\alpha 2}$L was expressed in oocytes from nuclear injected pTK$_2$C$_{\alpha 2}$L+ DNA there was a difference in the amount of protein detected in the oocyte fractions when αlys and αpC were used for immunoprecipitation (see VI.C). Further investigations were made of the immunoprecipitation by αlys and αpC of the translation products of SP6C$_{\alpha 2}$L *in vitro* and in the oocyte.

In the experiment shown in Figure VI.11 oocytes were coinjected with oviduct mRNA and capped SP6C$_{\alpha 2}$L RNA, and duplicate samples from the fractionated oocytes were immunoprecipitated separately with αlys and αpC. This experiment established that there was little recognition of wild type lysozyme, expressed from the oviduct mRNA, by the αpC antibodies; since, compared with the lysozyme precipitated by αlys (Figure VI.11 tracks 2, 4 & 6), very little protein was detected at the position of lysozyme in samples immunoprecipitated with αpC (tracks 3, 5 & 7). Approximately equal amounts of C$_{\alpha 2}$L are immunoprecipitated by αlys and αpC from the cytosol fraction of the oocytes (tracks 6 & 7, respectively), but more C$_{\alpha 2}$L was precipitated in the vesicle fraction by αpC (track 3) than by αlys (track 2). As in other experiments no C$_{\alpha 2}$L was detected in medium surrounding the injected oocytes by either antibody (tracks 4 & 5).

Figure VI.12 shows replicate aliquots from a wheat germ translation of capped SP6C$_{\alpha 2}$L immunoprecipitated with αlys (track 1) and αpC (track 2) separately, and together (track 3). Whilst it appears that αpC is almost as efficient as the two antibodies together at precipitating full-length preC$_{\alpha 2}$L (apparent molecular weight 22K), this precursor was poorly recognised by αlys alone. Curiously, when both antibodies were used together another much smaller major product (approx. 13.5K) is also detected, but this protein is not seen when either of the antibodies are used alone.

The experiments described here will be discussed in the next section (VI.F) together with other results concerning the expression of the C$_{\alpha}$L and C$_{\alpha 2}$L hybrid proteins which have been described in this chapter.

VI.F. Discussion

VI.F.1 Eukaryotic signal sequences function as autonomous units

The aim of the work described in this chapter was to determine if the cleavable eukaryotic signal sequence functions as an autonomous unit in effecting the translocation of secretory proteins. The approach taken was to construct a fusion gene which encoded a chimaeric protein in which the signal peptide of prelysozyme was replaced by the signal sequence of preprochymosin, then examine the translocation of this hybrid in *Xenopus* oocytes and also *in vitro*. It had been shown that signal-processed lysozyme, secreted by oocytes injected with oviduct mRNA, was not re-exported nor sequestered into the ER when injected into fresh oocytes (Lane *et al*,1980). The initial fusion gene I constructed, C_6L, encoded a hybrid protein in which the signal sequence and first 6 amino acids of prochymosin replaced the signal peptide and the first 7 residues of lysozyme (section B). Early work on the expression of this fusion protein in the oocyte from injected pTK_2C_6L+ plasmid DNA indicated that this region of preprochymosin was not sufficient to direct the translocation of the signal-minus lysozyme (VI.B); this led to the construction of another fusion gene, $C_{62}L$, which contained the same 122 codons of mature lysozyme but with a larger preprochymosin domain encoding the signal sequence plus the first 62 N-terminal residues of prochymosin (section VI.C). The experiments described in section D of this chapter showed that when both the chimaeric proteins $C_{62}L$ and C_6L are expressed in *Xenopus* oocytes following injection of capped synthetic RNA, they are translocated into the ER membrane judging by their association with the membrane and vesicle fraction of the oocyte and by their processing relative to the full-length precursor translation product. The observed cleavage of $preC_6L$ and $preC_{62}L$ by microsomal membranes *in vitro* (section E) also demonstrated that both hybrids contain the information necessary to initiate translocation across the ER. Therefore, since the signal sequence plus first 6 amino acids of prochymosin, present in $preC_6L$ can direct the translocation of signal-minus lysozyme this implies the signal sequence of eukaryotic secretory proteins does function as a discrete unit which is not protein-specific. When this work was initiated the functional autonomy of a eukaryotic signal sequence in its role in protein translocation had not been demonstrated. During the course of this work there were reports published by other groups which also provided evidence that the signal

sequence functions as a self-contained unit in initiating translocation across the ER.

Early in 1984 Lingappa *et al* reported that the leader peptide plus first 5 residues from the *E.coli* periplasmic protein pre-β-lactamase joined (via 4 amino acids derived from a linker) to chimpanzee α-globin resulted in a hybrid protein which was translocated across microsomal membranes *in vitro* and processed. Hence it appeared the cleavable signal sequence of a prokaryotic secretory protein could convert a eukaryotic cytoplasmic protein to a translocated protein in an *in vitro* eukaryotic system. These findings provided further support for the similarity of the process of protein translocation across the plasma membrane in prokaryotes and across the ER membrane in eukaryotes. However they were inconsistent with earlier reports that the C-terminal region of β-lactamase is required for its membrane transport in a bacterial system (Koshland & Botstein,1980 and 1982). This further highlights that comparison of, or inference from, results from prokaryotic systems to the process of protein translocation in eukaryotes should be made with caution, and *vica versa*. Two other reports published in 1984 used fusion proteins to define the region functioning as the signal sequence in proteins which were not processed on translocation. One publication concerned work carried out in this laboratory to identify the signal sequence of ovalbumin (see Tabe *et al*,1984 in Appendix); as noted earlier this is the only known eukaryotic secretory protein which does not contain a transient signal sequence (see I.E). This work firstly showed by deletion analysis that residues 20-41 of ovalbumin appeared to contain the information for translocation of this protein in oocytes; to complement these deletion experiments this region of ovalbumin was then fused to chimpanzee α-globin and it was found that the resulting chimaeric protein was translocated when expressed in oocytes from injected DNA. The transmembrane influenza virus protein neuraminidase also does not contain a cleavable signal sequence: in 1984 Bos *et al* reported that the N-terminal 40 amino acids of neuraminidase could direct the translocation of signal-minus haemagglutinin when the neuraminidase/ haemagglutinin fusion gene was expressed in CV1 cells. Both these results imply that uncleaved signal sequences are probably also functionally discrete units, although it is less easy to define the limits of the signal sequence in such cases.

As noted earlier (VI.A) at the time this work was initiated the functioning of the signal sequence as a self-contained unit was inherent in the concepts of the Signal Hypothesis for protein translocation, as it

currently stood. The finding that the signal sequence did indeed appear to contain the information to initiate the translocation of proteins across the ER was, therefore, in concordance with the concept of the co-translational translocation of proteins across this membrane. It was also consistent with the idea that, as the first step towards the translocation of eukaryotic secretory or membrane proteins, SRP in the cytoplasm recognises and interacts with the signal sequence of the nascent protein as it emerges from the translating ribosome (see I.A.2). It was originally envisaged that this interaction of the SRP with the nascent signal sequence resulted in an arrest in the translation of the particular protein which was released only on the subsequent interaction of this SRP/incomplete nascent protein/ribosome complex with the ER membrane via the SRP Receptor; then translocation of the nascent polypeptide chain occurred as its synthesis was continued. However, as a result of more recent data (*ie* Meyer,1985) there is now doubt whether this SRP-induced arrest of translation is an integral part of the mechanism of translocation *in vivo* (discussed in I.A.2). It could be argued that in the absence of a SRP-induced translational-arrest the structure and conformation of the nascent polypeptide chain beyond the signal sequence could be important for translocation *in vivo*. However the results I have presented and the others discussed above suggest that sequences beyond the signal peptide (except perhaps the first few amino acids of the mature protein) are not required to initiate translocation *in vivo* or *in vitro*; and that in unprocessed translocated proteins the same function resides in a discrete stretch of amino acids, which is of a similar length to cleaved signal peptides.

In the Theory of Topogenic Sequences, proposed by Blobel(1980), the signal sequence for translocation across the ER is just one subset of a wider class of 'signal sequences', defined as being responsible for initiating the translocation of proteins across translationally-competent cellular membranes (see Introduction I.B.1). The leader sequence of imported mitochondrial proteins and the transit peptide of proteins transported to the chloroplast are other subsets of the 'signal sequence' class of topogenic sequence; although the composition of these pre-sequences is distinct from the signal peptides of proteins translocated across the ER (see I.A.4). It is postulated in Blobel's theory that all these 'signal sequences' function as discrete, positively-acting units in initiating protein translocation. In contrast to translocation across the ER, the import of proteins to the mitochondrion and chloroplast has clearly been shown to be a post-translational process;

hence it is conceivable that interaction between the transient pre-sequence and the mature portion of the protein could be an integral part of the translocation process. However in 1984 two publications by Hurt *et al* established that the transient leader sequence of a mitochondrial matrix protein contained all the information for the correct post-translational localisation of a foreign cytosolic protein to this compartment of the mitochondrion; these workers were able to target the normally cytosolic murine DHFR to mitochondria by fusing it to N-terminal domains of yeast cytochrome oxidase subunit IV, a matrix protein. There is now good evidence from *in vitro* and *in vivo* studies that mitochondrial leader sequences and the chloroplast transit peptide can function as an autonomous unit in directing the translocation of proteins to the respective organelles (see Introduction I.B.4). The evidence is largely derived from examing the compartmentation of various chimaeric proteins encoded by fusion genes containing the 'signal sequence' of a translocated organelle protein joined to coding sequences of a cytosolic protein.

Hence the available evidence suggests that eukaryotic 'signal sequences' do contain all the information necessary to initiate the translocation of proteins across the translocationally-competent intracellular membranes; although the mechanism of protein translocation across each type of membrane must differ. In the context of the proposed signal sequence mutagenesis studies (outlined in I.E) a block in translocation caused by a mutation in the signal peptide region of preprochymosin can, therefore, be considered a consequence of alterations in the structure of the signal sequence *per se*, and not due to disruption of any interactions between the signal peptide and the mature part of the secretory protein.

VI.F.2. A functional signal sequence is not sufficient to direct the secretion of a protein

Although the preC$_{\alpha}$L and preC$_{\alpha 2}$L precursors were translocated and the signal sequence cleaved in the *Xenopus* oocyte they were apparently not secreted, since no immunospecific products could be detected in medium surrounding oocytes microinjected with the fusion genes contained in the pTK$_2$ vector or with synthetic RNA encoding the hybrid proteins (see sections B, C & D). Such a phenotype of signal sequence processing but not secretion has previously been observed for mutant secretory immunoglobulin light chains. Work by Mosmann and his coworkers found that in two cell clones derived from a myeloma cell line (MOPC315) the

immunoglobulin λ light chains were not secreted and were structurally altered, compared to the λ chains secreted by the parent cell line; both the mutant λ chains synthesized *in vivo* were found to be processed relative to the precursor polypeptides translated *in vitro* from polyribosomes isolated from each cell line (Mosmann *et al*,1979, Mosmann & Williamson,1980). The structural alteration was localised in each case to a different cyanogen bromide peptide and each is thought to be a point mutation. The same phenotype of translocation and signal-processing but no secretion was also observed when these mutant λ chains were expressed in *Xenopus* oocytes (Valle *et al*,1983). In the work mentioned in F.1, carried out in this laboratory, on the segregation in oocytes of deletion-mutants of ovalbumin (Tabe *et al*,1984, see Appendix), it was found that two truncated mutant ovalbumins were not secreted but were associated with the membrane and vesicle fraction of the oocyte and contained N-glycosylation (an event which takes place in the ER, see I.A.3). In addition ovalbumin-globin fusion proteins containing the N-terminal region from these deletion mutants were also not secreted but had translocated completely through the ER membrane, judging by the full protection of the membrane-associated ovalbumin/globin protein from digestion by exogenous proteases.

It appears, therefore, from the results described above that although a protein contains a signal sequence which directs its translocation across the ER membrane, with signal cleavage occurring where appropriate, this is not sufficient for subsequent secretion of the protein. These data imply that the role of the signal sequence is in the translocation of proteins as the first step along the secretory pathway and that other 'signals' or structural features are necessary to achieve passage of a protein beyond the ER to the cell surface for secretion. As noted in VI.F.1, within the context of the Signal Hypothesis and theory of topogenic sequences the role of the signal sequence is to initiate the translocation of the nascent chain across the ER. The data described in this chapter and the results of others, discussed above, are in agreement with the assignation of this role to the signal sequence.

In 1980 the theory of topogenic sequences proposed by Blobel had postulated a class of 'sorting sequences' which would act after translocation to direct proteins from the RER to their correct subcellular location or to the plasma membrane for export (I.B.). As discussed in the Introduction (I.A.3, I.B.5) it has also been proposed that once a protein is sequestered within the ER through the action of its signal sequence it is then routed to the cell surface, and no other specific sorting signals

are required for its subsequent secretion (*ie* see review of Kelly,1985); only the retention of the translocated protein at a particular location along the secretory pathway is envisaged to require the positive action of specific topogenic sequences. During recent years evidence has accumulated for discrete positively-acting topogenic sequences, of the type envisaged by Blobel, being involved in directing the subcellular compartmentation of proteins (discussed in I.B). In addition, during the course of the work of this thesis several lines of investigation were published which supported the view that in the absence of other topogenic information a protein translocated within the ER will ultimately be exported by the cell (*ie* Gething & Sambrook,1982; Wiedman *et al*,1984; Poruchinsky *et al*,1985, discussed in I.B). However, experimental data presented in this thesis and by others (described above) show that the action of a functional signal sequence is not sufficient to achieve the secretion of a protein. What then is the reason or basis for the non-secretion of the C_6L and $C_{62}L$ chimaeric proteins which translocate the ER membrane?

It is evident that neither of the constituent secretory proteins from which preC_6L and pre$C_{62}L$ are constructed contain a 'stop-transfer sequence' which could halt the translocation of the hybrid proteins (Blobel,1980; see I.B.5); it is therefore unlikely that the failure to detect a secreted C_6L or $C_{62}L$ product is a consequence of the chimaeric proteins being anchored in the membrane. Although this possibility cannot be rigorously excluded at this stage since the work described in this chapter did not examine whether the preprochymosin/lysozyme fusion proteins were wholly translocated across ER membranes (see later, F.4); however the complete translocation of C_6L and $C_{62}L$ is predicted as preprochymosin and prelysozyme are 'translocationally permissive' amino acid sequences.

Particular mutations in the signal sequence of certain secretory proteins have been found to result in mutant proteins in which the initiation of translocation into the ER is not prevented but signal peptide processing is inhibited; these translocated precursor proteins tend to remain associated with the ER membrane and are not secreted (Hortin & Boime,1980 and 1981; Schauer *et al*,1985). In these cases it seems the absence of signal processing is the basis for non-secretion, possibly due to this preventing the completion of the translocation of the mutant protein across the ER membrane; since, although the mutant signal peptide apparently no longer contains the functional wild-type signal peptidase cleavage site, some precursors are processed at a

substitute site and the abberantly cleaved proteins are slowly secreted. Since C_6L and $C_{62}L$ are efficiently processed on translocation the absence of cleavage of the signal peptide is not the reason that these proteins are not secreted.

It is likely that the block in the secretion of C_6L and $C_{62}L$ takes place after their translocation and processing in the ER. The stretches of amino acids which comprise any topogenic 'sorting sequences' which are necessary to direct the passage of secretory proteins from the ER are likely to be present in C_6L and $C_{62}L$ in the lysozyme derived domain, which is all but the N-terminal 7 residues of mature lysozyme (see VI.B and VI.C). The lack of secretion of C_6L seems curious since the primary sequence of the signal-processed C_6L product is nearly identical to that of mature lysozyme (see Fig.VI.1), differing in only the 6 amino acids at the N-terminus; yet lysozyme was found to be efficiently secreted when expressed in oocytes from cDNA or synthetic RNA (see VI.D). However it was noted that of the 7 codons of mature lysozyme which were replaced by prochymosin sequences in the construction of C_6L, one was a cysteine residue which in native lysozyme forms a disulphide bridge with another Cys at the C-terminus (residue 127), one of the four disulphide bridges found in lysozyme (Imoto *et al*,1973). Hence at the most only three correct disulphide bridges can be formed in the lysozyme domain of C_6L in the ER. The absence of this N-terminal Cys in C_6L therefore means the hybrid is unable to adopt a tertiary structure which resembles lysozyme and this is postulated to be the basis for the lack of secretion of C_6L. Since this Cys, the sixth residue of lysozyme, is also missing in $C_{62}L$ the consequent perturbation of the conformation of the lysozyme domain in this larger hybrid is also likely to be the reason why it was not exported from *Xenopus* oocytes. This conformational difference between native lysozyme and the lysozyme portion of the preprochymosin/lysozyme hybrids also has other implications on the interpretation of the experimental results of this chapter, these are discussed later (F.3). As discussed earlier, in other instances the presence of a functional signal sequence has been found to be sufficient to direct the secretion of a foreign or mutant protein. This fact indicates that it is perhaps the alteration of the physical properties *per se* of the secretory protein chimaeras, compared with those of the native constituent proteins, which prevents their passage along the secretory pathway; rather than perturbation of protein structure at the secondary or tertiary level causing the functional disruption of positively acting 'sorting sequences'. Therefore it is envisaged that the transit of a protein along the

secretory pathway depends not only on the presence of the appropriate, positively-acting topogenic sequences, but also on the polypeptide chain having a structure compatible with its passage through the different compartments on the secretory pathway. It has been found that the addition of a functional signal sequence will only direct the translocation of a cytosolic protein across membranes if the amino acid sequence of the cytosolic protein is 'permissive' for membrane translocation (see Garoff,1985). Following translocation other properties will be required if a protein is to be secreted, such as solubility in the interior of the ER and Golgi complex.

V.F.3. Compartmentation and immunoprecipitation of C_6L and $C_{62}L$ expressed in oocytes

It was surprising when $C_{62}L$ was first expressed in oocytes, from $pTK_2C_{62}L$+ DNA, that prochymosin antibodies apparently recognised $C_{62}L$ better than antilysozyme (see Fig.VI.5), since this hybrid contains 122 of the 129 residues of lysozyme but only 62 of the 365 amino acids of prochymosin. This original observation of the differential precipitation by antilysozyme (αlys) and antiprochymosin (αpC) of $C_{62}L$ led to the combined use of both antibodies to immunoprecipitate $C_{62}L$ when it was subsequently expressed from $SP6C_{62}L$ RNA (Fig.VI.6). The loss of the cysteine residue in the lysozyme domain of the hybrid proteins, described in VI.F.2, with the consequent effect on the protein conformation, is a strong candidate for the observation that the $C_{62}L$ chimaera, expressed *in vitro* and in the oocyte, is recognised better in immunoprecipitations by αpC than by αlys (V.C.2 & E.2). As the conformation adopted by the lysozyme domain of $C_{62}L$ is likely to be quite distinct from that of native lysozyme, many of the epitopes which are recognised by the antibodies raised against lysozyme may no longer be exposed for cross-reaction when $C_{62}L$ is immunoprecipitated with αlys. Since only αlys could be used to immunoprecipitate C_6L, there is a concern that the poor cross-reaction of the αlys antibodies with the mutant lysozyme domain of the hybrid proteins may result in incomplete immunoprecipitation of C_6L present in oocyte fractions. This possibility should be borne in mind when considering the compartmentation of the fusion proteins in the oocyte, and means a conclusive interpretation cannot be made. If antisera were raised against denatured lysozyme in which the disulphide bridges are reduced, this could prove more efficient at quantitatively immunoprecipitating both C_6L and $C_{62}L$.

In terms of the compartmentation of $C_{\epsilon 2}L$ expressed in the oocyte from synthetic RNA, it was consistently found when either αpC alone or αpC plus αlys antibodies were used for immunoprecipitation that significantly more $C_{\epsilon 2}L$ protein was detected in the membrane fraction than in the cytosol (see VI.D. and Figs.VI.6, 10 & 11). If these results represent quantitative immunoprecipitation of $C_{\epsilon 2}L$ present in the oocyte, the fractionation of $C_{\epsilon 2}L$ primarily with the membrane fraction reflects the normal distribution of preprochymosin and most other segregated foreign proteins expressed in *Xenopus* oocytes (Colman & Morser,1979; see III.F and V.C). As noted earlier (III.C), work carried out by Cutler had already found that the secretory protein lysozyme showed an anomalous fractionation pattern when expressed in *Xenopus* oocytes from mRNA (Cutler *et al*,1981; Cutler,1982). Following microinjection, when the oocytes were labelled with ^{35}S-methionine for 24h then fractionated, approximately equal amounts of lysozyme were found in the cytosol and vesicle fractions (ie see Fig.VI.11). Further studies showed that approximately 15% of the lysozyme synthesized and signal-processed could not be 'chased' out of the oocyte by newly-synthesized unlabelled lysozyme, unlike most sequestered foreign proteins. The work described in this thesis has confirmed that when lysozyme is expressed in oocytes from cDNA or synthetic SP6 RNA as much processed lysozyme is detected in the cytosol as in the membrane fraction (ie Figs. III.3 & VI.7). When $SP6C_{\epsilon}L$ was expressed in oocytes it was interesting to find that the $C_{\epsilon}L$ hybrid protein also consistently displayed this anomalous fractionation (ie see Fig.VI.7), with approximately equal amounts of the signal-processed $C_{\epsilon}L$ detected in the cytosol and the membrane fraction. As noted above it is recognised that poor recognition by αlys of the lysozyme domain of $C_{\epsilon}L$ may mean the $C_{\epsilon}L$ present in the oocyte fractions is not quantitatively immunoprecipitated in these experiments. However further evidence that $C_{\epsilon}L$ expressed in the oocyte is present in the cytosol and vesicle fractions in roughly equal amounts, came from an experiment carried out by Alan Colman subsequent to my own work. He injected oocytes with $SP6C_{\epsilon}L$ then labelled and fractionated them as usual; however the samples of the fractions were analysed without immunoprecipitation directly by SDS-PAGE and fluorography. The autoradiograph of this gel showed a relatively strong protein band the size of $C_{\epsilon}L$ present in approximately equal amounts amongst the proteins in both the cytosol and membrane fractions, but not amongst the secreted proteins. The basis for the unusual compartmentation in *Xenopus* oocytes of lysozyme and the $C_{\epsilon}L$ hybrid, but not apparently $C_{\epsilon 2}L$, could be the subject of further

investigations. For example, how much of the preprochymosin protein is required in a preprochymosin/lysozyme chimaera in order to cause the hybrid to display the more usual distribution in the oocyte and fractionate predominantly with the membranes? Is the relatively high association of signal-processed lysozyme and C_6L with the cytosolic fraction specific to lysozyme, or is it simply a consequence of the small size of these segregated polypeptides?

It was also found that the relative amounts of $C_{62}L$ protein detected in the cytosol and membrane fractions of oocyte expressing the $C_{62}L$ fusion gene was not the same when αlys and αpC were used as antibodies for immunoprecipitation (see Figs.VI.5, 10, 11). This differential immunoprecipitation suggests that the conformation of the $C_{62}L$ protein detected in the cytosol and membrane fractions of the oocyte is different. If it is assumed that the nascent pre$C_{62}L$ expressed in the oocyte translocates into the ER, and the subsequent detection of processed $C_{62}L$ in the cytosol fraction results from disruption of microsomes during homogenisation of the injected oocytes; then the change in the conformation adopted by the 'cytosol-associated' $C_{62}L$ will have taken place in the buffer used in the homogenisation process (see II.H). The relative amounts of $C_{62}L$ immunoprecipitated by αpC and αlys from injected oocytes was not, however, consistent in different experiments (compare Figs.VI.5, 10 & 11). In one experiment (shown in Fig.VI.11) approximately equal amounts of the chimaeric protein were detected by αlys and αpC in the cytosolic fraction, but in the membrane fraction more protein was immunoprecipitated by αpC than by αlys. Yet in another experiment in which $C_{62}L$ was also expressed in oocytes from SP6$C_{62}L$ RNA (Fig.VI.10), virtually no protein was immunoprecipitated by αlys from either the cytosol or vesicle fractions, although the use of αpC resulted in clearly detected $C_{62}L$ protein bands in both these fractions. This variation perhaps reflects further changes in the conformation of $C_{62}L$ which occur during storage of the oocyte fractions, which alter the subsequent recognition of the hybrid protein by the lysozyme antibodies.

VI.F.4. Future work

It would be of interest to investigate further the translocation, processing and subcellular localisation of C_6L and $C_{62}L$. Amino acid sequence analysis of the processed forms of the hybrid proteins detected in the oocyte would determine whether or not the preprochymosin signal peptide has been correctly cleaved on translocation of C_6L and $C_{62}L$; a similar analysis could be made of the cleaved hybrid proteins produced *in*

vitro in the presence of microsomes. As mentioned earlier (VI.F.2) the complete translocation of the preprochymosin/lysozyme fusion proteins across the ER membrane is predicted, but the work of this chapter did not experimentally test this. If on further study the processed membrane-associated C_6L and $C_{62}L$ produced both in the oocyte and on translocation *in vitro* were found to be resistant to digestion by exogenous proteases, this would demonstrate the complete polypeptide chain was translocated across the ER membrane. Using carbonate extraction of membranes to discriminate between integral and peripheral proteins would provide further information on the association of the hybrid proteins with the oocyte membranes (see Tabe *et al* in Appendix). Density gradient fractionation of homogenised oocytes injected with SP6 RNAs encoding the fusion proteins and preprochymosin and lysozyme would enable some comparison to be made of the intracellular location of the hybrid proteins and their constituent secretory proteins. For the experiments described above one would need to be confident of quantitatively detecting the chimeric proteins in the oocyte fractions. It may be necessary to raise antisera against denatured lysozyme containing no disulphide bridges to provide antibodies which efficiently immunoprecipitate C_6L and $C_{62}L$.

The use of microinjection of cells in tissue culture and immunofluorescence to visualise the foreign proteins could provide more information on the subcellular localisation of C_6L and $C_{62}L$ along the secretory pathway: whilst capped SP6 RNAs could be used for these experiments it has been found that these must be polyadenylated for expression in cells in culture (Drummond *et al*,1985).

VI.G. Summary

Two preprochymosin/lysozyme fusion genes were constructed in which the signal peptide and first 7 amino acids of prelysozyme were replaced by different regions of preprochymosin cDNA. The first fusion gene, C_6L, contained the signal sequence and first 6 amino acids of prochymosin. When this construct contained in the pTK_2 expression vector was initially injected into *Xenopus* oocytes no C_6L protein was detected, although in one subsequent experiment a product was detected from the C_6L fusion cDNA. The early results led to the construction of a second fusion gene, $C_{62}L$, in which a preprochymosin fragment encoding the signal sequence plus the first 62 residues of prochymosin was joined to the same 122 codons of mature lysozyme. The injection of $pTK_2C_{62}L+$ into oocytes resulted in only poor expression of a protein which displayed a mobility on SDS-PAGE

consistent with the expected size of the $C_{\epsilon 2}L$ chimaeric protein. The $C_{\epsilon}L$ and $C_{\epsilon 2}L$ fusion genes were subsequently cloned into the SP_{64} vector, and synthetic RNAs encoding the hybrid proteins were transcribed *in vitro* in the presence of m^7GG capping dinucleotide. The m^7GG-capped RNAs were found to be translated efficiently when injected into the cytoplasm of *Xenopus* oocytes, and in the wheat germ cell-free system. The injection into oocytes of the synthetic RNAs, $SP6C_{\epsilon}L$ and $SP6C_{\epsilon 2}L$, resulted in the synthesis of readily detectable amounts of both the $C_{\epsilon}L$ and $C_{\epsilon 2}L$ proteins which were associated with the membrane and cytosolic fractions of the oocyte but were not secreted. However, when oocytes were injected with synthetic RNA encoding prelysozyme, the lysozyme expressed was compartmented and secreted in a similar manner to lysozyme expressed in oocytes following the injection of lysozyme 'natural' mRNA and cDNA (described in Chapter III). The products detected in the oocyte following injection of $SP6C_{\epsilon}L$ and $SP6C_{\epsilon 2}L$ corresponded to the expected size of the signal-processed $C_{\epsilon}L$ and $C_{\epsilon 2}L$ proteins, and on SDS-PAGE displayed an increase in mobility relative to the full-length precursors produced on translation of the same capped RNAs *in vitro*; the increased migration was consistent with the cleavage of the preprochymosin signal peptide. The association of the processed $C_{\epsilon}L$ and $C_{\epsilon 2}L$ proteins with oocyte membranes showed the chimaeric proteins translocated the ER membrane and acted as a substrate for signal peptidase on the lumenal face of the ER membrane.

The translocation of $preC_{\epsilon}L$ and $preC_{\epsilon 2}L$ was also studied *in vitro* by translating the synthetic RNAs encoding the hybrid proteins in a wheat germ cell-free system in the presence of dog pancreatic microsomes. These *in vitro* experiments demonstrated that both $preC_{\epsilon}L$ and $preC_{\epsilon 2}L$ were translocated and processed in the presence of microsomes; and the cleaved proteins produced *in vitro* displayed the same migration on SDS-PAGE as the products expressed from the same RNAs in the oocyte. Hence the *in vitro* translocation studies reflected the situation seen *in vivo* in the oocyte and indicated that both $C_{\epsilon 2}L$ and $C_{\epsilon}L$ contained the information which specifies the translocation and processing of these hybrid proteins. Therefore the signal sequence and first 6 amino acids of preprochymosin is sufficient to direct the translocation of signal-minus lysozyme into the ER membrane *in vivo* and *in vitro*. However, it appears that the presence of a functional signal sequence is not sufficient to direct the subsequent secretion of the protein *in vivo*. These results provided the first demonstration that the cleavable signal sequence of eukaryotic secretory proteins functions as an autonomous unit in initiating the translocation of the nascent polypeptide chain.

VII.A. CHAPTERS III-VI CONCLUSIONS

VII.A.1. The *Xenopus* Oocyte and *In Vitro* Assay Systems

The experimental work described in the Results & Discussion chapters (III-VI) has employed the *Xenopus* oocyte as an *in vivo* system in which to study the segregation and secretion of eukaryotic proteins; in part consolidating further the usefulness of this system for such studies, but also raising some of its limitations.

Chapter III showed that cloned cDNA genes encoding the eukaryotic secretory proteins lysozyme, prochymosin and ovalbumin could be expressed in oocytes by inserting the cDNA into specific expression vectors so that the cDNA was under the control of a viral promoter which is active in *Xenopus* oocytes. A vector which contained the thymidine kinase promoter region from Herpes Simplex Virus was chosen as a suitable vector for the expression of cDNAs in oocytes, this was designated pTK_2. Following injection of specific pTK_2 constructs into the nucleus of *Xenopus* oocytes, the foreign secretory proteins encoded by the cDNA were found to be sequestered in membranes within the oocyte and secreted into the medium, in a manner analagous to the compartmentation of the same proteins expressed from cytoplasmically injected mRNA. In particular prochymosin was secreted by *Xenopus* oocytes injected with preprochymosin cDNA (PPChy) contained in the pTK_2 vector, and within the oocyte the prochymosin fractionated predominately with membranes and vesicles with little detected in the cytosolic fraction. This distribution is the same as that previously observed for most foreign secretory proteins expressed in *Xenopus* oocytes from injected mRNA (eg Colman & Morser, 1979). When oocytes were injected with the pTK_2Lys+ plasmid which encoded prelysozyme, approximately equal amounts of apparently signal-processed lysozyme were detected in the cytosol and the membrane fractions of the oocyte; this reflected the anomalous distribution of lysozyme seen when *Xenopus* oocytes are injected with prelysozyme mRNA (Cutler,1982). In theory the use of the pTK_2 vector is not limited to the expression of secretory proteins in *Xenopus* oocytes, but could also be employed to express the translation products encoded by other cDNAs or by specific recombinant DNA constructs. Indeed the pTK_2 vector was used by Davey *et al*(1985) to study the localisation in oocytes of protein products encoded by influenza virus nucleoprotein cDNA and several deletion mutants derived from this cDNA, with a view to identifying the karyophilic signal in this protein. However, whilst the wild-type

secretory proteins lysozyme, prochymosin and ovalbumin were expressed at readily detectable levels in the oocyte following nuclear injection of the corresponding pTK_2 construct; it was found that certain DNA constructs derived from these cDNA genes resulted in either no detectable polypeptide product in the oocyte, or gave low and variable levels of the foreign protein. As described in Chapter III, two constructs derived from the preprochymosin cDNA were also cloned into the pTK_2 vector and injected into *Xenopus* oocytes; one of these (PChy) encoded a methionyl-prochymosin product and the other (Chy) methionyl-chymosin. As neither of these derivatives contains the signal sequence of preprochymosin it is expected that the PChy and Chy translation products will be compartmented in the cytoplasm of the oocyte. The construct pTK_2PChy+ only produced detectable levels of prochymosin protein in batches of oocytes in which the full-length PPChy DNA gave particularly good expression of preprochymosin; the small amount of PChy product which was detected was localised exclusively in the cytosol of the oocyte. However, no chymosin-specific protein was ever detected following injection of pTK_2Chy+ into *Xenopus* oocytes. It was also found that neither of the preprochymosin/lysozyme fusion genes, C_6L and $C_{62}L$ (described in Chapter VI), gave good expression in the oocyte of the corresponding chimaeric proteins following injection of pTK_2C_6L+ or $pTK_2C_{62}L$+ DNA. Similarly certain ovalbumin cDNA deletion mutants which were injected into *Xenopus* oocytes in the pTK_2 vector gave low or undetectable levels of the mutant ovalbumin protein; described in Tabe *et al*,1984 (see Appendix).

The SP6 system developed by Melton *et al*(1984) first became available during the course of this work, and provided an alternative means of studying the compartmentation in oocytes of the translation products of wild-type cDNA, or of specific DNA constructs for which there is no natural mRNA. By cloning particular DNA fragments into the SP6 vector it is then possible to generate *in vitro* large amounts of specific transcripts of the DNA insert. The work of Krieg & Melton(1984) demonstrated that if a 5' cap is added post-transcriptionally to such SP6 transcripts, using guanyltransferase, the synthetic RNA is expressed on injection into the cytoplasm of *Xenopus* oocytes. Experiments described in this thesis (Chapters V and VI) showed that if the *in vitro* transcription was carried out in the presence of the capping dinucleotide m^7GpppG, these m^7GG-capped SP6 transcripts were also active in *Xenopus* oocytes. The proteins expressed in the oocyte from synthetic SP6 RNAs were compartmented in the same way as those expressed from the corresponding natural mRNA or from cDNA; ie. the secretory proteins prelysozyme and

preprochymosin expressed in oocytes from SP6 RNAs were segregated into membranes, processed and secreted. It was found that injection of capped synthetic SP6 RNA into oocytes resulted in the production of larger amounts of the foreign protein than was obtained following nuclear injection of the corresponding pTK_2 DNA. This meant that the PChy, $C_{\epsilon}L$ and $C_{\epsilon 2}L$ translation products, which showed poor and variable expression in oocytes from injected DNA, were clearly and consistently detected following injection of the corresponding SP6 transcripts. The injection of RNA into the cytoplasm of *Xenopus* oocytes is also technically easier than the nuclear injection of DNA. These factors mean that the injection of SP6 transcripts synthesized *in vitro* from specific DNA inserts is the preferred method of expressing cloned cDNA genes, and constructs derived from them, in *Xenopus* oocytes. The work described in Chapters V and VI showed that the injection of synthetic SP6 RNAs into oocytes can be used to characterise the compartmentation of polypeptides encoded by hybrid genes or derivatives of a secretory protein cDNA. However, it was also found that not all SP6 transcripts which are capable of producing a translation product will necessarily result in the expression of detectable levels of the specific protein on injection into oocytes. When SP6Chy RNA, encoding signal-minus chymosin, was injected into *Xenopus* oocytes no chymosin was detected. In terms of using *Xenopus* oocytes as a system in which to study the compartmentation of eukaryotic proteins it is possible that derivatives of secretory proteins which cannot translocate into ER membranes may be unstable as 'miscompartmentalised' proteins in the cytoplasm of the oocyte, and will never be detected due to their rapid degredation. It is also possible that in certain SP6 RNAs the sequences upstream from the initiation AUG may be unfavourable for translation in the oocyte. The absence of a detectable *in vivo* translation product from a particular SP6 RNA could also result from instability of the synthetic transcript in the oocyte. These situations are probably more likely to arise when the synthetic RNA is not transcribed from a DNA template which is exactly complementary to natural mRNA sequences; for example, in transcripts of mutated cDNAs, or those of inserts containing sequences derived from synthetic oligonucleotides.

The advent of the SP6 *in vitro* transcription system also enables the protein products encoded by specific DNA constructs to be analysed using eukaryotic *in vitro* translation systems. In Chapters V & VI it was shown that when each of the six SP6 RNAs synthesized (including SP6Chy) was translated in a wheat germ cell-free system this resulted in readily detectable amounts of the appropriate polypeptides, even for SP6PChy and

SP6Chy transcripts in which the nucleotide sequences flanking the initiation AUG are 'unfavourable' in terms of the consensus sequence derived by Kozak (Kozak,1983b; see III.E.2). It was also found there was no necessity to specifically immunoprecipitate the protein expressed from the exogenous SP6 RNA prior to analysis of the *in vitro* translation product by SDS-PAGE; although in the absence of immunoprecipitation RNase digestion of ^{32}P-labelled transcripts was required to avoid a high background of radioactivity on the gel. The above factors mean that initially the translation product encoded by a particular synthetic transcript can conveniently be examined using *in vitro* translation systems.

Experiments in which SP6 RNAs encoding prelysozyme, preC$_{\epsilon}$L and preC$_{\epsilon 2}$L were translated *in vitro* in the presence of microsomal vesicles (VI.E) demonstrated that the signal sequence present in these precursor polypeptides could direct the translocation of the nascent polypeptide *in vitro*; in agreement with the data obtained from *Xenopus* oocyte *in vivo* assays. These findings pave the way for an extension of the use of *in vitro* translation-translocation assays to cover the study of the translation products encoded by SP6 RNAs transcribed from specific DNA constructs, particularly those for which there is no natural mRNA counterpart. These experiments would be analogous to those carried out in the past with authentic eukaryotic mRNAs; for example, work described in IV.C used protease resistance assays to demonstrate that both the prochymosin proteins encoded by the preprochymosin mRNA were fully translocated within microsomal vesicles *in vitro*. However, it was found that the conditions required to assay for protease resistance is not the same for all translocated proteins; the precise conditions need to be determined empirically for the particular polypeptide to be studied. It is also anticipated that it will be possible to use SP6 RNAs in 'depletion and reconstitution' *in vitro* translocation studies, similar to those carried out by Walter & Blobel(1981a and b), Gilmore *et al*(1982a and b) and Meyer *et al*(1982a and b) (see I.A.2). In these experiments specific protein components are extracted from the *in vitro* translation-translocation system and the effect this has on the translocation of nascent polypeptides is characterised. It can also be determined how the translocation activity is affected by the addition of particular extracts to the depleted system.

Hence the segregation of proteins encoded by cloned cDNAs for secretory proteins, and by derivatives of these cDNAs, can be assayed in *Xenopus* oocytes and *in vitro* through the use of synthetic SP6 RNAs. *Xenopus* oocytes provide an *in vivo* system in which not only translocation, but also secretion of the polypeptides can be studied; although there is a problem of the possible instability of 'miscompartmentalised' proteins. Degredation of newly-synthesized proteins does not appear to be a drawback when using *in vitro* translation-translocation assays. These *in vitro* systems allow a closer examination to be made of the process of translocation of nascent polypeptides across ER-derived microsomal membranes; although no information can be gained on the subsequent intracellular localisation of the translocated polypeptides.

VII.A.2. **The Function and Autonomy of Eukaryotic Signal Sequences**

Experiments with preprochymosin mRNA, described in III.D and Chapter IV, showed that preprochymosin behaves like a 'classical' eukaryotic secretory protein with a cleaved N-terminal signal peptide. Firstly, when preprochymosin mRNA was injected into *Xenopus* oocytes this resulted in the expression of a protein, immunoprecipitated by prochymosin antisera, which was segregated into membranes and secreted by the oocyte; this protein comigrated with authentic prochymosin. On closer examination it was found, unexpectedly, that when preprochymosin mRNA was expressed either in oocytes or *in vitro* two proteins were detected which showed close migration on SDS-PAGE, and both proteins of the doublet were precipitated by antibodies raised against calf prochymosin. Compared with the primary translation products expressed from the preprochymosin mRNA *in vitro* (approx. 42.5K and 43K), the prochymosin proteins segregated and secreted by oocytes were smaller (approx. 41.5K and 42K); and it is likely this processing is due to cleavage of a N-terminal signal peptide on translocation of the preprochymosin precusors *in vivo*. This is in accord with the observation that preprochymosin cDNA clones encode a polypeptide with 16 amino acids N-terminal to the known amino acid sequence of prochymosin (Harris *et al*,1982; Noir *et al*,1982); this N-terminal extension is similar in structure to the signal peptides of other eukaryotic secretory proteins (Watson,1984). The segregation of preprochymosin in oocytes was supported by the results of *in vitro* translation-translocation assays, which showed that if microsomes were present during translation both the preprochymosin precursors encoded by the mRNA were translocated

completely within these vesicles and processed. Thus it appears that the two primary translation products of preprochymosin mRNA contain a transient signal sequence which directs the translocation of nascent preprochymosin across ER membranes *in vitro* and *in vivo*, with cleavage of the signal peptide occurring on translocation. The translation of preprochymosin polyA RNA, extracted from the abomasum of unweaned calves, had previously been reported to produce only one prochymosin-specific protein (Nishimori *et al*,1981; Harris *et al*,1982; Moir *et al*,1982; Nicholson & Jones,1984; McConnell *et al*,1984). It is suggested that the two electrophoretically distinct preprochymosins encoded by the mRNA have slightly different amino acid sequences, and represent the translation products of allelic chymosin genes; one is probably preprochymosin B and the other may be preprochymosin A (Foltmann *et al*,1977 and 1979).

Work discussed in III.E and V.B confirmed that the N-terminal signal sequence of preprochymosin is essential for translocation of the nascent protein. These experiments compared the compartmentation in *Xenopus* oocytes of polypeptides expressed from full-length preprochymosin cDNA (PPChy) and from the construct PChy, derived from PPChy cDNA, which encodes methionyl-prochymosin without the signal peptide region. When PChy was expressed in oocytes from injected DNA (pTK$_2$PChy+) or SP6PChy RNA, the prochymosin protein detected was localised exclusively in the cytoplasm and showed no association with membranes. Whereas in the same experiments the pTK$_2$PPChy+ DNA or SP6PPChy RNA expressed a protein which was segregated, processed and secreted, like the prochymosins expressed following injection of preprochymosin mRNA.

At the outset of the work covered in this thesis there was an open question as to whether the signal sequence of eukaryotic secretory proteins contained all the information to direct the translocation of the nascent protein across the ER membrane. The work described in Chapter VI was concerned with investigating the functional autonomy of eukaryotic signal sequences. The question addressed was whether the cleaved signal peptide region of a eukaryotic secretory protein was a functionally self-contained unit which, on transfer to a different signal-sequence-minus polypeptide, could direct the translocation of the resultant hybrid protein.

The two fusion genes which were constructed from preprochymosin and prelysozyme cDNA encoded chimaeric polypeptides in which the signal sequence plus the first 7 amino acids of lysozyme were replaced by different N-terminal regions of preprochymosin. In the preC$_6$L chimaera

the signal peptide and first 6 amino acids of prochymosin are joined to the lysozyme domain; whilst the preC$_{\epsilon 62}$L hybrid contains the 'pre' sequence plus the first 62 residues of prochymosin. When the C$_{\epsilon}$L and C$_{\epsilon 62}$L fusion gene products were expressed in *Xenopus* oocytes they were apparently segregated in the membranes of the oocyte, but were not secreted. The fusion proteins detected within the oocyte (in the membrane and the cytosol fractions) were processed relative to their respective primary translation products, obtained by translation of SP6 RNAs *in vitro*. This indicated that *in vivo* both the preC$_{\epsilon}$L and preC$_{\epsilon 62}$L nascent polypeptides gain access to the lumenal side of the ER membrane where the signal peptide is cleaved by signal peptidase. *In vitro* translation-translocation studies provided further evidence to support the hypothesis that the C$_{\epsilon}$L and C$_{\epsilon 62}$L gene products contain the information specifying the translocation and processing of these hybrid proteins. Translation of SP6C$_{\epsilon}$L and SP6C$_{\epsilon 62}$L RNAs *in vitro* in the presence of microsomes resulted in the processing of preC$_{\epsilon}$L and preC$_{\epsilon 62}$L, producing polypeptides of the same size as those detected in *Xenopus* oocytes from the same RNAs.

Hence it appears that the signal sequence plus the first 6 residues of prochymosin is sufficient to direct the translocation of signal-minus lysozyme into the ER membrane. These results indicate that the signal peptide region of preprochymosin does function as an autonomous unit, and that specific interactions between the signal sequence and the mature portion of the protein are not important for translocation of the nascent polypeptide. The functional autonomy of the signal sequence is in agreement with the concept of co-translational translocation of polypeptides across the ER membrane, as contained in the Signal Hypothesis, and also provides support for some of the postulates of the Theory of Topogenic Sequences (Blobel,1980; Walter *et al*,1984; discussed in I.A.2, I.B and VI.F).

Although the chimaeric precursors are segregated into membranes and processed, it was not possible to detect *in vivo* any of the C$_{\epsilon}$L and C$_{\epsilon 62}$L proteins in the medium surrounding the injected oocytes; yet the constituent preprochymosin and prelysozyme precursors are processed and secreted, following injection of the appropriate cDNA or SP6 RNA. This observation indicates that the role of the signal sequence is in initiating the translocation of the nascent polypeptide as the first step along the secretory pathway, and that other features of the polypeptide chain are responsible for its subsequent secretion. The presence of a functional signal peptide in a polypeptide is clearly insufficient to ensure a protein is routed through the subsequent stages of the secretory

pathway, following its translocation into the ER. It is probable that the non-secretion of the C_6L and $C_{62}L$ hybrids is not a consequence of the absence of positively-acting determinants which are required for intracellular protein sorting after the action of the signal sequence. Instead it is suggested that the signal-processed C_6L and $C_{62}L$ proteins cannot be secreted since their physical properties *per se* are incompatible with their further passage along the secretory pathway, because the gross conformation adopted by the lysozyme and/or prochymosin domains of the chimaera is markedly different from that in the native constituent proteins. The results with the preprochymosin/lysozyme fusion proteins do not accord with the theory that all proteins which translocate the ER membrane are routed to the cell-surface for export, unless they are retained or divert en route by the action of positively-acting topogenic sequence (ie Kelly,1985).

VII.B. TOWARDS *IN VITRO* MUTAGENESIS OF THE SIGNAL PEPTIDE OF PREPROCHYMOSIN

VII.B.1 Objectives

It has been mentioned several times in this thesis that the long term objective of the project was to introduce mutations into the signal sequence of a eukaryotic secretory protein; then study the translocation of such mutants *in vivo*, using *Xenopus* oocytes, and *in vitro* (ie see I.E). It was hoped this would provide information on the molecular structure of the signal peptide necessary for the translocation process. Due to the constraint of time it was not possible for me to fulfil this objective, but in this section I will discuss some of the practical and conceptual considerations involved in carrying out such a study.

It takes little analysis of published signal sequences (as compiled by Watson,1984) to conclude that there is great variation in the amino acid structure of the transient signal peptides of eukaryotic proteins which translocate the ER (as discussed in I.B.3). Yet each signal peptide distinguishes the nascent polypeptide for export from the cytoplasm, and apparently functions as a self-contained unit which interacts with a common translocation machinery. As part of the process of translocation across the ER membrane seems to involve the signal peptide interacting with at least one specific protein component (SRP), there must be constraints on the structure of the signal sequence to enable these specific ligand-receptor interactions to occur. It is particularly interesting to compare the products of multigene families which have

evolved from a single gene, in order to see what amino acid changes have been 'permitted' in these related signal peptides without destroying their function of initiating translocation across the ER. For example, comparing the signal peptides of six human α-interferons - A,B,C,D,F,G (Watson,1984), there is a common amino acid at 9 positions which are scattered throughout the signal sequence (22 residues in total, excluding the initiation Met). At the other 13 positions two or three different types of residue are found; the alternative amino acids are as diverse as Trp/Pro/Ser (position -20), or more similar such as Leu/Val (positions -2 and -13).

Theoretical considerations have led to the proposal of certain 'rules' governing the structure required for a functional signal sequence (ie Von Heijne,1985; discussed in I.B.3). Several groups have experimentally characterised the effect of mutations in the signal sequence region of exported prokaryotic proteins (reviewed in Benson *et al*,1985, Oliver,1985a; discussed in I.B.2), and these studies have drawn some consensus of the conformational requirements for a prokaryotic leader peptide. To what extent will studies on the effect of mutagenising the signal sequence of a eukaryotic secretory protein agree with the theoretical calculations, or with work on prokaryotic secretory proteins? Can the experimental observations with eukaryotic signal sequence mutants support or extend current theories on the successive stages in the translocation of nascent proteins across the ER membrane (discussed in I.A.2)?

In the prospective study it was initially intended to generate a range of mutations in the signal peptide region of cDNA encoding a eukaryotic secretory protein, and determine how effectively these mutant signal sequences functioned in the translocation of the nascent precursor. This would build up a picture of the effect of different amino acid substitutions on the maintenance of signal sequence function. This could lead later to a more detailed analysis of the flexibility at certain positions in the signal peptide.

The work described in Chapters III-VI examined the segregation of two wild-type secretory proteins with cleaved N-terminal signal peptides, preprochymosin and prelysozyme. Many of these experimental results were discussed in the context of the proposed signal sequence mutagenesis studies. From these studies, summarised in the previous section (VII.A.), it was decided that preprochymosin would be selected initially for signal sequence mutagenesis. In particular it had been demonstrated that preprochymosin behaved as a 'classical' eukaryotic secretory protein when

evolved from a single gene, in order to see what amino acid changes have been 'permitted' in these related signal peptides without destroying their function of initiating translocation across the ER. For example, comparing the signal peptides of six human α-interferons - A,B,C,D,F,G (Watson,1984), there is a common amino acid at 9 positions which are scattered throughout the signal sequence (22 residues in total, excluding the initiation Met). At the other 13 positions two or three different types of residue are found; the alternative amino acids are as diverse as Trp/Pro/Ser (position -20), or more similar such as Leu/Val (positions -2 and -13).

Theoretical considerations have led to the proposal of certain 'rules' governing the structure required for a functional signal sequence (ie Von Heijne,1985; discussed in I.B.3). Several groups have experimentally characterised the effect of mutations in the signal sequence region of exported prokaryotic proteins (reviewed in Benson *et al*,1985, Oliver,1985a; discussed in I.B.2), and these studies have drawn some consensus of the conformational requirements for a prokaryotic leader peptide. To what extent will studies on the effect of mutagenising the signal sequence of a eukaryotic secretory protein agree with the theoretical calculations, or with work on prokaryotic secretory proteins? Can the experimental observations with eukaryotic signal sequence mutants support or extend current theories on the successive stages in the translocation of nascent proteins across the ER membrane (discussed in I.A.2)?

In the prospective study it was initially intended to generate a range of mutations in the signal peptide region of cDNA encoding a eukaryotic secretory protein, and determine how effectively these mutant signal sequences functioned in the translocation of the nascent precursor. This would build up a picture of the effect of different amino acid substitutions on the maintenance of signal sequence function. This could lead later to a more detailed analysis of the flexibility at certain positions in the signal peptide.

The work described in Chapters III-VI examined the segregation of two wild-type secretory proteins with cleaved N-terminal signal peptides, preprochymosin and prelysozyme. Many of these experimental results were discussed in the context of the proposed signal sequence mutagenesis studies. From these studies, summarised in the previous section (VII.A.), it was decided that preprochymosin would be selected initially for signal sequence mutagenesis. In particular it had been demonstrated that preprochymosin behaved as a 'classical' eukaryotic secretory protein when

expressed in *Xenopus* oocytes and *in vitro* systems; the precursor is segregated within membranes, processed and (in oocytes) secreted. It has also been shown the preprochymosin signal sequence is essential for translocation of nascent prochymosin. Furthermore the signal peptide of preprochymosin apparently functions as an autonomous unit in initiating translocation across the ER membrane *in vivo* and *in vitro*. From the experimental work to date it appears the main disadvantage to using preprochymosin for signal sequence mutagenesis studies, is the difficulty in resolving the preprochymosin precursor from signal-processed prochymosin by SDS-PAGE. As discussed in V.E & IV.D.2 by experimenting further with different gel electrophoresis systems it may be possible to surmount this problem. In the following sections I will therefore be discussing the proposed signal sequence mutagenesis study with respect to preprochymosin, but similar considerations would apply if other eukaryotic secretory proteins were used.

In the next section I will discuss briefly some of the site-directed mutagenesis techniques which could be employed to introduce mutations into the signal peptide coding region of preprochymosin cDNA. The following section (VII.B.3) will briefly outline how it was intended to analyse the 'translocation phenotype' of the mutant preprochymosins; and consider the information this might yield on the process of the translocation of nascent eukaryotic proteins across the ER membrane.

VII.B.2 **Site-Directed Mutagenesis of the Preprochymosin Signal Sequence**

A full discussion of the subject of the site-directed mutagenesis of DNA can be found in the reviews of Shortle *et al*(1981); Smith & Gillam(1981) and Smith(1982). In terms of the prospective study under consideration (outlined above) several criteria and questions need to be considered in deciding the mutagenesis strategy to be adopted. Clearly a means of producing a range of mutations relatively rapidly and simply is preferred. How are mutations to be directed and limited to the preprochymosin signal sequence? How can one identify and isolate the mutated preprochymosin cDNAs? Following mutagenesis the mutant preprochymosin genes will need to be inserted into the SP6 vectors in order to study their biological activity (see VII.B.3). It is not feasible here to present a full or detailed consideration of the relative merits and disadvantages of each mutagenesis procedure, in terms of producing the required preprochymosin signal sequence mutants.

It was originally proposed to generate a range of mutations in the signal sequence region of the secretory protein under study using sodium

(a)		Met	Arg	Cys	Leu	Val	Val	Leu	Leu	Ala	Val	Phe	Ala	Leu	Ser	Gln	Gly
coding		ATG	AGG	TGT	CTC	GTG	GTG	CTA	CTT	GCT	GTC	TTC	GCT	CTC	TCC	CAG	GGC
non-coding		TAC	TCC	ACA	GAG	CAC	CAC	GAT	GAA	CGA	CAG	AAG	CGA	GAG	AGG	GTC	CCG
(b)		...	...	...	Phe	...	...	Leu	Phe	Val	Val	Phe	Val	Phe	Phe	End	Gly
coding					TTC			TTA	TTT	GTT	GTT	TTT	GTT	TTC	TTC	TAG	GGT
														Phe	Phe		
coding					Phe									TTT	TTT		
														Leu	Ser		
coding					Leu									CTT	TCT		
(c)		Ile	Lys	Tyr	...	Met	Met	...	...	Thr	Ile	...	Thr	...	...	Gln	Ser
coding		ATA	AAG	TAT		ATG	ATG			ACT	ATC		ACT			CAA	AGC
non-coding		TAT	TTC	ATA		TAC	TAC			TGA	TAG		TGA			GTT	TCG
			Lys			Ile	Ile										Asp
coding			AAA			ATA	ATA										GAC
non-coding			TTT			TAT	TAT										CTG
			Arg			Val	Val										Asn
coding			AGA			GTA	GTA										AAC
non-coding			TCT			CAT	CAT										TTG

Figure VII.1 Possible mutations in the signal sequence of preprochymosin by the action of sodium bisulphite
(a) The nucleotide and amino acid sequence of the wild-type preprochymosin signal sequence encoded by the PPChy cDNA (Harris *et al*,1982; see Figs.III.4 & V.2 and V.B.2). Below are the base changes induced by sodium bisulphite mutagenesis of the coding strand (b) and non-coding strand (c) of the wild-type PPChy DNA; together with the amino acid substitutions these base changes would cause.

bisulphite mutagenesis (Shortle & Nathans,1978; Shortle & Botstein,1983). Sodium bisulphite can be used to specifically deaminate cytosine bases in single stranded DNA, producing uracil which subsequently pairs with adenosine in repair and replication. Hence it is possible to generate two types of base substitutions in the coding sequence of a particular DNA fragment: if the coding strand is exposed to the bisulphite C to T transitions occur, and if the noncoding strand is mutagenised this results in G→A substitutions in the complementary coding strand. The mutagenesis conditions can be adjusted so that only one or two cytosines within the single stranded DNA are mutated, generating a mixed population of fragments with different base substitutions. Following treatment with sodium bisulphite the mutated DNA is then used to transform *E.coli*. This method has been used for site-directed mutagenesis by several groups (eg Giza *et al*,1981; Weiher & Schaller,1982; Folk & Hofstetter,1983). Figure VII.1 shows the mutations which could be produced in the preprochymosin signal sequence by the action of sodium bisulphite on the coding and noncoding strands of the PPChy preprochymosin cDNA (described in Fig.III.4). Work of Chapter VI had revealed an error in the published sequence of this preprochymosin cDNA (Harris *et al*,1982); the nucleotide sequence encoding the preprochymosin signal peptide shown in Fig.VII.1 is from my own DNA sequence analysis (see VI.B.2, Fig.VI.2). It was envisaged that a modification of the procedure used by Folk & Hofstetter(1983), involving M13 vectors, would be employed to produce a mutagenesis template in which the signal sequence of preprochymosin was exposed as single stranded DNA. It was hoped this could basically be achieved by annealing complementary strands of M13-PPChy and M13-PChy DNA.

The use of the nucleotide analogue [N⁴-hydroxy]dCTP (OHdCTP) provides an alternative means of generating a range of mutations within a particular stretch of DNA (ie see Muller *et al*,1978; Dierks *et al*,1981 and 1983; Wieringa *et al*,1983). N⁴-hydroxycytosine can base pair with both guanine and adenine, with roughly equal efficiency. Hence OHdCTP can substitute for dCTP and/or dTTP in the synthesis of the complementary strand to single stranded DNA. Following replication, where OHdCTP had originally been incorporated opposite guanine about 50% of the progeny will contain a G:C to A:T mutation; similarly, pairing of OHdCTP with adenine will result in some mutant progeny with A:T to G:C substitutions. Depending on whether the coding or noncoding strand is mutagenised in this way, and according to which nucleotides are replaced by OHdCTP in the strand synthesis, it is possible to produce A→G, T→C, G→A and C→T mutations in the coding sequence. The latter two of these are the same as

(a)		Met	Arg	Cys	Leu	Val	Val	Leu	Leu	Ala	Val	Phe	Ala	Leu	Ser	Gln	Gly
	coding	ATG	AGG	TGT	CTC	GTG	GTG	CTA	CTT	GCT	GTC	TTC	GCT	CTC	TCC	CAG	GGC
	non-coding	TAC	TCC	ACA	GAG	CAC	CAC	GAT	GAA	CGA	CAG	AAG	CGA	GAG	AGG	GTC	CCG

(b)		...	...	...	Phe	...	...	Leu	Phe	Val	Val	Phe	Val	Phe	Phe	End	Gly
	coding				TTC			TTA	TTT	GTT	GTT	TTT	GTT	TTC	TTC	TAG	GGT
	coding				Phe									Phe	Phe		
					TTT									TTT	TTT		
	coding				Leu									Leu	Ser		
					CTT									CTT	TCT		

(c)		Ile	Lys	Tyr	...	Met	Met	...	...	Thr	Ile	...	Thr	...	...	Gln	Ser
	coding	ATA	AAG	TAT		ATG	ATG			ACT	ATC		ACT			CAA	AGC
	non-coding	TAT	TTC	ATA		TAC	TAC			TGA	TAG		TGA			GTT	TCG
			Lys			Ile	Ile										Asp
	coding		AAA			ATA	ATA										GAC
	non-coding		TTT			TAT	TAT										CTG
			Arg			Val	Val										Asn
	coding		AGA			GTA	GTA										AAC
	non-coding		TCT			CAT	CAT										TTG

Figure VII.1 Possible mutations in the signal sequence of preprochymosin by the action of sodium bisulphite
(a) The nucleotide and amino acid sequence of the wild-type preprochymosin signal sequence encoded by the PPChy cDNA (Harris *et al*,1982; see Figs.III.4 & V.2 and V.B.2). Below are the base changes induced by sodium bisulphite mutagenesis of the coding strand (b) and non-coding strand (c) of the wild-type PPChy DNA; together with the amino acid substitutions these base changes would cause.

bisulphite mutagenesis (Shortle & Nathans,1978; Shortle & Botstein,1983). Sodium bisulphite can be used to specifically deaminate cytosine bases in single stranded DNA, producing uracil which subsequently pairs with adenosine in repair and replication. Hence it is possible to generate two types of base substitutions in the coding sequence of a particular DNA fragment: if the coding strand is exposed to the bisulphite C to T transitions occur, and if the noncoding strand is mutagenised this results in G→A substitutions in the complementary coding strand. The mutagenesis conditions can be adjusted so that only one or two cytosines within the single stranded DNA are mutated, generating a mixed population of fragments with different base substitutions. Following treatment with sodium bisulphite the mutated DNA is then used to transform *E.coli*. This method has been used for site-directed mutagenesis by several groups (eg Giza *et al*,1981; Weiher & Schaller,1982; Folk & Hofstetter,1983). Figure VII.1 shows the mutations which could be produced in the preprochymosin signal sequence by the action of sodium bisulphite on the coding and noncoding strands of the PPChy preprochymosin cDNA (described in Fig.III.4). Work of Chapter VI had revealed an error in the published sequence of this preprochymosin cDNA (Harris *et al*,1982); the nucleotide sequence encoding the preprochymosin signal peptide shown in Fig.VII.1 is from my own DNA sequence analysis (see VI.B.2, Fig.VI.2). It was envisaged that a modification of the procedure used by Folk & Hofstetter(1983), involving M13 vectors, would be employed to produce a mutagenesis template in which the signal sequence of preprochymosin was exposed as single stranded DNA. It was hoped this could basically be achieved by annealing complementary strands of M13-PPChy and M13-PChy DNA.

The use of the nucleotide analogue [N⁴-hydroxy]dCTP (OHdCTP) provides an alternative means of generating a range of mutations within a particular stretch of DNA (ie see Muller *et al*,1978; Dierks *et al*,1981 and 1983; Wieringa *et al*,1983). N⁴-hydroxycytosine can base pair with both guanine and adenine, with roughly equal efficiency. Hence OHdCTP can substitute for dCTP and/or dTTP in the synthesis of the complementary strand to single stranded DNA. Following replication, where OHdCTP had originally been incorporated opposite guanine about 50% of the progeny will contain a G:C to A:T mutation; similarly, pairing of OHdCTP with adenine will result in some mutant progeny with A:T to G:C substitutions. Depending on whether the coding or noncoding strand is mutagenised in this way, and according to which nucleotides are replaced by OHdCTP in the strand synthesis, it is possible to produce A→G, T→C, G→A and C→T mutations in the coding sequence. The latter two of these are the same as

(a)		Met	Arg	Cys	Leu	Val	Val	Leu	Leu	Ala	Val	Phe	Ala	Leu	Ser	Gln	Gly
	coding	ATG	AGG	TGT	CTC	GTG	GTG	CTA	CTT	GCT	GTC	TTC	GCT	CTC	TCC	CAG	GGC
(b)		Val	Gly	...	...	...	...	Leu	...	...	...	...	...	...	...	Arg	...
	coding	GTG	GGG					CTG								CGG	
(c)		Thr	...	Arg	Pro	Ala	Ala	Pro	Pro	Ala	Ala	Leu	Ala	Pro	Pro	...	...
	coding	ACG		CGT	CCC	GCG	GCG	CCA	CCT	GCC	GCC	CTC	GCC	CCC	CCC		
	coding			Arg					Pro			Ser					
				CGC					CCC			TCC					
	coding			Cys					Leu			Pro					
				TGC					CTC			CCC					

Figure VII.2 Possible mutations in the signal peptide of preprochymosin using [N^4-hydroxy]dCTP

(a) The wild-type signal sequence of preprochymosin (see Fig.VII.1).

(b & c) The amino acid changes generated by (b) A→G and (c) C→T changes in the nucleotide sequence of the signal peptide region of preprochymosin (PPChy). These base substitutions could be generated if complementary strands to the coding and non-coding strands, respectively, of PPChy are synthesized using [N^4-hydroxy]dCTP, dATP, dGTP and dCTP. If [N^4-hydroxy]dCTP, dATP, dGTP and dTTP were used in the same way the possible amino acid changes in the preprochymosin signal peptide are the same as those generated by sodium bisulphite mutagenesis (see Fig.VII.1).

the mutations induced by sodium bisulphite mutagenesis (see Fig.VII.1). Figure VII.2 shows the new amino acid changes which could be generated in the preprochymosin signal peptide through the use of OHdCTP. A drawback of this mutagenesis method is that the nucleotide analogue OHdCTP is not commercially available and must be synthesized by the reaction of hydroxylamine with dCTP. It is necessary to produce a suitable primer-template complex in which only sequences complementary to all, or part of, the preprochymosin signal sequence are synthesized in the presence of OHdCTP. In devising a strategy to generate a suitable template for the mutagenesis of such a small stretch of DNA (approx. 50bp) one is hindered by the paucity of restriction sites. Consequently it may well be necessary to also employ modifying enzymes such as Bal31, T_4 polymerase or Klenow DNA polymerase to produce a suitable mutagenesis template.

Recently Myers *et al*(1985) reported a new procedure for generating and isolating random single base substitutions in cloned DNA fragments; using various chemicals to mutagenise the DNA. This may prove a convenient method of producing a range of mutations in the preprochymosin signal sequence.

The elongation of an appropriate primer-template complex, similar to those suitable for 'OHdCTP mutagenesis', is also involved in site-directed mutagenesis by 'forced nucleotide misincorporation'. In this technique single noncomplementary nucleotides are inserted at specific sites by error-directed DNA polymerisation, and one can in principle generate any base substitution (ie Traboni *et al*,1893; Zakour *et al*,1984). Although this mutagenesis method does not produce a range of mutants from a single reaction, it is possible to produce several mutations within a short stretch of nucleotides by limited manipulation of a basic primer-template complex.

Oligonucleotide-directed mutagenesis would provide the best means of producing specific base changes with high efficiency, at any position in the signal sequence (for review see Zoller & Smith,1983). Specific oligonucleotides, containing a single mutation of the PPChy sequence, would be annealed to single stranded wild-type PPChy, and then extended and replicated to produce mutant preprochymosin genes. However each mutant would entail the synthesis of a different oligonucleotide; so to produce a number of mutants would tend to be costly and time consuming. Although improvements in technology in recent years have resulted in cheaper and better machines and reagents for oligonucleotide synthesis;

so that it is no longer always necessary to obtain oligonucleotides from commercial laboratories.

It is noted that some of the preprochymosin mutants generated by 'random' mutagenesis methods (ie. bisulphite, OHdCTP, and the saturation mutagenesis protocol of Myers *et al*,1985) will not be suitable for translocation studies. These include nucleotide substitutions which do not change the preprochymosin signal peptide amino acid sequence, mutants in which the initiation codon is destroyed, and base changes which introduce a termination codon - see Figs.VII.1 & 2. Similarly, mutations in the 5' noncoding sequence may alter the translational efficiency of the cDNA, which could complicate the analysis of these mutants. On the other hand, if individual site-specific mutations are to be produced one needs to consider what rationale should be adopted to decide which particular amino acid changes are to be studied.

Following mutagenesis of the DNA, the identification of preprochymosin signal sequence mutants amongst the transformants obtained would involve DNA sequence analysis; and in many cases there will be no other means of discriminating transformants containing mutated, as opposed to the original wild-type, preprochymosin DNA. It is clearly highly desirable to use a method which produces a high frequency of mutants, to reduce the time taken to screen for mutants of interest. Whichever method is used it would be convenient if the preprochymosin DNA were contained in one of the M13 vectors for mutagenesis. Therefore, as a preliminary step towards the site-directed mutagenesis of the preprochymosin signal sequence, both the PPChy and PChy *HindIII* inserts (Figs.III.4 & 5) were cloned into M13mp10 and mp11, and transfectants containing these inserts in either orientation were isolated (as described in the Materials & Methods, II.C). These M13 clones could be used, in various ways, in the construction of a template for the mutagenesis of the preprochymosin signal peptide, as discussed in this section.

VII.B.3 Characterisation of Signal Sequence Mutants

The primary step in the biological analysis of preprochymosin signal sequence mutants will be to clone the mutated PPChy cDNAs into a SP6 vector, and produce synthetic transcripts *in vitro*. In characterising the translocation phenotype of preprochymosins with mutated signal sequences, each experiment will need to include wild-type SP6PPChy and PChy SP6 RNA (encoding Met-prochymosin) as positive and negative controls for translocation. It is not anticipated that there will necessarily be a clear division between signal sequence mutants which can

and cannot translocate the ER membrane; for some mutants translocation may be impaired but not abolished. The primary concern is how the structure of the signal sequence relates to its role in enabling the nascent polypeptide to translocate the ER membrane.

It is known that the PChy polypeptide, which contains no signal peptide, is localised exclusively in the cytoplasm when expressed in *Xenopus* oocytes, and shows no association with the membrane and vesicle fraction; whilst full-length preprochymosin is segregated within the membranes and secreted (see VII.A). Oocyte microinjection experiments using transcripts of mutant preprochymosins will identify which mutants are detected in the membrane fraction, and those which are also secreted. In some mutants processing of the signal peptide will provide a useful indication of translocation of the ER membrane. As noted previously, it is anticipated that certain mutations, towards the C-terminus of the signal peptide, would inhibit cleavage of the signal sequence but not abolish initiation of translocation; and that inhibition of signal processing could prevent completion of translocation or other steps in the export pathway. Mutant preprochymosins which have functionally defective signal peptides and do not translocate the ER membrane, may not necessarily be detected in the cytosol, as discussed in VII.A. Parallel experiments in which the mutant preprochymosin SP6 RNAs are translated *in vitro* in the presence of microsomes, will provide further evidence as to whether the mutated signal sequence is able to direct the translocation of the nascent precursor into ER-derived vesicles; and whether the translocated proteins are fully protected from digestion by exogenous proteases (discussed in VII.A.1).

Following the identification of nontranslocated mutants it is hoped to carry out a closer examination of the basis for the failure of particular preprochymosins to segregate. *In vitro* translation-translocation experiments would be used, in which the process of translocation can be biochemically dissected. Adopting experimental approaches similar to others, as cited below, it is envisaged that the translocation of wild-type and mutant preprochymosins will be compared in studies which investigate the following questions. Does signal recognition particle recognise and bind to polysomes synthesizing the nascent preprochymosin? (Walter *et al*,1981). In a wheat germ system primed with SP6 RNA is synthesis of the preprochymosins blocked by the addition of SRP? (Walter & Blobel,1981b; Rottier *et al*,1984). Is the translation arrest induced by SRP released by SRP Receptor present in whole or salt-extracted (SRP-depleted) microsomes? (Walter &

Blobel,1981b; Meyer *et al*,1982b; Gilmore *et al*,1982b). From the results of such studies it may be found that nontranslocated mutants can be divided into groups with different characteristics. For example, mutants which apparently interact with SRP but which remain in translation arrest even in the presence of microsomes. The translation of certain mutants may be unaffected by SRP. Others may show SRP-mediated elongation arrest and then continue translation when microsomes are added, but still not translocate the ER membrane. It is noted that recent studies have shown that the phenomenum of SRP-induced translation arrest is limited to the wheat germ *in vitro* translation system, and does not occur in mammalian cell-free systems (Hortsch & Meyer,1984; Meyer,1985). Hence it is probably an artefact of the reconstituted wheat germ system, and not an integral part of the translocation process *in vivo*. However, it may still be useful to use SRP-mediated elongation arrest to characterise preprochymosin signal sequence mutants, if it is found that the translation of wild-type preprochymosin is arrested by SRP in a wheat germ cell-free system. Since this could enable the interaction of SRP with the signal peptide to be separated from subsequent stages of translocation. It may also provides some information about the structural features of the signal sequence which allow it to interact with SRP. The existence or absence of subgroups amonst the nontranslocated mutants will indicate whether the sole function of the signal sequence is to bind SRP; with the effect that if this interaction occurs the nascent polypeptide will translocate the ER membrane. Alternatively it may appear that the signal sequence contains information which specifies its interaction with several components of the translocation machinery.

Analysis of the effect of particular amino acid substitutions in the preprochymosin signal peptide on the 'translocation phenotype' of the nascent precursor, will gradually build up a picture of the structure-function relationship of the preprochymosin signal sequence. Not only will a study of signal sequence mutants whose translocation is impaired contribute to this picture, but information is also to be gained from those signal sequence mutants in which the amino acid substitution(s) does not affect translocation.

-REFERENCES-

The references are cited according to the conventions in 'Instruction to authors' issued by the Biochemical Journal, except that the title of each reference is also given

A:

Abrahmsén, L., Moks, T., Nilsson, B., Hellman, U. & Uhlén, M. (1985) EMBO J. 4, 3901-3906. Analysis of signals for secretion in the staphylococcal protein A gene.

Adams, G.A. & Rose, J.K. (1985a) Cell 41, 1007-1015. Structural requirements of a membrane-spanning domain for protein anchoring and cell surface transport.

Adams, G.A. & Rose, J.K. (1985b) Mol.Cell.Biol. 5, 1442-1448. Incorporation of a charged amino acid into the membrane-spanning domain blocks cell-surface transport but not membrane-anchoring of a viral protein.

Anderson, D.J., Walter, P. & Blobel, G. (1982) J.Cell Biol 93, 501-506. Signal recognition protein is required for the integration of acetylcholine receptor δ subunit, a transmembrane glycoprotein, into the endoplasmic reticulum membrane.

Anderson, D.J., Mostov, K.E. & Blobel, G. (1983) Proc.Natl.Acad.Sci.USA 80, 7249-7253. Mechanisms of integration of *de-novo*-synthesized polypeptides into membranes: Signal-recognition particle is required for integration into microsomal membranes of calcium ATPase and of lens MP26 but not of cytochrome b_5.

Asato, N. & Rand, A.G.Jr. (1971) Anal.Biochem. 44, 32-41. Resolution of prorennin and rennin by polyacylamide gel electrophoresis.

Asato, N. & Rand, A.G.Jr. (1972) Biochem.J. 129, 841-846. Fractionation and isolation of the multiple forms of prorennin (prochymosin).

Asselbergs, F.A.M. (1979) Mol.Biol.Rep. 5, 199-208. Protein synthesis on heterologous messenger RNA in *Xenopus* oocytes.

Austen, B.M., Hermon-Taylor, J., Kaderbhai, M.A. & Ridd, D.H. (1984) Biochem.J. 224, 317-325. Design and synthesis of a consensus signal sequence that inhibits protein translocation into rough microsomal vesicles.

B:

Banerjee, A.K. (1980) Microbiol.Rev. 44, 175-205. 5'-terminal cap structure in eukaryotic messenger ribonucleic acids.

Bankaitis, V.A., Rasmussen, B.A. & Bassford Jr, P.J. (1984) Cell 37, 243-252. Intragenic suppressor mutations that restore export of maltose binding protein with a truncated signal peptide.

Bar-Nun, S., Kreibich, G., Adesnik, M., Alterman, L., Negishi, M. & Sabatini, D.D. (1980) Proc.Natl.Acad.Sci.USA 77, 965-969. Synthesis and insertion of cytochrome P-450 into endoplasmic reticulum membranes.

Bassford, P. & Beckwith, J. (1979) Nature 277, 538-541. *Escherichia coli* mutants accumulating the precursor of a secreted protein in the cytoplasm.

Bassford Jr, P.J., Emr, S.D., Silhavy, T.J., Beckwith, J., Bedouelle, H., Clement, J.M., Hedgepeth, J. & Hofnung, M. (1981) Meth.Cell Biol. 23, 27-38. The genetics of protein secretion in *Escherichia coli*.

Bedouelle, H., Bassford, P.J., Fowler, A.V., Zabin, I. & Beckwith, J. (1980) Nature 285, 78-81. Mutations which alter the function of the signal sequence of the maltose binding protein of *Escherichia coli*.

Bendzko,P., Prehn, S., Pfeil, W. & Rapoport, T.A. (1982)
Eur.J.Biochem. 123, 121-126. Different modes of membrane interactions of the signal sequence of carp preproinsulin and of the insertion sequence of rabbit cytochrome b_5.

Benson, S.A. & Silhavy, T.J. (1983) Cell 32, 1325-1335.
Information within the mature Lam B protein necessary for localization to the outer membrane of E.coli K12.

Benson, S.A. Bremer, E. & Silhavy, T.J. (1984) Proc.Natl.Acad.Sci.USA 81, 3830-3834. Intragenic regions required for Lam B export.

Benson, S.A., Hall, M.N. & Schwartz, T.J. (1985) Ann.Rev.Biochem. 54, 101-134. Genetic analysis of protein export in *Escherichia coli K12*.

Bergmann, J. & Singer, S.J. (1983) J.Cell Biol. 97, 1777-1787. Immunoelectron microscope studies of the intracellular transport of the membrane glycoprotein (G) of Vesicular Stomatitis Virus in infected Chinese Hamster ovary cells.

Birnboim, H.C. & Doly, J. (1979) Nucl.Acid.Res. 7, 1513-1523. A rapid alkaline extraction procedure for screening recombinant plasmid DNA.

Blobel, G. (1980) Proc.Natl.Acad.Sci.USA 77, 1496-1500.
Intracellular protein topogenesis.

Blobel, G. & Dobberstein, B. (1975a) J.Cell Biol. 67, 835-851.
Transfer of proteins across membranes. I Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma.

Blobel, G. & Dobberstein, B. (1975b) J.Cell Biol. 67, 852-862.
Transfer of proteins across membranes. II Reconstitution of functional rough microsomes from heterologous components.

Blobel, G. & Sabatini, D.D. (1971)
In 'Biomembranes', (Manson, L.A. ed) vol.2, pp193-195, Plenum Press, New York. 'Ribosome-membrane interaction in eukaryotic cells'.

Boeke, J.D. & Model, P. (1982) Proc.Natl.Acad.Sci.USA 79, 5200-5204.
A prokaryotic membrane anchor sequence: carboxyl terminus of bacteriophage f1 gene III protein retains it in the membrane.

Bonner, W.M. (1978) In 'The cell nucleus 6' (Busch, H. ed), Part C pp97-148. Academic Press, London.
Protein migration and accumulation in nuclei.

Bonner, W.M. & Laskey, R.A. (1974) Eur.J.Biochem. 46, 83-88.
A film detection method for tritium-labelled proteins and nucleic acids in polyacrylamide gels.

Borgese, N. & Gaetani, S. (1980) FEBS Letts. 112, 216-222.
Site of synthesis of rat liver NADH-cytochrome b_5 reductase, an integral membrane protein.

Bos, T.J., Davis, A.R. & Nayak, D.P. (1984) Proc.Natl.Acad.Sci.USA 81, 2327-2331. NH_2-terminal hydrophobic region of influenza virus neuraminidase provides the signal function in translocation.

Braell, W.A. & Lodish, H.F. (1982a) Cell 28, 23-31. The erythrocyte anion transport protein is co-translationally inserted into microsomes.

Braell, W.A. & Lodish, H.F. (1982b) J.Biol.Chem. 257, 4578-4582. Ovalbumin utilizes an NH_2-terminal signal sequence.

Brown, W.J. & Farquar, M.G. (1984) Cell 36, 295-307. The mannose-6-phosphate receptor for lysosomal enzymes is concentrated in the *cis* Golgi cisternae.

Burke, B., Matlin, K., Bause, E., Leggler, G., Peyrieras, N. & Ploegh, H. (1984) EMBO J. 3, 551-556. Inhibition of N-linked oligosaccharide trimming does not interfere with surface expression of certain integral membrane proteins.

Bursztain, S. & Fischbach G.D. (1984) J.Cell Biol. 98, 498-506. Evidence that coated vesicles transport acetylcholine receptors to the surface membrane of chick myotubes.

C:
Cancedda, R. & Schlesinger, M.J. (1974) J.Bact. 117, 290-301. Localization of polyribosomes containing alkaline phosphatase nascent polypeptides on membranes of *Escherichia coli*.

Chou, P.Y. & Fasman, G.D. (1978) Ann.Rev.Biochem. 47, 251-276. Empirical predictions of protein conformation.

Chua, N-H. & Schmidt, G.W. (1979) J.Cell Biol. 81, 461-483. Transport of proteins into mitochondria and chloroplasts.

Clemens, M.J. (1984) In 'Transcription and translation; a practical approach' (Hames, B.D. & Higgins, S.J., eds.), pp231-270, IRL Press Ltd., Oxford. Translation of eukaryotic messenger RNA in cell-free extracts.

Coleman, J., Inukai, M. & Inouye, M. (1985) Cell 43, 351-360. Dual function of the signal peptide in protein transfer across the membrane.

Colman, A. (1982) TIBS. 7, 435-437. Cells that secrete foreign proteins.

Colman, A. (1984a) In 'Transcription and translation: a practical approach.' *ibid* pp271-303. Translation of eukaryotic messenger RNA in *Xenopus* oocytes.

Colman, A. (1984b) *ibid* pp46-69. Expression of exogenous DNA in *Xenopus* oocytes.

Colman, A. & Morser, J. (1979) Cell 17, 517-526. Export of proteins from oocytes of *Xenopus laevis*.

Colman, A., Lane, C., Craig, R., Boulton, A., Mohun, T. & Morser, J. (1981) Eur.J.Biochem 113, 339-348. The influence of topology and glycosylation on the fate of heterologous secretory proteins made in *Xenopus* oocytes.

Colman, A., Cutler, D., Krieg, P. & Valle, G. (1983) In 'Molecular biology of egg maturation' pp249-267, Ciba Foundation Symposium 98. Pitman Books, London. The oocyte as a secretory cell.

Cutler, D. (1982) PhD. thesis, University of Warwick.
Secretion of chick proteins from *Xenopus* oocytes.

Cutler, D., Lane, C. & Colman, A. (1981) J.Mol.Biol. 153, 917-931.
Non-parallel kinetics and the role of tissue-specific factors in the secretion of chicken ovalbumin and lysozyme from *Xenopus* oocytes.

D:
Date, T., Zwizinski, C., Ludmerer, S. & Wickner, W. (1980a)
Proc.Natl.Acad.Sci.USA 77, 827-831. Mechanisms of membrane assembly.The effects of energy poisons on the conversion of soluble M13 procoat to membrane-bound coat protein.

Date, T., Goodman, J.M. & Wickner, W. (1980b)
Proc.Natl.Acad.Sci.USA 77, 4669-4673. Procoat, the precursor of M13 coat protein, requires an electrochemical potential for membrane insertion.

Daum, G., Gasser, S.M. & Schatz, G. (1982) J.Biol.Chem. 257. 13075-13080.
Import of proteins into mitochondria. Energy-dependent two-step processing of the intermembrane space enzyme cytochrome b_2 by isolated yeast mitochondria.

Davey, J., Dimmock, N.J. & Colman, A. (1985) Cell 40, 667-675.
Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in *Xenopus* oocytes.

Davis, B.D. & Tai, P-C. (1980) Nature 283. 433-438.
The mechanism of protein secretion across membranes.

Davis, N.G., Boeke, J.D. & Model, P. (1985) J.Mol.Biol. 181, 111-121.
Fine structure of a membrane anchor domain.

Davis, R.W. (1982) In 'Gel electrophoresis of nucleic acids: a practical approach' (Rickwood, D. & Hames, B.D. eds) pp117-172. IRL Press Ltd. Oxford, UK. 'DNA sequencing'.

De Robertis, E.M. (1983) Cell 32, 1021-1025.
Nucleocytoplasmic segregation of proteins and mRNAs.

Dhar, R., McClements, W.L., Enquist, L.W. & Vande Woude, G.F. (1980)
Proc.Natl.Acad.Sci.USA 77, 3937-3941. Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions.

Dierks, P., Wieringa, B., Marti, D., Reiser, J., Van Ooyen, A., Meyer, F., Weber, H. & Weissmann, C. (1981) In 'Developmental biology using purified genes'. (Brown, D.D. & Fox, C.F) ICN-UCLA Symp.Mol.Cell.Biol. vol 23, pp347-366. Expression of β-globin genes modified by restructuring and site-directed mutagenesis.

Dierks, P., Van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser,J. & Weissmann,C. (1983) Cell 32, 695-706. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin in mouse 3T6 cells.

Dingwall, C. (1985) TIBS, 10, 64-66.
The accumulation of proteins in the nucleus.

Dobberstein, B. (1978) Hoppe Seyler's Z.Physiol.Chem. 359, 1469-1470.
Reassembly of functional rough microsomal membranes.

Dobberstein, B., Blobel, G. & Chua, N-H. (1977) Proc.Natl.Acad.Sci.USA 74, 1082-1085. *In vitro* synthesis and processing of a putative precursor for the small subunit of ribulose-1,5-bisphosphate carboxylase of *Chlamydomonas reinhardtii*.

Docherty, P.A. & Aronson, N.N.Jr. (1985) J.Biol.Chem. 260, 10847-10855. Effect of the threonine analog β-hydroxynorvaline on the glycosylation and secretion of α_1-glycoprotein by rat hepatocytes.

Donnelly, W.J., O'Callaghan, D.M., Carroll, D.P., McConnell, D.J. & Gannon, F. (1984) Biochem.Soc.Transactions 12, 440-441. Multiple forms of calf prochymosin and chymosin.

Doyle, C., Roth, M.G., Sambrook, J. & Gething M-J (1985) J.Cell Biol. 100, 704-714. Mutations in the cytoplasmic domain of the influenza virus haemagglutinin affect different stages of intracellular transport.

Drummond, D., Armstrong, J. & Colman, A. (1985) Nucl.Acids Res. 13, 7375-7394. The effect of capping and polyadenylation on the stability, movement and translation of synthetic messenger RNAs in *Xenopus* oocytes.

Dunphy, W.G. & Rothman, J.E. (1985) Cell 42, 13-21.
Compartmental organisation of the Golgi stack.

E:
Ellis, J. (1985) Nature 313, 353-354.
Eukaryotic proteins retargetted among cell compartments.

Emr, S.D. & Silhavy, T.J. (1982) J.Cell Biol. 95, 689-696.
Molecular components of the signal sequence that function in the initiation of protein export.

Emr, S.D. & Silhavy, T.J. (1983) Proc.Natl.Acad.Sci.USA 80, 4599-4603. Importance of secondary structure in the signal sequence for protein secretion.

Emr, S.D., Hall, M.N. & Silhavy, T.J. (1980) J.Cell Biol. 86, 701-711. The signal hypothesis and bacteria.

Emtage, J.S., Angal, S., Doel, M.T., Harris, T.J.R., Jenkins, B., Lilley, G. & Lowe, P. (1983) Proc.Natl.Acad.Sci.USA 80, 3671-3675. Synthesis of calf prochymosin (prorennin) in *Escherichia coli*.

Engelman, D.M. & Steitz, T.A. (1981) Cell 23, 411-422.
The spontaneous insertion of proteins into and across membranes: the helical hairpin hypothesis.

Enquist, H.G., Hirst, T.R., Harayama, S., Hardy, S.J.S. & Randall, L.L. (1981) Eur.J.Biochem 116, 227-233. Energy is required for maturation of exported proteins in *Escherichia coli*.

Erickson, A.H., Conner, G.E. & Blobel, G. (1981)
J.Biol.Chem. 256, 11224-11231. Biosynthesis of a lysosomal enzyme.

F:
Farquar, M.G. & Palade, G.E. (1981) J.Cell Biol. 91, 77s-103s. The Golgi apparatus (complex) - (1954-1981) - from artefact to centre stage.

Folk, W.R. & Hofstetter, H. (1983) Cell 33, 585-593. A detailed mutational analysis of the eucaryotic $tRNA_1^{met}$ gene promoter.

Foltmann, B. (1970) Meths.Enzymol. 19, 421-436.
Prochymosin and chymosin (prorennin and rennin).

Foltmann, B. (1981) In 'Essays in Biochemistry' vol.17 (Campbell,P.N. & Marshall,R.D. eds) pp52-84. Academic Press, London, UK.
Gastric proteases- structure, function, evolution and mechanism of action.

Foltmann, B., Pederson, V.B., Jacobsen, H., Kauffman, D. & Wybrandt, G. (1977) Proc.Natl.Acad.Sci.USA 74, 2321-2324. The complete amino acid sequence of prochymosin.

Foltmann, B., Pederson, V.B., Kauffman, D. & Wybrandt, G. (1979) J.Biol.Chem. 254, 8447-8456. The primary structure of calf prochymosin.

Foreman, R.C., Judah, J.D. & Colman. A. (1984) FEBS Letts. 168, 84-88.
Xenopus oocytes can synthesize but do not secrete the Z variant of human α_1-antitrypsin.

Fraser, T.H. & Bruce, B.J. (1978) Proc.Natl.Acad.Sci.USA 75, 5936-5940.
Chicken ovalbumin is synthesized and secreted by *Escherichia coli*.

Friedlander, M. & Blobel, G. (1985) Nature 318, 338-343.
Bovine opsin has more than one signal sequence.

Fries, E., Gustafsson, L. & Peterson, P.A. (1984) EMBO J. 3, 147-152.
Four secretory proteins synthesized by hepatocytes are transported from endoplasmic reticulum to Golgi complex at different rates.

Fuller, S.D., Bravo, R. & Simons, K. (1985) EMBO J. 4, 297-307.
An enzymatic assay reveals that proteins destined for the apical or basolateral domains of an epithelial cell line share the same late Golgi compartments.

Furuichi, Y., LaFiandra, A. & Shatkin, A.J. (1977) Nature 266, 235-239.
5'-terminal structure and mRNA stability.

G:
Garoff, H. (1985) Ann.Rev.Cell Biol. 1, ?-??. Using recombinant DNA techniques to study protein targetting in the eukaryotic cell.

Garoff, H., Kondor-Koch, C., Petterson, R. & Burke, B. (1983) J.Cell Biol. 97, 652-658. Expression of Semliki Forest Virus proteins from cloned complementary DNA. II The membrane-spanning glycoprotein E2 is transported to the cell surface without its normal cytoplasmic domain.

Gasser, S.M. & Schatz, G. (1983) J.Biol.Chem. 258, 3427-3430.
Import of proteins into mitochondria: *In vitro* studies on the biogenesis of the outer membrane.

Gasser, S.M., Daum, G. & Schatz, G. (1982) J.Biol.Chem. 257, 13034-13041.
Import of proteins into mitochondria.

Gething, M-J. & Sambrook, J. (1982) Nature 300, 598-603.
Construction of influenza haemagglutinin genes that code for intracellular and secreted forms of the protein.

Geuze, H.J., Slott, J.W., Strous, G.J.A.M., Hasilik, A. & Von Figura, K. (1985) J.Cell Biol. 101, 2253-2262. Possible pathways for lysosomal enzyme delivery.

Gibson, R., Leavitt, R., Kornfeld, S. & Schlesinger, S. (1978)
Cell 13, 671-679. Synthesis and infectivity of vesicular stomatitis virus containing nonglycosylated G protein.

Gilmore, R. & Blobel, G. (1983) Cell 35, 677-685.
Transient involvement of signal recognition particle and its receptor in the microsomal membrane prior to protein translocation.

Gilmore, R. & Blobel, G. (1985) Cell 42, 497-505.
Translocation of secretory proteins across the microsomal membrane occurs through an environment accessible to aqueous perturbants.

Gilmore, R., Blobel, G. & Walter, P. (1982a) J.Cell Biol. 95, 463-469.
Protein translocation across the endoplasmic reticulum I. Detection in the microsomal membrane of a receptor for the signal recognition particle.

Gilmore, R., Blobel, G. & Walter, P. (1982b) J.Cell Biol. 95, 470-477.
Protein translocation across the endoplasmic reticulum II Isolation and characterization of the signal recognition particle.

Giza, P.E., Schmidt, D.M. & Murr, B.L. (1981) Gene 15, 331-342.
Region and strand specific mutagenesis of a recombinant plasmid.

Glabe, C.G., Hanover, J.A. & Lennarz, W.J. (1980) J.Biol.Chem 255, 9236-9242.
Glycosylation of ovalbumin nascent chains.

Goff, C.G., Moir, D.T., Kohno, T., Gravius, T.C., Smith, R.A., Yamasaki, E. & Taunton-Rigby, A. (1984) Gene 27, 35-46. Expression of calf prochymosin in *Saccharomyces cerevisiae*.

Gray, G.L., Baldridge, J.S., McKeown, K.S., Heynekker, H.L. & Chang, C.N. (1985) Gene 39, 247-254. Periplasmic production of correctly processed human growth hormone in *Escherichia coli*: natural and bacterial signal sequences are interchangeable.

Green, R. & Shields, D. (1984) J.Cell Biol. 99, 97-104.
Somatostatin discriminates between the intracellular pathways of secretory and membrane proteins.

Griffiths, G., Quinn, P. & Warren, G. (1983) J.Cell Biol. 96, 835-850.
Dissection of the Golgi Complex I. Monensin inhibits the transport of viral membrane proteins from *medial* to *trans* Golgi cisternae in Baby Hamster Kidney cells infected with Semliki Forest virus.

Gronenborn, B. & Messing, J. (1978) Nature 272, 375-377.
Methylation of a single-stranded DNA *in vitro* introduces new restriction endonuclease sites.

Guan, J-L. & Rose, J.K (1984) Cell 37, 779-787.
Conversion of a secretory protein into a transmembrane protein results in its transport to the Golgi complex but not to the cell surface.

Guan, J-L., Machamer, C.E. & Rose, J.K.(1985) Cell 42, 489-496.
Glycosylation allows cell-surface transport of an anchored secretory protein.

Gumbiner, B.B. & Kelly, R.B. (1982) Cell 28, 51-59.
Two distinct intracellular pathways transport secretory and membrane glycoproteins to the surface of pituitary tumor cells.

Gundelfinger, E.D., Krause, E., Melli, M. & Dobberstein, B. (1983) Nucl.Acid.Res. 11, 7363-7374. The organization of the 7SL RNA in the signal recognition particle.

Gurdon, J.B. & Melton, D.A. (1981) Ann.Rev.Genet. 15, 189-218. Gene transfer in amphbian eggs and oocytes.

Gurdon, J.B., Lane, C.D., Woodland, H.R. & Marbaix, G. (1971) Nature 233, 177-182. Use of frog eggs and oocytes for the study of messenger RNA and its translation in living cells.

H:
Hall, M.H., Schwartz, M. & Silhavy, T.S. (1982) J.Mol.Biol 156, 93-112. Sequence information within the *lamB* gene is required for proper routing of the bacteriophage λ receptor protein to the outer membrane of *Escherichia coi*.

Hamer, D.H. & Leder, P. (1979) Cell 18, 1299-1309. Splicing and the formation of stable RNA.

Hames, B.D. & Rickwood, D. (1981) 'Gel electrophoresis of proteins: a practical approach'. IRL Press Ltd. Oxford, UK.

Hanover, J.A. & Lennarz, W.J. (1981) Arch.Biochem.Biophys. 211. 1-19. Transmembrane assembly of membrane and secretory glycoproteins.

Harris, T.J.R., Lowe, P.A., Lyons, A., Thomas, P.G., Eaton, M.A.W., Millican, T.A., Patel, T.P., Bose, C.C., Carey, N.H. & Doel, M.T. (1982) Nucl.Acid.Res. 10, 2177-2187. Molecular cloning and nucleotide sequence of cDNA coding for calf preprochymosin.

Hase, T., Müller, U., Riezman, H. & Schatz, G. (1984) EMBO J. 3, 3157-3161. A 70kd protein of the yeast mitochondrial outer membrane is targetted and anchored via its extreme amino terminus.

Hasilik, A. & Neufeld, E.F. (1980) J.Biol.Chem. 255, 4946-4950. Biosynthesis of lysosmal enzymes in fibroblasts.

Hay, R., Böhni, P. & Gasser, S. (1984) Biochim.Biophys.Acta 779. 65-87. How mitochondria import proteins.

Heide, K. & Schwick, H.G. (1978). In 'Handbook of Experimental Immunology' (Weir, D.M. ed.) pp7.1-7.11. Blackwell Science Publications, Oxford, UK. 'Salt fractionation of immunoglobulins'.

Hennig, B., Koehler, H. & Neupert, W. (1983) Proc.Natl.Acad.Sci.USA 80, 4963-4967. Receptor sites involved in post-translational transport of apocytochrome c into mitochondria: specificty, affinity and number of sites.

Highfield, P.E. & Ellis, R.J. (1978) Nature 271, 420-424. Synthesis and transport of the small subunit of chloroplast ribulose bisphosphate carboxylase.

Hindley, J. (1983) 'DNA Sequencing'. In series Laboratory Techniques in Biochemistry and Molecular Biology (Work, T.S. & Burdon, R.H., eds) Elsevier Science Publishers, Oxford, UK.

Hortin, G. & Boime, I. (1980) Proc.Natl.Acad.Sci.USA 77, 1356-1360. Inhibition of preprotein processing in ascites tumour lysates by incorporation of a leucine analog.

Hortin, G. & Boime, I. (1981) Cell 24, 453-461.
Miscleavage at the presequence of rat preprolactin synthesized in pituitary cells incubated with a threonine analog.

Hortsch, M. & Meyer, D. (1984) Biol.Cell 52, 1-8.
Pushing the signal hypothesis: What are the limits?.

Hortsch, M., Avossa, D. & Meyer, D.I. (1985) J.Biol.Chem. 260, 9137-9145.
A structural and functional analysis of the docking protein.

Horwich, A.L., Fenton, W.A., Firgaira, F.A., Fox, J.E., Kolansky, D., Mellman, I.S. & Rosenberg, L.E. (1985a) J.Cell Biol. 100, 1515-1521.
Expression of amplified DNA sequences for ornithine transcarbamylase in HeLa cells: Arginine residues may be required for mitochondrial import of enzyme precursor.

Horwich, A.L., Kalousek, F., Mellman, I.S. & Rosenberg, L.E. (1985b) EMBO J. 4, 1129-1135. A leader peptide is sufficient to direct mitochondrial import of a chimeric protein.

Housman, D., Jacobs-Lorena, M., Rajbhandary, U.L. & Lodish, H.F.(1970) Nature 227, 913-918. Initiation of haemoglobin synthesis by methionyl-tRNA.

Hurt, E.C., Pesold-Hurt, B. & Schatz, G. (1984a) EMBO J. 3, 3149-3156.
The amino terminal of an imported mitochondrial precursor polypeptide can direct cytoplasmic dihydrofolate reductase into the mitochondrial matrix.

Hurt, E.C., Pesold-Hurt, B. & Schatz, G. (1984b) FEBS Letts. 178, 306-310.
The cleavable prepiece of an imported mitochondrial protein is sufficient to direct cytosolic dihyrofolate reductase into the mitochondrial matrix.

Hurt, E.C., Müller, U. & Schatz, G. (1985a) EMBO J. 4, 3509-3518.
The first twelve amino acids of a yeast mitochondrial outer membrane protein can direct a nuclear-endcoded cytochrome oxidase subunit to the mitochondrial inner membrane.

Hurt, E.C., Pesold-Hurt, B., Suda, K., Oppliger, W. & Schatz, G. (1985b) EMBO J. 4, 2061-2068. The first twelve amino acids (less than half of the pre-sequence of an imported mitochondrial protein can direct mouse cytosolic dihydrofolate reductase into the yeast mitochondrial matrix.

I:
Imoto, T., Johnson, L.N., North, A.C.T., Phillips, D.C. & Rupley, J.A. (1973) In 'The enzymes' (Boyer,P.D. ed) vol 7, pp665-868. Academic Press, London.
Vertebrate lysozymes.

Innis, M.A., Tokunaga, A.M., Williams, M.E., Loranger, J.M., Chang, S-Y., Chang, S. & Wu, H.C. (1984) Proc.Natl.Acad.Sci.USA 81, 3708-3712. Nucleotide sequence of the *Escherichia coli* prolipoprotein signal peptidase (lsp) gene.

Inouye, M. & Halegoua, S. (1980) CRC Crit.Rev.Biochem. 7, 339-371.
Secretion and membrane localization of proteins in *Escherichia coli*.

Inouye, S., Soberon, X., Franceschini, T., Nakamura, K., Itakura, K. & Inouye, M. (1982) Proc.Natl.Acad.Sci.USA 79, 3438-3441. Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane.

Inouye, S., Vlasuk, G.P., Hsiung H. & Inouye, M. (1984) J.Biol.Chem. 259, 3729-3733. Effects of mutations at glycine residues in the hydrophobic region of the *Escherichia coli* prolipoprotein signal peptide on the secretion across the membrane.

Ito, K., Mandell, G. & Wickner, W. (1979)
Proc.Natl.Acad.Sci.USA 76, 1199-1203. Soluble precursor of an integral membrane protein: synthesis of procoat protein in *Escherichia coli* infected with bacteriophage M13.

Ito, K., Date, T. & Wickner, W. (1980) J.Biol.Chem. 255, 2123-2130.
Synthesis, assembly into the cytoplasmic membrane, and proteolytic processing of the precursor of coliphage M13 coat protein.

J:
Jackson, M.E., Pratt, J.M., Stoker, N.G. & Holland, B.I. (1985)
EMBO J. 4, 2377-2383. An inner membrane N-terminal signal sequence is able to promote efficient localisation of an outer membrane protein in *Escherichia coli*.

Jackson, R.C. & Blobel, G. (1977) Proc.Natl.Acad.Sci.USA 74, 5598-5602. Post-translational cleavage of presecretory proteins with an extract of rough microsomes from dog pancreas containing signal peptidase activity.

Jackson, R.C. & White, W.R. (1981) J.Biol.Chem. 256. 2545-2550.
Phospholipid is required for the processing of presecretory proteins by detergent solubilized canine pancreatic signal peptidase.

Jackson, R.C., Walter, P. & Blobel, G. (1980) Nature 286. 174-176. Secretion requires a cytoplasmically disposed sulphydryl of the RER membrane.

Jansen, M., Van Schaik, F.M.A., Ricker, A.T., Bullock, B., Woods, D.E., Gabbay, K.H., Nussbaum, A.L., Sussenbach, J.S. & Van Den Brande, J.L. (1983) Nature 306. 609-611. Sequence of a cDNA encoding human insulin-like growth factor.

Jones, P. & Nicholson, B.H. (1979) Biochem.Soc.Trans. 7, 1253-1255. Translational activity of RNA from ruminant abomasum.

Josefsson, L-G. & Randall, L.L. (1981) Cell 25, 151-157.
Different exported proteins in *E.coli* show differences in the temporal mode of processing in vivo.

Jung, A., Sippel, A.E., Gretz, M. & Schutz G. (1980)
Proc.Natl.Acad.Sci.USA 77, 5759-5763. Exons encode the functional and structural units of chicken lysozyme.

K:
Kato, Y., Iwase, H. & Hotta, K. (1984) Anal.Biochem. 138, 437-441. Preparation of eight ovalbumin subfractions by combined lectin affinity chromatography.

Kelly, R.B. (1985) Science 230, 25-32.
Pathways of protein secretion in eukaryotes.

Kessler, S.W. (1975) J.Immunol. 115, 1617-1624. Rapid isolation of antigens from cells with a staphylococcal protein A-antibody adsorbent: Parameters of the interaction of antibody-antigen complexes with protein A.

Konarska, M.M., Padgett, R.A. & Sharp, P.A. (1984) Cell 38, 731-736. Recognition of cap structure in splicing *in vitro* of mRNA precursors.

Kondor-Koch, C., Burke, B. & Garoff, H. (1983) J.Cell Biol. 97, 644-651. Expression of Semliki Forest Virus proteins from cloned complementary DNA. I. The fusion activity of the spike glycoprotein.

Koren, R., Burstein, Y. & Soreq, H. (1983) Proc.Natl.Acad.Sci.USA 80, 7205-7209. Synthetic leader peptide modulates secretion of proteins from microinjected *Xenopus* oocytes.

Kornfeld, R. & Kornfeld, S. (1985) Ann.Rev.Biochem. 54, 631-664. Assembly of asparagine-linked oligosaccharides.

Koshland, D. & Botstein, D. (1980) Cell 20, 749-760. Secretion of beta-lactamase requires the carboxy end of the protein.

Koshland, D. & Botstein, D. (1982) Cell 30, 893-902. Evidence for posttranslational translocation of β-lactamase across the bacterial inner membrane.

Kozak, M. (1983a) Cell 34, 971-978. Translation of insulin-related polypeptides from messenger RNAs with tandemly reiterated copies of the ribosome binding site.

Kozak, M, (1983b) Nucl.Acid.Res. 12, 857-872. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs.

Kozak, M. (1984) Nature 308, 241-246. Point mutations close to the AUG initiation codon affect the efficiency of translation of rat preproinsulin.

Kreibich, G., Ulrich, B.L. & Sabatini D.D. (1978a) J.Cell Biol. 77, 464-487. Proteins of rough microsomal membranes related to ribosome binding I Identification of ripbophorins I and II, membrane proteins characteristic of rough microsomes.

Kreibich, G., Freienstein, C.M., Pereyra, B.N., Ulrich, B.L. & Sabatini, D.D. (1978b) J.Cell Biol. 77, 488-506. Proteins of rough microsomal membranes related to ribosome binding II Cross-linking of bound ribosomes to specific membrane proteins exposed at the binding sites.

Kreil, G. (1981) Ann.Rev.Biochem. 50, 317-348. Transfer of proteins across membranes.

Krieg, P.A. & Melton, D.A. (1984) Nucl.Acid.Res. 12, 7075-7070. Functional messenger RNAs are produced by SP6 *in vitro* transcription of cloned cDNAs.

Krieg, P., Strachan, R., Wallis, E., Tabe, L. & Colman, A. (1984) J.Mol.Biol. 180. 615-643. Efficient expression of cloned complementary DNAs for secretory proteins after injection into *Xenopus* oocytes.

Kronenberg, H.M., Fennick, B.J. & Vasicek, T.J. (1983) J.Cell Biol. 96, 1117-1119. Transport and cleavage of bacterial pre-β-lactamase by mammalian microsomes.

L:
Laemmli, U.K. (1970) Nature 227. 680-685. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4.

Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. (1981) Nucl.Acid.Res. 9, 2251-2266. 5'-terminal sequences of eukaryotic mRNA can be cloned with high efficiency.

Lane, C. (1983) In 'Current Topics in Developmental Biology'. Vol.18. Academic Press Inc. UK. pp89-116. The fate of genes, messengers and proteins introduced into *Xenopus* oocytes.

Lane, C.D., Shanon, S. & Craig, R. (1979) Eur.J.Biochem. 101, 485-495. Sequestration and turnover of guinea-pig milk proteins and chicken ovalbumin in *Xenopus* oocytes.

Lane, C.D., Colman, A., Mohun, T., Morser, J., Champion, J., Kourides, I., Craig, R., Higgins, S., James, T.C., Appelbaum, S.W., Ohlsson, R.I., Paucha, E., Houghton, M., Matthews, J. & Miflin, B.J. (1980) Eur.J.Biochem. 111. 225-235. The *Xenopus* oocyte as a surrogate secretory system.

Lane, C., Champion, J., Haiml, L. & Kreil, G. (1981) Eur.J.Biochem. 113. 273-281. The sequestration, processing and retention of honey-bee promellittin made in amphibian oocytes.

Lane, C.D., Champion, J. & Craig, R. (1983) Eur.J.Biochem. 130. 529-535. Signal sequences, secondary modification and the turnover of miscompartmenalized secretory proteins in *Xenopus* oocytes.

Lang, L., Reitman, M.R., Tang, J., Roberts, R.M. & Kornfeld, S.(1984) J.Biol.Chem. 259. 14663-14671. Lysosomal enzyme phosphorylation. Recognition of a protein-dependent determinant allows specific phosphorylation of oligosaccharides present on lysosomal enzymes.

Laskey, R.A. & Mills, A.D. (1975) Eur.J.Biochem. 56, 335-341. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography.

Lauffer, L., Garcia, P.D., Harkins, R.N., Coussens, L., Ullrich, A. & Walter, P. (1985) Nature 318. 334-338. Topology of signal recognition particle receptor in endoplasmic reticulum membrane.

Li, W-Y., Reddy, R., Henning, D., Epstein, P. & Busch, H. (1982) J.Biol.Chem. 257. 5136-5142. Nucleotide sequence of 7S RNA.

Lingappa, V.R., Katz, F.N., Lodish, H.F. & Blobel, G. (1978a) J.Biol.Chem. 253. 8667-8670. A signal sequence for the insertion of a transmembrane glycoprotein.

Lingappa, V.R., Shields, D., Woo, S.L.C. & Blobel, G. (1978b) J.Cell Biol. 79, 567-572. Nascent ovalbumin contains the functional equivalent of a signal sequence.

Lingappa, V.R., Lingappa, J.R. & Blobel, G. (1979) Nature 281, 117-121. Chicken ovalbumin contains an internal signal sequence.

Lingappa, V.R., Chaidez, J., Yost, C.S. & Hedgpeth, J. (1984) Proc.Natl.Acad.Sci.USA 81, 456-460. Determinants for protein localization: β-lactamase signal sequence directs globin across microsomal membrane.

Lively, M.O. & Walsh, K. (1983) J.Biol.Chem. 258, 9488-9495. Hen oviduct signal peptidase is an integral membrane protein.

Lodish, H.F. & Kong, N. (1984) J.Cell Biol. 98, 1720-1729. Glucose removal from N-linked oligosaccharides is required for efficient maturation of certain secretory glycoproteins from the rough endoplasmic reticulum to the Golgi complex.

Lodish, H.F., Braell, W.A., Schwartz, A.L., Strous, G.J.A.M. & Zilberstein, A. (1981) Int.Rev.Cytol.Suppl. 12, 247-307. Synthesis and assembly of membrane and organelle proteins.

Lodish, H.F., Kong, N., Snider, M. & Strous, G.J.A.M. (1983) Nature 304, 80-83. Heptoma secretory proteins migrate from rough endoplasmic reticulum to Golgi at characteristic rates.

Mc:
McConnell, D.J., Hughes, D., Chen, Y-F., Carroll, D.P., Donnelly, W.J., Ollington, J.F., Quinn, D., Barry, T. & Gannon, F. (1984) Biochem.Soc Trans. 12, 487-488. Towards the cloning of the calf prochymosin gene.

McKnight, S. & Kingsbury, R. (1982) Science, 217, 316-324. Transcriptional control signals of a eukaryotic protein-coding gene.

M:
Maccecchini, M-L., Rudin, Y., Blobel, G. & Schatz, G. (1979) Proc.Natl.Acad.Sci.USA 76, 343-347. Import of proteins into mitochondria: Precursor forms of the extramitochondrially made F_1-ATPase subunit in yeast.

Machamer, C., Florkiewicz, R.Z. & Rose, J.K. (1985) Mol.Cell.Biol. 5, 3074-3083. A single N-linked oligosaccharide at either of the two normal sites is sufficient for transport of Vesicular Stomatitis Virus G protein to the cell surface.

Majzoub, J.A., Rosenblatt, M., Fennick, B., Maunus, R., Kronenberg, H.M., Potts, J.T. & Habener, J.F. (1980) J.Biol.Chem. 255, 11478-11483. Synthetic pre-proparathyroid hormone leader sequence inhibits cell-free processing of placental, parathyroid, and pituitary prehormones.

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982) Cold Spring Harbor Laboratory, USA. 'Molecular cloning: a laboratory manual'.

Markoff, L., Lin, B-C., Sveda, M.M. Lai, C-J. (1984) Mol.Cell.Biol. 4, 8-16. Glycosylation and surface expression of the influenza virus neuraminidase requires the N-terminal hydrophobic region.

Maxam, A. & Gilbert, W. (1980) Meths.Enzymol. 65, 499-599. Sequencing end-labeled DNA with base specific chemical cleavages.

Meek, R.L., Walsh, K.A. & Palmiter, R.D. (1982) J.Biol.Chem. 257, 12245-12251. The signal sequence of ovalbumin is located near the NH_2 terminus.

Mellor, J., Dobson, M.J., Roberts, N.A., Tuite, M.F., Emtage, E.J.S., White, S., Lowe, P.A., Patel, T., Kingsman, A.J. & Kingsman, S.M. (1983) Gene 24, 1-14. Efficient synthesis of enzymatically active calf chymosin in *Saccharomyces cerevisiae*.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. & Green, M.R. (1984) Nucl.Acid.Res. 12, 7035-7056. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter.

Mertz, J.E. & Gurdon, J.B. (1977) Proc.Natl.Acad.Sci.USA 74, 1502-1506. Purified DNAs are transcribed after microinjection into *Xenopus* oocytes.

Messing, J. (1983) Meths.Enzymol. 101, 20-78. New M13 vectors for cloning.

Messing, J., Gronenborn, B., Muller-Hill, B. & Hofschneider, P.H. (1977) Proc.Natl.Acad.Sci.USA 74, 3642-3646. Filamentous coliphage M13 as a cloning vehicle: insertion of a HindII fragment of the *lac* regulatory region in M13 replicative form *in vitro*.

Meyer, D.I. (1982) TIBS 7, 320-321.
The signal hypothesis - a working model.

Meyer, D.I. (1985) EMBO J.4, 2031-2033.
Signal recognition particle (SRP) does not mediate a translational arrest of nascent secretory proteins in mammalian cell-free systems.

Meyer, D.I. & Dobberstein, B. (1980a) J.Cell Biol. 87, 498-502.
A membrane component essential for vectorial translocation of nascent proteins across the endoplasmic reticulum: requirements for its extraction and reassociation with membrane.

Meyer, D.I. & Dobberstein, B. (1980b) J.Cell Biol. 87, 503-508. Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum.

Meyer, D.I., Louvard, D. & Dobberstein, B. (1982a) J.Cell Biol. 92, 579-583. Characterization of molecules involved in protein translocation using specific antibody.

Meyer, D.I., Krause, E. & Dobberstein, B. (1982b) Nature 297, 647-650. Secretory protein translocation across membranes - the role of the docking protein.

Michaelis, S. & Beckwith, J. (1982) Ann.Rev.Microbiol. 36, 435-465. Mechanism of incorporation of cell envelope proteins in Escherichia coli.

Milstein, C., Brownlee, G.G., Harrison, T.M. & Matthews, M.B. (1972) Nat.New Biol. 239, 117-120. A possible precursor of immunoglobulin light chains.

Mishina, M., Tobimatsu, T., Imoto, K., Tanaka, K., Fujita, Y., Fukuda, K., Kurasaki, M., Takahashi, H., Morimoto,Y., Hirose,T., Inayama,S., Takahashi,T., Kuno, M. & Numa, S. (1985) Nature 313, 364-369. Location of functional regions of acetylcholine receptor α-subunit by site-directed mutagenesis.

Mishkind, M.L., Wessler, S.R. & Schmidt, G.W. (1985) J.Cell Biol. 100, 226-234. Functional determinants in transit sequences: import and partial maturation by vasular plant chloroplasts of the ribulose-1,5,-bisphosphate carboxylase small subunit of *Chlamydomonas*.

Moir, D., Mao, J., Schumm, J., Vovis, G.F., Alford, B.L. & Taunton-Rigby, A. (1982) Gene 19, 127-138. Molecular cloning and characterization of double-stranded cDNA coding for bovine chymosin.

Moreno, F., Fowler, A.V., Hall, M., Silhavy, T.J. Zabin, I. & Schwartz, M. (1980) Nature 286, 356-359. A signal sequence is not sufficient to lead ß-galactosidase out of the cytoplasm.

Morrison, T.G., McQuain, C.O. & Simpson, D. (1978) J.Virol. 28, 368-374. Assembly of viral membranes: maturation of the vesicular stomatitis virus glycoprotein in the presence of tunicamycin.

Mosmann, T.R. & Williamson, A.R. (1980) Cell 20, 283-292. Structural mutations in a mouse immunoglobulin light chain resulting in a failure to be secreted.

Mosmann, T.R., Baumal, R. & Williamson, A.R. (1979) Eur.J.Immunol. 9, 511-516. Mutations affecting immunoglobulin light chain secretion by myeloma cells: I. Functional analysis by cell fusion.

Müller, M., Ibrahami, I., Nan Chang, C., Walter, P. & Blobel, G. (1982) J.Biol.Chem. 257, 11860-11863. A bacterial secretory protein requires signal recognition particle for translocation across mammalian endoplasmic reticulum.

Müller, W., Weber, H., Meyer, F. & Weissmann, C. (1978) J.Mol.Biol. 124, 343-358. Site-directed mutagenesis in DNA: generation of point mutations in cloned β globin complementary DNA at the positions corresponding to amino acids 121-123.

Mulligan, R.C. & Berg, P. (1980) Science 209, 1422-1427.
Expression of a bacterial gene in mammalian cells.

Myers, R.M., Lerman, L.S. & Maniatis, T. (1985) Science 229, 242-247.
A general method for saturation mutagenesis of cloned DNA fragments.

N:
Nagahari, K., Kanaya, S., Munakata, K., Aoyagi, Y. & Mizushima, S. (1985) EMBO J. 4, 3589-3592. Secretion into the culture medium of a foreign product from *Escherichia coli*: use of the *ompF* gene for secretion of human-β-endorphin.

Neupert, W. & Schatz, G. (1981) TIBS. 6, 1-4.
How proteins are transported into mitochondria.

Nicholson, B.H. & Jones, P. (1984) J.Dairy Res. 51, 79-89. Isolation and *in vitro* translation of messenger-RNA from the calf abomasal mucosa and identification of a messenger-RNA coding for a precursor of prochymosin.

Nishimori, K., Kawaguchi, Y., Hidaka, M., Uozumi, T. & Beppu, T. (1981) J.Biochem. 90, 901-904. Cloning in *Escherichia coli* of the structural gene of prorennin, the precursor of calf milk-clotting enzyme rennin.

Nishimori, K., Kawaguchi, Y., Hidaka, M., Uozumi, T. & Beppu, T. (1982a) Gene 19, 337-344. Expression of cloned calf prochymosin gene sequence in *Escherichia coli*.

Nishimori, K., Kawaguchi, Y., Hidaka, M., Uozumi, T. & Beppu, T. (1982b) J.Biochem. 91, 1085-1088. Nucleotide sequence of calf prorennin cloned in *Escherichia coli*.

Nishimori, K., Shimuzu, N., Kawaguchi, Y., Hidaka, M., Uozumi, T. & Teruhiko, B. (1984) Gene 29, 41-49. Expression of cloned calf prochymosin cDNA under control of the tryptophan promoter.

O:
Ohashi, A., Gibson, J., Gregor, I. & Schatz, G. (1982) J.Biol.Chem. 257, 13042-13047. Import of proteins into mitochondria. The precursor of cytochrome c_1 is processed in two steps, one of them heme-dependent.

Old, R.W., Woodland, H.R., Ballantine, J.E.M., Aldridge, T.C., Newton, C.A., Bains, W.A. & Turner, P.C. (1982) Nucl.Acid.Res. 10, 7561-7580. Organization and expression of cloned histone gene cluster from *Xenopus laevis* and *X.borealis*.

Olden, K., Parent, J.B. & White, S.L. (1982)
Biochim.Biophys.Acta 650, 209-232. Carbohydrate moieties of glycoproteins a re-evaluation of their function.

Oliver, D.B. (1985a) Ann.microbiol. 39, 615-648.
Protein secretion in E.coli.

Oliver, D.B. (1985b) J.Bact. 161, 285-291.
Identification of five new essential genes involved in the synthesis of a secreted protein in *Escherichia coli*

Oliver, D.B. & Beckwith, J. (1981) Cell 25, 765-772.
E.coli mutant pleiotropically defective in the export of secreted proteins.

P:
Palade,G. (1975) Science 189. 347-358.
Intracellular aspects of the process of protein synthesis.

Palmiter, R.D., Gagnon, J., Ericsson, L.H. & Walsh, .K.A. (1977) J.Biol.Chem. 252. 6386-6393. Precursor of egg white lysozyme.

Palmiter, R.D., Gagnon, J. & Walsh, .K.A. (1978)
Proc.Natl.Acad.Sci.USA 75, 94-98. Ovalbumin: a secreted protein without a transient hydrophobic leader sequence.

Pederson, V.B. & Foltmann, B. (1975) Eur.J.Biochem. 55, 95-103. Amino acid sequence of the peptide segment liberated during activation of prochymosin.

Pelham, H.R.B. & Jackson, R.S. (1976) Eur.J.Biochem. 67, 247-256. An efficient mRNA-dependent translation system from reticulocyte lysates.

Pelletier, J. & Sonenberg, N. (1985) Cell 40, 515-526. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency.

Perara, E. & Lingappa, V.R. (1985) J.Cell Biol. 101. 2292-2301.
A former amino terminal signal sequence engineered to an internal location directs translocation of both flanking protein domains.

Perlman, D. & Halvorson, H.O. (1983) J.Mol.Biol. 167, 391-409.
A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides.

Pfeiffer, S., Fuller, S.D. & Simons, K. (1985) J.Cell Biol. 101. 470-476. Intracellular sorting and basolateral appearance of the G protein of Vesicular Stomatitis Virus in Madin-Darby canine kidney cells.

Pfisterer, J., Lachmann, P. & Kloppstech, K. (1982) Eur.J.Biochem. 126, 143-148. Transport of proteins into chloroplasts: Binding of nuclear-encoded chloroplast proteins to the chloroplast envelope.

Poruchynsky, M.S., Tyndall, C., Both, G.W., Sato, F., Bellamy, A.R. & Atkinson, P.H. (1985) J.Cell Biol. 101, 2199-2209. Deletions into an NH_2-terminal hydrophobic domain result in secretion of a rotavirus VP7, a resident endoplasmic reticulum membrane glycoprotein.

Prehn, S., Tsamaloukas, A. & Rapoport, T.A. (1980) Eur.J.Biochem 107, 185-195. Demonstration of specific receptors of the rough endoplasmic reticulum membrane for the signal sequence of carp preproinsulin.

Prehn, S., Nunberg, P. & Rapoport, T.A. (1981) FEBS Letts. 123. 79-84.
A receptor for signal segments of secretory proteins in rough endoplasmic reticulum membranes.

R:
Rachubinski, R.A., Verma, D.P.S. & Bergeron, J.J.M. (1980)
J.Cell Biol. 84, 705-716. Synthesis of rat liver microsomal cytochrome b_5 by free ribosomes.

Randall, L.L. (1983) Cell 33, 231-240. Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation.

Randall, L.L. & Hardy, S.J.S. (1977) Eur.J.Biochem 75, 43-53. Synthesis of exported proteins by membrane-bound polysomes from *Escherichia coli*

Rapoport, T.A. (1981) Eur.J.Biochem. 115, 665-669.
Intracellular compartmentation and secretion of carp proinsulin synthesized in *Xenopus* oocytes.

Redman, C.M. & Sabatini, D.D. (1966) Proc.Natl.Acad.Sci.USA. 56, 608-615. Vectorial discharge of peptides released by puromycin from attatched ribosomes.

Reitman, M.L. & Kornfeld, S. (1981) J.Biol.Chem. 256, 4275-4281. UDP-N-acetylglucosamine: glycoprotein N-acetylglucosamine-1-phospho-transferase. Proposed enzyme for the phosphorylation of the high mannose oligosaccharide units of lysosomal enzymes.

Riezman, H., Hay, R., Witte, C., Nelson, N. & Schatz, G. (1983) EMBO J. 2, 1113-1118. Yeast mitochondrial outer membrane specifically binds cytoplasmically synthesized precursors of mitochondrial proteins.

Rindler, M.J., Ivanov, I.E., Plesken, H. & Sabatini, D.D. (1985)
J.Cell Biol. 100. 136-151. Polarized delivery of viral glycoproteins to the apical and basolateral plasma membranes of Madin-Darby canine kidney cells infected with temperature-sensitive viruses.

Robbins, A.R. & Myerowitz, R. (198) J.Biol.Chem. 256, 10623-10627. The mannose-6-phosphate receptor of Chinese Hamster ovary cells.

Roberts, B.E. & Paterson, B.M. (1973) Proc.Natl.Acad.Sci.USA 70, 2330-2334. Efficient translation of tobacco mosiac virus RNA and globin 9S RNA in a cell-free system from commercial wheat germ.

Robinson, C. (1984) PhD. thesis, University of Warwick.
Uptake and processing of chloroplast polypeptides.

Robinson, C. & Ellis, R.J. (1984) Eur.J.Biochem. 142, 343-346. Transport of proteins into chloroplasts: the precursor of small subunit of ribulose bisphosphate carboxylase is processed to the mature size in two steps.

Rose, J.K. & Bergmann, J.E., (1982) Cell 30, 753-762.
Expression from cloned cDNA of cell-surface and secreted forms of the glycoprotein of vesicular stomatitis virus in eukaryotic cells.

Rosenblatt, M., Beaudette, N.V. & Fasman, G.D. (1980)
Proc.Natl.Acad.Sci.USA 77, 3983-3987. Conformational studies of the synthetic precursor specific region of preproparathyroid hormone.

Rothman, J.E. (1985) Sci.Amer. 253, 84-95.
The compartmental organization of the Golgi apparatus.

Rothman, J.E. & Lodish, H.F. (1977) Nature 269, 775-780. Synchronised transmembrane insertion and glycosylation of a nascent membrane protein.

Rothman, J.E., Katz, F.N. & Lodish, H.F. (1978) Cell 15, 1447-1454. Glycosylation of a membrane protein is restricted to the growing polypeptide chain but is not necessary for insertion as a transmembrane protein.

Rothman, J.E., Miller, R.L. & Urbani, L.J. (1984) J.Cell Biol. 99, 260-271. Intercompartmental transport in the Golgi complex is a dissociative process: facile transfer of membrane protein between two Golgi populations.

Rottier, P., Brandenburg, D., Armstrong, J., Van Der Zeijst, B. & Warren G. (1984) Proc.Natl.Acad.Sci.USA 81, 1421-1425. Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: The E1 glycoprotein of coronavirus mouse hepatitis virus A59.

Rottier, P., Armstrong, J. & Meyer, D.I. (1985) J.Biol.Chem. 260, 4648-4652. Signal recognition particle dependent insertion of coronavirus E1, an intracellular membrane protein.

S:
Sabatini, D.D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982)
J.Cell Biol. 92, 1-22. Mechanisms for the incorporation of proteins in membranes and organelles.

Sakaguchi, M., Mihara, K. & Sato, R. (1984) Proc.Natl.Acad.Sci.USA 81, 3361-3364. Signal recognition particle is required for co-translational insertion of cytochrome P450 into microsomal membranes.

Sanger, F., Nicklēn, S. & Coulson, A.R. (1977) Proc.Natl.Acad.Sci.USA 74, 5463-5467. DNA sequencing with chain terminating inhibitors.

Schatz, G & Butow, R.A. (1983) Cell 32, 316-318.
How are proteins imported into mitochondria.

Schauer, I., Emr, S., Gross, C. & Scheckman, R. (1985)
J.Cell Biol 100, 1664-1675. Invertase signal and mature sequence substitutions that delay intercompartmental transport of active enzyme.

Scheckman, R. (1982) TIBS. 7, 243-246. The secretory pathway in yeast.

Schechter, I., McKean, D.J., Guyer, R. & Terry, W. (1975)
Science 188, 160-162. Partial amino acid sequence of the precursor of immunoglobulin light chain programmed by messenger RNA in vitro.

Scheele, G. & Tartakoff, A. (1985) J.Biol.Chem. 260, 926-931.
Exit of nonglycosylated proteins from the rough endoplasmic reticulum is asynchronous in the exocrine pancreas.

Schleyer, M. & Neupert, W. (1985) Cell 43, 339-350.
Transport of proteins into mitochondria: translocational intermediates spanning contact sites between outer and inner membranes.

Schreier, P.H., Seftor, E.A., Schell, J. & Bohnert, H.J. (1985) EMBO J 4, 25-32. The use of nuclear-encoded sequences to direct the light-regulated synthesis and transport of a foreign protein into plant chloroplasts.

Sekikawa, K. & Lai, C-J. (1983) Proc.Natl.Acad.Sci.USA 80, 3563-3567. Defects in functional expression of an influenza virus haemagglutinin lacking the signal peptide sequence.

Shortle, D. & Nathans, D. (1978) Proc.Natl.Acad.Sci.USA 75, 2170-2174. Local mutagenesis: A method for generating viral mutants with base subsitutions in preselected regions of the viral genome.

Shortle, D. & Botstein, D. (1982) Meths.Enzymol. 100. 457-468. Directed mutagenesis with sodium bisulphite.

Shortle, D., Dimaio, D. & Nathans, D. (1981) Ann.Rev.Genet. 15, 265-294. Directed mutagenesis.

Shuman, H.A. (1981) J.Memb.Biol. 61, 1-11. The use of gene fusions to study bacterial transport proteins.

Silhavy, T.J., Benson, S.A. & Emr, S.D. (1983) Microbiol. Rev. 47, 313-344. Mechanisms of protein localization.

Sly, W.S. & Fischer, W.D. (1982) J.Cell.Biochem. 18, 67-85. The phosphomannosyl recognition system for intracellular transport of lysosomal enzymes.

Smith, M. (1982) TIBS 7, 440-442. Site-directed mutagenesis.

Smith, M. & Gillam, S. (1981) In 'Developmental biology using purified genes' ICN-UCLA Symp.Mol.Cell.Biol. (Brown,D.D. & Fox, C.F. eds) vol 23, pp671-682. Academic Press, New York. 'In vitro construction of specific mutants.'

Smith, S.M. & Ellis, R.J. (1979) Nature 278, 662-664. Processing of small subunit precursor of ribulose bisphosphate carboxylase and its assembly into whole enzyme are stromal events.

Smith, W.P. (1980) J.Bact. 141. 1142-1147. Cotranslational secretion of diptheria toxin and alkaline phosphatase in vitro: involvement of membrane protein(s).

Smith, W.P., Tai, P-C., Thompson, R.C. & Davis, B.D. (1977) Proc.Natl.Acad.Sci.USA 74, 2830-2834. Extracellular labelling of nascent polypeptides traversing the membrane of *Escherichia coli*.

Soreq, H. (1985) CRC.Crit.Rev.Biochem. 18, 199-238. Biosynthesis of biologically active proteins in mRNA - microinjected oocytes of *Xenopus laevis*.

Strauss, A.W. & Boime, I. (1982) CRC.Crit.Rev.Biochem. 12, 205-235. Compartmentation of newly synthesized proteins.

Strous, G.J.A.M., Willemsen, R., Van Kerkhof, P., Slot, J.W., Geuze, H.J. & Lodish, H.F. (1983) J.Cell Biol. 97, 1815-1822. Vesicular stomatitis virus glycoprotein, albumin and transferrin are transported to the cell surface via the same Golgi vesicles.

Sugimoto, K., Sugisaki, H., Okamoto, T. & Takanami, M. (1977) J.Mol.Biol. 111, 487-507. Studies on bacteriophage fd DNA. IV. The sequence of messenger RNA for the major coat protein gene.

Suissa, M. & Schatz, G. (1982) J.Biol.Chem. 257, 13048-13055. Import of proteins into mitochondria. Translatable mRNAs for imported mitochondrial proteins are present in free as well as mitochondrial-bound cytoplasmic polysomes.

Szczesna, E. & Boime, I (1976) Proc.Natl.Acad.Sci.USA 73, 1179-1183. mRNA dependent synthesis of authentic precursor to human placental lactogen: conversion to its mature form in ascites cell-free extract.

T:
Tabe, L., Krieg, P., Strachan, R., Jackson, D., Wallis, E. & Colman, A. (1984) J.Mol.Biol. 180, 645-666. Segregation of mutant ovalbumins and ovalbumin-globin fusion proteins in *Xenopus* oocytes.

Talmadge, K., Stahl, S. & Gilbert, W. (1980)
Proc.Natl.Acad.Sci.USA 77, 3369-3373. Eukaryotic signal sequence transports insulin antigen in *Escherichia coli*.

Talmadge, K., Brosius, J. & Gilbert, W. (1981) Nature 294, 176-178. An 'internal' signal sequence directs secretion and processing of proinsulin in bacteria.

Tartakoff, A.M. (1982) TIBS. 7, 174-176. Simplifying the complex Golgi.

Tokunaga, M., Loranger, J.M., Wolfe, P.B. & Wu, H.C. (1982)
J.Biol.Chem. 257. 9922-9925. Prolipoprotein signal peptidase in *Escherichia coli* is distinct from the M13 procoat signal peptidase.

Traboni, C., Cortese, R., Cilberto, G & Cesareni, G. (1983) Nucl.Acids Res. 11, 4229-4239. A general method to select for M13 clones carrying base pair substitution mutants constructed *in vitro*.

U:
Uchiyama, H., Uozumi, T., Beppu, T. & Arima, K. (1980) Agric.Biol.Chem. 44, 1373-1381. Purification of prorennin mRNA and its translation *in vitro*.

Ullu, E., Murphy, S. & Melli, M. (1982) Cell 29, 195-202. Human 7SL RNA consists of a 140 nucleotide middle-repetitive sequence inserted in an Alu sequence.

V:
Valle, G., Besley, J. & Colman, A. (1981) Nature 291. 338-340. Synthesis and secretion of mouse immunoglobulin chains from *Xenopus* oocytes.

Valle, G., Besley, J., Williamson, A.R., Mosmann, T.R. & Colman, A. (1983) Eur.J.Biochem. 132. 131-138. Post-translational fate of variant MOPC 315 λ chains in *Xenopus* oocytes and mouse myeloma cells.

Van Den Broeck, V., Timko, M.P., Kausch, A.P., Cashmore, A.R., Van Montagu,M. & Herrera-Estrella, L. (1985) Nature 313. 358-363. Targetting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase.

Varki, A. & Kornfeld, S. (1981) J.Biol.Chem. 256, 9937-9943.
Purification and characterization of rat liver α-N-acetylglucosaminyl phosphodiesterase.

Viebrock, A., Pertz, A. & Sebald, W. (1982) EMBO J 1, 565-571. The imported preprotein of the proteolipid subunit of the mitochondrial ATP synthase from *Neurospora crassa*. Molecular cloning and sequencing of mRNA.

Vlasuk, G.P., Inouye, S., Ito, H., Itakura, K. & Inouye, M. (1983)
J.Biol.Chem. 258, 7141-7148. Effects of the complete removal of basic amino acid residues from the signal peptide on secretion of lipoprotein in *Escherichia coli*.

Von Heijne, G. (1981a) Biochem.Soc.Symp. 46, 259-273.
Models for transmembrane translocation of proteins.

Von Heijne, G. (1981b) Eur.J.Biochem 120, 275-278. Membrane proteins. The amino acid composition of membrane-penetrating segments.

Von Heijne, G. (1983) Eur.J.Biochem. 133, 17-21.
Patterns of amino acids near signal-sequence cleavage sites.

Von Heijne, G. (1984a) J.Mol.Biol. 173. 243-251.
How signal sequences maintain cleavage specificity.

Von Heijne, G. (1984b) EMBO J. 3, 2315-2318. Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells.

Von Heijne, G. (1985) J.Mol.Biol. 184, 99-105.
Signal sequences: the limits of variation.

Von Heijne, G. & Blomberg, C. (1979) Eur.J.Biochem 97, 175-181.
Trans-membrane translocation of proteins. The direct transfer model.

W:
Waheed, A., Pohlmann, R., Hasilik, A. & Von Figura, K. (1981a)
J.Biol.Chem. 256, 4150-4152. Subcellular location of two enzymes involved in the synthesis of phosphorylated recognition markers in lysosomal enzymes.

Waheed, A., Pohlmann, R., Hasilik, A. & Von Figura, K. (1981b)
J.Biol.Chem. 256. 5717-5721. Processing of the phosphorylated recognition markers in lysosomal enzymes.

Walter, P. & Blobel, G. (1980) Proc.Natl.Acad.Sci.USA 77, 7112-7116.
Purification of a membrane-associated protein complex required for protein translocation across the endplasmic reticulum.

Walter, P. & Blobel, G. (1981a) J.Cell Biol 91, 551-556.
Translocation of proteins across the endoplasmic reticulum. II Signal recognition protein (SRP) mediates the selective binding to microsomal membranes of in-vitro-assembled polysomes synthesizing secretory protein.

Walter, P. & Blobel, G. (1981b) J.Cell Biol 91, 557-561. Translocation of proteins across the endoplasmic reticulum. III Signal recognition protein (SRP) causes signal sequence dependent and site-specific arrest of chain elongation that is released by microsomal membranes.

Walter, P. & Blobel, G. (1982) Nature 299, 691-698.
Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum.

Walter, P. & Blobel, G. (1983a) Cell 34, 525-533.
Disassembly and reconstitution of signal recognition particle.

Walter, P. & Blobel, G. (1983b) J.Cell Biol. 97, 1693-1699. Subcellular distribution of signal recognition particle and 7SL-RNA determined with polypeptide-specific antibodies and complementary DNA probe.

Walter, P., Jackson, R.C., Marcus, M.M., Lingappa, V.R. & Blobel, G. (1979) Proc.Natl.Acad.Sci.USA 76, 1795-1799. Tryptic digestion and reconstitution of translocation activity for nascent presecretory proteins across microsomal membranes.

Walter, P., Ibrahami, I. & Blobel, G. (1981) J.Cell Biol 91, 545-550.
Translocation of proteins across the endoplasmic reticulum. I Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein.

Walter, P., Gilmore, R. & Blobel, G. (1984) Cell 38, 5-8.
Protein translocation across endoplasmic reticulum.

Warren, G. & Dobberstein, B. (1978) Nature 273, 569-571.
Protein transfer across microsomal membranes reassembled from separated membrane components.

Watson, M.E.E. (1984) Nucl.Acid.Res. 12, 5145-5164.
Compilation of published signal sequences.

Watts, C., Silver, P. & Wickner. W. (1981) Cell 25, 347-353.
Membrane assembly from purified compnonents. II Assembly of M13 procoat into liposomes reconstituted from pufified leader peptidase.

Watts, C., Wickner, W. & Zimmermann, R. (1983) Proc.Natl.Acad.Sci.USA 80, 2809-2813. M13 procoat and a pre-immunoglobulin share processing specifity but use different membrane receptor mechanisms.

Weiher, H. & Schaller, H. (1982) Proc.Natl.Acad.Sci.USA 79, 1408-1412.
Segment-specific mutagenesis: extensive mutagenesis of a lac promoter/ operator element.

Weiss, R.A., Teich, N.N., Varwes, H. & Coffin, J.M. (1984) In 'Molecular biology of tumor viruses'. Cold Spring Harbor Laboratory Press. USA. 2nd Edn. part 3.

Wickens, M.P. & Gurdon, J.B. (1983) J.Mol.Biol. 163, 1-26.
Post-transcriptional processing of Simian Virus 40 late transcripts in injected frog oocytes.

Wickens, M.P. & Laskey, R.A. (1981) In 'Genetic engineering' Vol.1 (Williamson, R. ed.) pp103-167, Academic Press, London. Expression of cloned genes in cell free systems and in *Xenopus* oocytes.

Wickens, M.P., Woo, S., O'Malley, B.W. & Gurdon, J.B. (1980)
Nature 258, 628-634. Expression of a chick chromosomal ovalbumin gene injected into frog oocyte nuclei.

Wickner, W. (1979) Ann.Rev.Biochem. 48, 23-45. The assembly of proteins into biological membranes. The membrane trigger hypothesis.

Wickner, W. (1980) Science 210, 861-868.
Assembly of proteins into membranes.

Wickner, W. (1983) TIBS. 8, 427-428.
Separate signal and trigger steps for bacterial protein export.

Wickner, W. & Lodish, H. (1985) Science 230, 400-407.
Multiple mechanisms of protein insertion into and across membranes.

Wiedman, M., Huth, A. & Rapoport, T.A. (1984) Nature 309, 637-639.
Xenopus oocytes are able to secrete bacterial β-lactamase.

Wieringa, B., Meyer, F., Reiser, J. & Weissmann, C. (1983) Nature 301, 38-43
Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β-globin gene.

Wilkie, N.M., Eglin, R.P., Sanders, P.G. & Clements, J.B. (1980) Proc.R.Soc.Lond.B. 210, 411-421. The association of Herpes Simplex Virus with squamous carcinoma of the cervix and studies of the virus thymidine kinase gene.

Wolfe, P.B. & Wickner, W. (1984) Cell 36, 1067-1072. Bacterial leader peptidase, a membrane protein without a leader peptide uses the same export pathway as pre-secretory proteins.

Wolfe, P.B., Wickner, W. & Goodman, J.M (1983) J.Biol.Chem. 258. 12073-12080. Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope.

Y:
Yamada, H., Yamagata, H. & Mizushimi, S. (1983) FEBS Letts. 166, 179-182. The major outer membrane lipoprotein and new lipoproteins share a common signal peptidase that exists in the cytoplasmic membrane of *Escherichia coli*.

Yost, C.S., Hedgepeth, J. & Lingappa, V.R. (1983) Cell 34, 759-766. A stop-transfer sequence confers predictable transmembrane orientation to a previously secreted protein in cell-free systems.

Z:
Zakour, R.A., James, E.A. & Lobb, L.A. (1984) Nucl.Acids Res. 12, 6615-6628. Site specific mutagenesis: insertion of single noncomplementary nucleotides at specific sites by error-directed DNA polymerization.

Zehavi-Willner, T. & Lane, C. (1977) Cell 11, 683-693. Subcellular compartmentation of albumin and globin made in oocytes under the direction of injected messenger RNA.

Zimmerman, R., Paluch, U. & Neupert, W. (1979) FEBS Letts. 108, 141-146. Cell-free synthesis of cytochrome c.

Zimmerman, R., Hennig, B. & Neupert, W. (1981) Eur.J.Biochem. 116, 455-460. Different transport pathways of individual precursor proteins in mitochondria.

Zoller, M.J. & Smith, M. (1983) Meths.Enzymol. 100. 468-500. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors.

Zwizinski, C. & Wickner, W. (1980) J.Biol.Chem. 255,7973-7977. Purification and characterization of leader (signal) peptidase from *Escherichia coli*.

Zwizinski, C., Date, T. & Wickner, W. (1981) J.Biol.Chem. 256, 3593-3597. Leader peptidase is found in both the inner and outer membranes of *Escherichia coli*.

-APPENDIX-

Reprinted from *J. Mol. Biol.* (1984) **180**, 615–643

Efficient Expression of Cloned Complementary DNAs for Secretory Proteins after Injection into *Xenopus* Oocytes

P. Krieg, R. Strachan, E. Wallis, L. Tabe
and A. Colman

J. Mol. Biol. (1984) **180**, 615–643

Efficient Expression of Cloned Complementary DNAs for Secretory Proteins after Injection into *Xenopus* Oocytes

P. Krieg†, R. Strachan, E. Wallis, L. Tabe and A. Colman

Medical Research Council Developmental Biology Group
Department of Biological Sciences, University of Warwick
Coventry CV4 7AL, U.K.

(Received 5 June 1984, and in revised form 3 September 1984)

Cloned complementary DNAs encoding chicken ovalbumin, chicken prelysozyme and calf preprochymosin, prochymosin and chymosin were inserted downstream from various viral promoters in modified recombinant "shuttle" vectors. Microinjection of the ovalbumin, prelysozyme and preprochymosin constructs into the nuclei of *Xenopus laevis* oocytes resulted in the synthesis, segregation in membranes and secretion into the extracellular medium of ovalbumin, lysozyme and prochymosin, respectively. Judging from molecular weight estimations, lysozyme and prochymosin were correctly proteolytically processed while ovalbumin, which lacks a cleavable signal sequence, was glycosylated. Injection of the DNA construct encoding prochymosin without its signal sequence resulted in synthesis of prochymosin protein that was localized exclusively in the oocyte cytoplasm. No immunospecific protein was detected after injection of the DNA encoding mature chymosin.

In terms of protein expression in oocytes, the *Herpes simplex* thymidine kinase (TK) promoter was up to sevenfold more effective than the simian virus 40 (SV40) early promoter, and equally as effective as the Moloney murine sarcoma virus long terminal repeat element. Where tested, protein expression in oocytes was much reduced if DNA sequences encoding the SV40 small t intron and its flanking sequences were present in the constructs.

S_1 nuclease mapping of transcripts produced after injection of DNAs containing the TK promoter indicated that the majority of transcripts initiated at, or within, two bases of the known "cap" site. However, minor transcripts initiating upstream from this site were observed and one (or more) of these transcripts was responsible for the synthesis of an ovalbumin polypeptide containing a 51 amino acid N-terminal extension. This extended protein remained in the oocyte cytosol.

When ovalbumin cDNA was inserted into the vectors with opposite polarity to the viral promoter, expression in oocytes resulted in the predominant synthesis and secretion of a variant ovalbumin with a 21 amino acid N-terminal extension, although some full-length ovalbumin was also synthesized and secreted. S_1 mapping revealed the presence, in these oocytes, of transcripts of predicted polarity initiating 118 bases upstream from the wild type ovalbumin initiator

† Present address: Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Mass. 02138, U.S.A.

0022–2836/84/350615–29 $03.00/0

ATG, at a previously unreported SV40 "promoter". No protein synthesis was detected after the injection of these reverse-orientation constructs into baby hamster kidney (BHK-21) cells.

1. Introduction

In recent years the *Xenopus* oocyte has achieved prominence as an efficient translational assay system (reviewed by Lane & Knowland, 1975; Asselbergs, 1979; Marbaix & Huez, 1980; Lane, 1983; Colman, 1984). The ability of the oocyte to correctly post-translationally modify many foreign proteins encoded by the injected mRNAs represents a particular advantage over cell-free translational systems (reviewed by Lane, 1981; Soreq, 1984). A further advantage of this system is that microinjection is a direct and quantitative method for introducing new genetic material into a cell, and the large size and metabolic activity of the oocyte makes biochemical analysis of even one cell a feasible proposition.

We have been studying the secretion of foreign proteins from *Xenopus* oocytes (reviewed by Lane *et al.*, 1980; Colman, 1982). In qualitative terms, the behaviour of nearly all proteins tested in this way is similar to that exhibited in the parental cell type. These results accord with the prevailing view that certain sorting sequences involved in protein segregation are of ancient evolutionary origin and can be effective in distant cell types (Sabatini *et al.*, 1982). Much of the evidence to support this view has come from studies in bacteria, where eukaryotic secretory proteins, including chick ovalbumin or rat preproinsulin, can be expressed and translocated through a membrane under the direction of their own, or a bacterial "signal" sequence (Fraser & Bruce, 1978; Talmadge *et al.*, 1980). However, the bacterial export pathway has only a limited resemblance to that operating in an animal cell, so many studies have been initiated where cloned eukaryotic DNA, manipulated *in vitro*, is introduced into animal cells. The effect of various mutations on the segregation of the expressed protein can then be followed (Gething & Sambrook, 1982; Sveda *et al.*, 1982; Rose & Bergman, 1983; Garoff *et al.*, 1983). The most successful studies of this type have used transient transfection or infection of cultured mammalian cells with the modified DNAs. In this paper we show how this strategy can be extended to the use of the oocyte.

Oocytes will transcribe exogenous DNAs after microinjection (reviewed by Gurdon & Melton, 1981; Colman, 1984). In several cases this transcription is faithful, and in the case of intron-containing transcripts, correct splicing has been observed (Wickens & Gurdon, 1983) or inferred (Rungger & Turler, 1978), although the efficiency is both sequence-specific and variable (D. A. Melton, personal communication). However, many eukaryotic promoters, including those for chick ovalbumin (Wickens *et al.*, 1980) and *Xenopus* α and β globins (Melton, personal communication), do not function in *Xenopus* oocytes, whereas others (e.g. SV40 (Wickens & Gurdon, 1983), thymidine kinase (McKnight & Kingsbury, 1982), sea urchin histone (Probst *et al.*, 1979) and *Xenopus* histone (Old *et al.*, 1982)) function quite well. We report here the efficient expression in oocytes of various secretory proteins after microinjections of DNAs containing fusions between viral promoters and cloned cDNAs. In the accompanying paper (Tabe *et al.*, 1984) we show how this expression can be exploited to investigate polypeptide regions involved in the segregation of one of these proteins, chick ovalbumin.

2. Materials and Methods

(a) *Chemicals and reagents*

All chemicals were of analytical grade and were purchased from British Drug Houses (Poole, U.K.). [^{35}S]methionine (150 to 400 Ci/mmol), [γ-^{32}P]ATP (approx. 5000 Ci/mmol), [α-^{32}P]dCTP (2000 Ci/mmol) and ^{14}C-labelled protein markers were obtained from the Radiochemical Centre (Amersham, U.K.).

Restriction enzymes, phage T4 polynucleotide kinase, *Hin*dIII linkers (G-C-A-A-G-C-T-T-G-C) and *Eco*RI linkers (G-G-A-A-T-T-C-C) were purchased from Bethesda Research Laboratories (U.K.). Phage T4 DNA ligase was a kind gift from C. Darnborough (Glasgow University). S_1 nuclease was from Sigma (U.K.) and calf intestinal phosphatase was from Boehringer (West Germany).

Rabbit anti-ovalbumin complete sera, and FITC-conjugated goat anti-rabbit immunoglobin were purchased from Miles (U.K.). Rabbit anti-prochymosin was a gift from P. Lowe (Celltech, U.K.), and rabbit anti-chick lysozyme was a kind gift from D. Cutler (University of Warwick, U.K.).

(b) *Construction of expression vectors and preparation of mRNAs*

Polyadenylated oviduct mRNA from laying Rhode Island Red hens was prepared as described (Cutler *et al.*, 1981). Preprochymosin mRNA was prepared from unweaned calves as described by Harris *et al.* (1982), and was a kind gift from T. Harris (Celltech, U.K.).

(c) *Source of recombinant DNAs*

The recombinant plasmid pOv230, which contains a full-length chicken ovalbumin cDNA† (McReynolds *et al.*, 1977) was a kind gift from M. Wickens (University of Wisconsin). The plasmid pls184, which contains all the chicken lysozyme coding sequence (Land *et al.*, 1981), was a gift from A. Sippel (Cologne, West Germany). *Bcl*I DNA fragments containing the coding sequences for calf preprochymosin, prochymosin and chymosin, as described by Mellor *et al.* (1983), were gifts from T. Harris (Celltech). Recombinants pSV2 SFV(d-1) and pSV 8-SFV(d-1), which contain Semliki Forest virus sequences (Garoff *et al.*, 1983; Kondor-Koch *et al.*, 1983), were gifts from H. Garoff (Heidelberg). Plasmids pTKMoLTR$_1$, which contains the Moloney murine sarcoma virus long terminal repeat (Dhar *et al.*, 1980) and pTK1, which contains the *Herpes simplex* thymidine kinase gene (Wilkie *et al.*, 1980), were both kind gifts from N. Wilkie (Beatson, Glasgow, U.K.).

(d) *Construction of expression vectors and DNA manipulation techniques*

In general, the techniques used to construct the different expression vectors were as described by Maniatis *et al.* (1982). The specific techniques and any modifications of the techniques described by Maniatis *et al.* (1982) are described in Fig. 1. Similar techniques were also used to manipulate the various eukaryotic DNAs before insertion into the vectors, and these are described in the appropriate Figures and legends.

† Abbreviations used: cDNA, complementary DNA; PMSF, phenylmethylsulphonyl fluoride; SDS, sodium dodecyl sulphate; TK, thymidine kinase; LTR, long terminal repeat; BHK, baby hamster kidney.

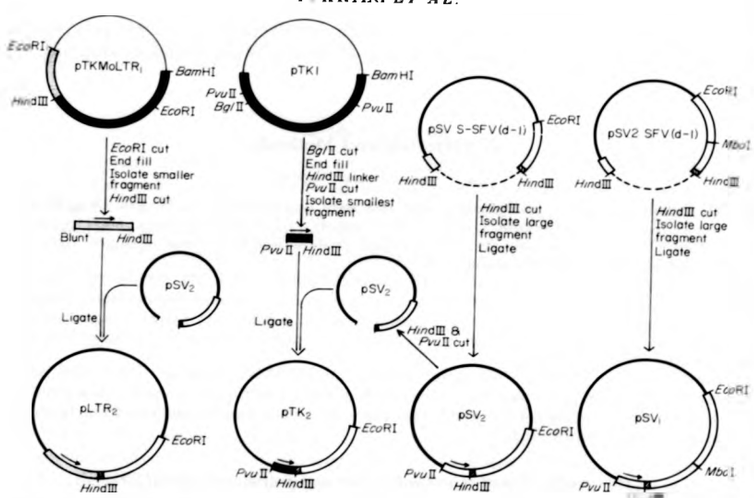


FIG. 1. Construction of expression vectors pSV_1, pSV_2, pTK_2 and $pLTR_2$. The plasmids pSV_2 and pSV_1 were derived from pSV S-SFV (d-1) and pSV_2 SFV (d-1), respectively, by excision of the Semliki Forest virus sequence (broken line) using the strategy indicated. These plasmids resemble the parental pSV_2 vector as described by Mulligan & Berg (1980) in that they contain 2300 bases of pBR sequences (thick lines) specifying the origin for plasmid replication and the ampicillinase gene. They also both contain SV40 sequences (open boxes) including a *Pvu*II-*Hin*dIII fragment containing the SV40 early promoter region. Downstream from the unique *Hin*dIII site, relative to this promoter, both vectors have a fragment carrying the early transcripts termination and polyadenylation signals. In pSV_2, this region lies almost adjacent to the *Hin*dIII site but in pSV_1 there is, between the *Hin*dIII site and the polyadenylation region, an additional *Mbo*I fragment containing the SV40 small t intron. The hatched box indicates a 21 base synthetic oligonucleotide that contains translational stop codons in all 3 reading frames (see Garoff *et al.*, 1983). The vectors pTK_2 and $pLTR_2$ were derived from pSV_2 after excision from pSV_2 of the SV40 early promoter and its replacement with either a region of the *Herpes simplex* thymidine kinase gene (shaded box), or the Moloney murine sarcoma virus long terminal repeat (LTR, stippled box); manipulations were performed as indicated in the Figure. For clarity, the various regions of the vectors are not drawn to scale. The position and polarity of the various promoters are shown by the small arrows.

(e) *Microinjection of DNA and RNA into* Xenopus *oocytes*

The microinjection techniques and oocyte culture conditions were as described by Colman (1984). An important departure from the method used for previous DNA injections (e.g. see Wickens & Gurdon, 1983) is the 45 min post-injection recovery period at 0°C, which we find increases oocyte survival. Plasmid DNAs were injected in 88 mM-NaCl, 10 mM-HEPES (pH 7·6), at 150 μg/ml, whereas mRNAs were injected in distilled water at 0·25 to 1·0 mg/ml. Injection volumes were approx. 40 nl/oocyte. For protein labelling, injected oocytes were always cultured for 24 h before removal of unhealthy oocytes, then incubated in modified Barths' saline containing 1 mCi [^{35}S]methionine/ml, in microtitre wells with 30 μl of medium/5 oocytes (Colman & Morser, 1979; Colman, 1984). After 24 h incubation in radioactive media, oocytes and media were collected to await processing. Similar procedures were adopted where transcript analysis on oocyte RNA was the experimental aim, except that no radioactive media were used.

(f) *Oocyte fractionation*

Labelled oocytes were usually homogenized as groups of 25 to 35 in 500 μl of T buffer (50 mM-NaCl, 10 mM-magnesium acetate, 20 mM-Tris·HCl, pH 7·6) supplemented with 10% (w/v) sucrose, 100 mM-NaCl and 1 mM-phenylmethylsulphonyl fluoride at 4°C. Homogenates were layered onto 1 ml of T buffer containing 20% (w/v) sucrose, in 5-ml polycarbonate tubes (M.S.E., Crawley, U.K.) and spun in an 8 × 5 ml rotor at 17,000 g_{max} for 30 min at 4°C. The supernatants representing the oocyte cytosol (C) were retained for immunoprecipitation. The pellets containing yolk and oocyte membranes were further extracted with 500 μl of resuspension buffer (100 mM-Tris·HCl (pH 7·6), 5 mM-magnesium acetate, 1% (v/v) Triton X-100 and 1 mM-PMSF) followed by centrifugation for 1 min at 10,000 g_{max} in an Eppendorf microcentrifuge at 4°C. The supernatant (M) containing solubilized membranes, was retained for immunoprecipitation.

In some cases, oocytes were homogenized directly in resuspension buffer and clarified by centrifugation as described above.

(g) *Immunoprecipitation and electrophoresis*

In early experiments, immunoprecipitations of the membrane (M), cytosol (C) or labelled incubation media (S) were performed exactly as described by Valle *et al.* (1983). In later experiments, all samples were first clarified by centrifugation in a Beckman airfuge at 1·4 kg/cm^2 for 5 min at 4°C. This step resulted in much cleaner immunoprecipitates. All immunoprecipitates were resuspended in sample buffer (200 mM-Tris·HCl (pH 8·8), 1 M-sucrose, 0·01% (w/v) bromophenol blue, 5 mM-EDTA, 3% (w/v) SDS, 8 mM-dithiothreitol), boiled for 3 min, alkylated at room temperature for 15 min in the presence of 70 mM-iodoacetamide and electrophoresed on 12·5% (w/v) polyacrylamide gels (Laemmli, 1970). Usually, each track contained the extract of 2·5 (M and C) or 5 (S) oocytes. Gels were fixed and then fluorographed (Bonner & Laskey, 1974).

(h) *Partial peptide mapping*

Selected bands were excised from the dried, fluorographed gels and subjected to partial cleavage with *N*-chlorosuccinimide (Sigma) as described by Lischwe & Ochs (1982). Each slice was washed in 25 ml of distilled water for 20 min with 1 change and then in 10 ml of urea/water/glacial acetic acid (1 g/1 ml/1 ml) for 20 min (1 change). Slices were then soaked in 5 ml of 15 mM-*N*-chlorosuccinimide dissolved in the urea solution described above, for 30 min, followed by two 10 min washes with 25 ml of water. After equilibration for 1·5 h in 3 × 10 ml of 10% (v/v) glycerol, 15% (v/v) mercaptoethanol, 3% (w/v) SDS, 62·5 mM-Tris·HCl (pH 6·8), slices were inserted into slots of an 18% (w/v) polyacrylamide gel. Electrophoresis proceeded at 12 mA overnight.

(i) *Enucleation of oocytes*

Injected oocytes cultured in non-radioactive media for 48 h were enucleated under culture medium by a shallow tangential incision with a 25-G needle at the animal (pigmented) pole, followed by gentle squeezing of the oocyte around the equatorial region (see Colman, 1984). The extruded nucleus was immediately homogenized in RNA extraction media (see below). The whole procedure took less than 30 s.

(j) *Extraction of RNA from whole or enucleated oocytes*

Oocytes or isolated nuclei and enucleated cytoplasms were homogenized in a proteinase K (Boehringer)-containing buffer and RNA was extracted as described by Kressmann *et al.* (1978). Transfer RNA (*E. coli*, Sigma) was added as carrier (final conen 100 μg/ml) before homogenization of the nuclei. Extracted RNA from oocytes or nuclei was resuspended in S_1 hybridization buffer (see section (1), below) at an approximate concentration of 4 mg/ml.

(k) *Translation* in vitro

Wheat germ extracts were prepared and assayed in the presence of 50 μg added mRNA/ml by the procedure of Roberts & Paterson (1973).

(l) *S_1 nuclease mapping*

In each hybridization assay, 4 to 8 μg of extracted RNA were mixed with approximately 200 ng of double-stranded DNA probe, in a total volume of 7 μl of S_1 hybridization buffer (80% (v/v) deionized formamide, 0·4 M-NaCl, 40 mM-PIPES, 1 mM-EDTA, pH 6·4). Each mixture was sealed in a glass capillary tube and the contents denatured by 10 min incubation at 80°C, followed by immediate transfer to 52°C. After hybridization for 15 h at 52°C, the contents of each capillary were diluted into 150 μl of S_1 assay buffer (0·28 M-NaCl, 4·5 mM-zinc acetate, 20 μg sonicated, denatured salmon sperm DNA/ml (Sigma), 0·05 M-sodium acetate, pH 4·6) containing 150 units of S_1 nuclease. The reactions were incubated at 37°C for 30 min before termination was effected by addition of 6 μl of 0·2 M-EDTA. Samples were then extracted with 150 μl of phenol/chloroform (1 : 1, v/v) followed by precipitation with ethanol. Precipitates were dissolved in 80% (v/v) deionized formamide, 0·3% (w/v) xylene cyanol FF, 0·3% (w/v) bromophenol blue, 20 mM-EDTA (pH 7·6) and electrophoresed on 6% or 8% (w/v) acrylamide/urea sequencing gels as described by Sanger *et al.* (1977).

Preparation of labelled DNA probes. The two 5′ termini of a *Hin*dIII-restricted pTK_2 vector (Fig. 1) were "phosphatased" with calf intestinal phosphatase and then "kinased" using [γ-^{32}P]ATP and T4 polynucleotide kinase. Probes 1 and 2 (see Fig. 6) were prepared *via* a gel purification step after restriction of the kinased fragment with *Bam*HI or *Pru*II, respectively.

A labelled 3′ terminus at the *Acc*I site at position 341 bp into the *Hin*dIII ovalbumin insert was prepared by restricting pTK_2OV+ with *Acc*I and isolating the 4000 base fragment. One end of this fragment was then end filled using Klenow polymerase and [α-^{32}P]dCTP alone. Under these conditions, the other end is not labelled. This probe (probe 3) is shown schematically in Fig. 8.

(m) *Microinjection and staining of cultured cells*

Microinjection of DNAs or mRNA (both at 1 mg/ml in distilled water) into cultured baby hamster kidney (BHK-21) cells was performed as described by Kondor-Koch *et al.* (1982) or Graessmann *et al.* (1980). Injected cells were cultured at 37°C for 8 h before fixing in 3% (v/v) paraformaldehyde and stained for indirect immunofluorescence by the method of Ash *et al.* (1977), using FITC-conjugated goat antirabbit antisera as second antibody. All antibodies were used at a concentration of 10 μg/ml. Stained cells were examined with a Zeiss microscope using epifluorescence. Kodak Tri-X film was used for photography.

Several of the injections were performed by Beate Timm (EMBL, Heidelberg), whose assistance is gratefully acknowledged.

3. Results

(a) *Wild type ovalbumin expression from viral promoters in oocytes*

Oocytes injected with cloned genomal ovalbumin DNA synthesize ovalbumin but the synthesized ovalbumin represented only 0·01% of the total newly synthesized oocyte protein (Wickens *et al.*, 1980). No ovalbumin was detected after injection of pOV230, a recombinant plasmid containing a full-length ovalbumin cDNA (Wickens *et al.*, 1980), although the reasons for this lack of protein production were not clear. In Figure 2, we describe the excision of the

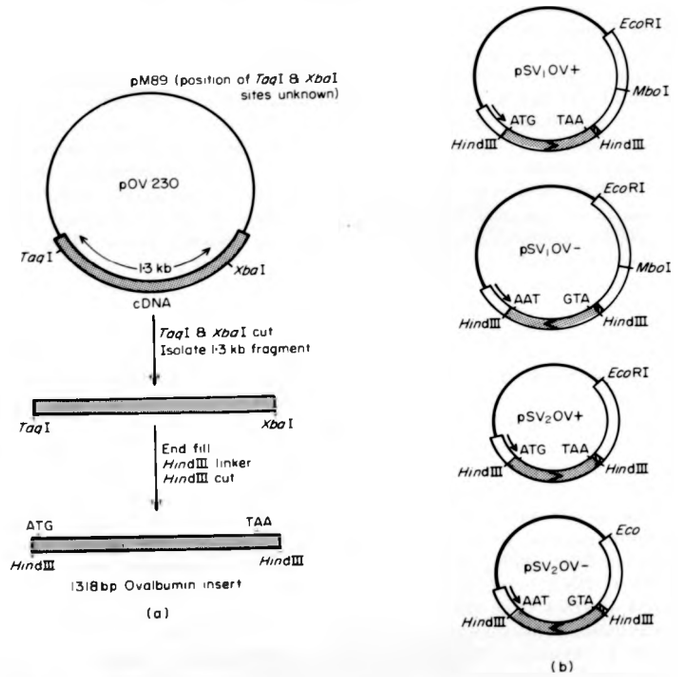


FIG. 2. Construction of ovalbumin cDNA fragment and insertion into the expression vectors pSV_1 and pSV_2. (a) A 1300 base *Taq*I-*Xba*I fragment encompassing the coding region of the chick ovalbumin cDNA was excised from pOv230. After the overhanging ends were filled in, *Hin*dIII linkers were added to this fragment, which then enabled *Hin*dIII ends to be generated for the insertion of the ovalbumin cDNA into the *Hin*dIII site of the expression vectors described in Fig. 1. (b) The constructs obtained when this *Hin*dIII ovalbumin fragment was ligated into the *Hin*dIII-cut expression vectors pSV_1 and pSV_2. The vector components are as described in Fig. 1.

ovalbumin cDNA from pOV230 and its manipulation for insertion into the pSV_1, pSV_2 and pTK_2 vectors (see Fig. 1). The insertion of the ovalbumin cDNA into both vectors occurred in each of two orientations, generating six constructs, pSV_1OV+, pSV_1OV-, pSV_2OV+ and pSV_2OV- (Fig. 2(b)), and pTK_2OV+ and pTK_2OV- (not shown), the plus and minus signs denoting the orientation of the insert relative to the viral promoter (see Fig. 2(b)). All these constructs were injected into *Xenopus* oocyte nuclei, and the oocytes were cultured in the presence of [^{35}S]methionine as described in Materials and Methods. After collection of the incubation media, oocytes were homogenized and fractionated into a membrane

fraction containing microsomes and other vesicles, and a soluble, cytosolic fraction. Samples were then immunoprecipitated and submitted to electrophoresis (Fig. 3). It is clear that injection of both pSV_1OV+ and pSV_2OV+ results in the appearance of two immunospecific polypeptides in the membrane fraction. Polypeptides of similar mobility are also found in the media. These proteins are similar to those found after mRNA injection (Fig. 3(b); and Colman *et al.*, 1981). Previous work has shown the two bands to be glycosylated derivatives of ovalbumin (Colman *et al.*, 1981). These two bands are also present in the cytosol at a low level and represent leakage from the membraneous vesicles during fractionation (Colman *et al.*, 1981). A faint band of greater mobility than the glycosylated proteins is also present in the cytosol track. The band is unglycosylated ovalbumin (Colman *et al.*, 1981), and its presence in the cytosol is presumably caused by a failure in translocation of some newly synthesized ovalbumin into the endoplasmic reticulum.

The results in Figure 3(a) also show that pSV_2OV+ elicits the production of more protein than pSV_1OV+. This result has been obtained in four similar experiments. Since the only difference between the two constructs is in the provision of an intron and its flanking sequences in the 3′ untranslated region of any transcripts initiating at the SV40 early promoter (see Figs 1 and 2(b)), we conclude that the presence of this intron and/or flanking regions is deleterious to expression of ovalbumin in oocytes.

The highest level of ovalbumin expression obtained with pSV_2OV+ over four experiments was 0·07%, measured as the incorporation of [^{35}S]methionine into ovalbumin relative to its incorporation into total protein. When ovalbumin expression from pTK_2OV+ DNA was directly compared with that from pSV_2OV+ DNA, over sevenfold more protein was reproducibly found using pTK_2OV+ DNA (Fig. 3(b)). In recent experiments, as much as 0·42% of the [^{35}S]methionine incorporation into protein was found in ovalbumin after pTK_2OV+ injection. This represents a 42-fold increase in the synthesis of ovalbumin as compared to that found after injection of ovalbumin genomal DNA (Wickens *et al.*, 1980).

(b) *Variant ovalbumin expression in oocytes*

On analysis of the oocytes injected with pSV_2OV- (Figs 3(a) and 4(a)), we were surprised to note both the appearance of immunospecific bands and their abundance in comparison to products from pSV_2OV+ injected oocytes. No such bands were found after pSV_1OV- injection (Fig. 3(a)). Several bands are evident in both cytosol and membrane fractions of pSV_2OV- injected oocytes. For the clarity of the following discussion, we have labelled some of these bands a to f (Fig. 4(a)). The membrane-associated bands (Fig. 4(a), bands a and b) have different mobilities from the cytosol forms (Fig. 4(a), bands c and d), and only those bands with similar mobility to bands a and b appear in the medium. The apparent molecular weight of the most abundant membrane form synthesized after pSV_2OV- DNA injection (band a) is about 3000 greater than the glycosylated, wild type form found in mRNA (band f) or pSV_2OV+ (band e)

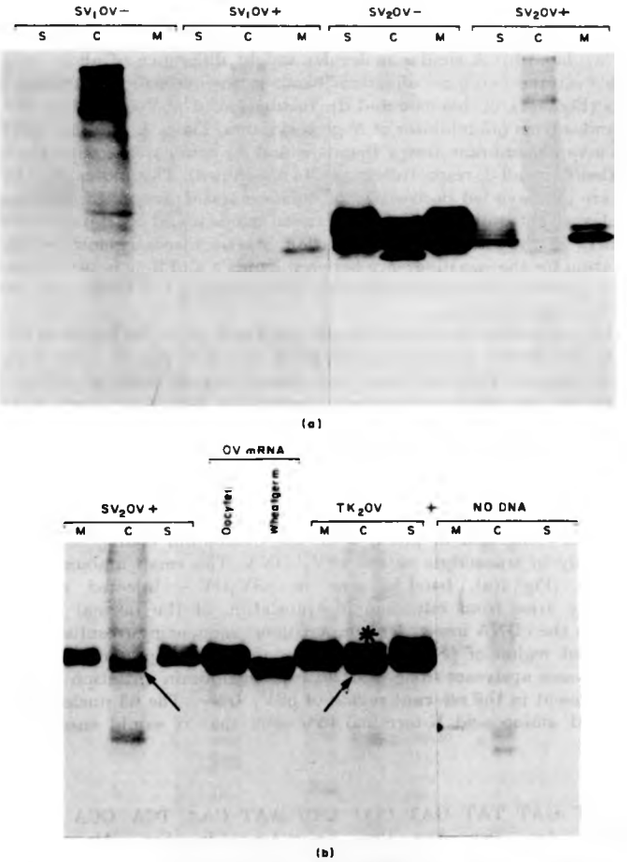


FIG. 3. Synthesis of wild type ovalbumin after DNA injections. Oocytes were injected with the indicated DNAs and incubated in [^{35}S]methionine as indicated in Materials and Methods. Oocytes were then fractionated into membrane (M) and cytosol (C) components. Incubation media (S) and fractions M and C were then immunoprecipitated with anti-ovalbumin antibody and electrophoresed as described in Materials and Methods. Oviduct mRNA was also translated by either injection into oocytes or in a wheat germ cell free system. mRNA-injected oocytes were homogenized directly in resuspension buffer (see Materials and Methods) before immunoprecipitation, whereas samples from the wheat germ assay were diluted 10-fold into resuspension buffer before immunoprecipitation. The arrows point to the "miscompartmented" normal length ovalbumin (see the text and Fig. 4(b)), whereas the asterisk indicates a novel extended ovalbumin. (a) and (b) are separate experiments with different batches of oocytes.

fraction containing microsomes and other vesicles, and a soluble, cytosolic fraction. Samples were then immunoprecipitated and submitted to electrophoresis (Fig. 3). It is clear that injection of both pSV_1OV+ and pSV_2OV+ results in the appearance of two immunospecific polypeptides in the membrane fraction. Polypeptides of similar mobility are also found in the media. These proteins are similar to those found after mRNA injection (Fig. 3(b); and Colman *et al.*, 1981). Previous work has shown the two bands to be glycosylated derivatives of ovalbumin (Colman *et al.*, 1981). These two bands are also present in the cytosol at a low level and represent leakage from the membraneous vesicles during fractionation (Colman *et al.*, 1981). A faint band of greater mobility than the glycosylated proteins is also present in the cytosol track. The band is unglycosylated ovalbumin (Colman *et al.*, 1981), and its presence in the cytosol is presumably caused by a failure in translocation of some newly synthesized ovalbumin into the endoplasmic reticulum.

The results in Figure 3(a) also show that pSV_2OV+ elicits the production of more protein than pSV_1OV+. This result has been obtained in four similar experiments. Since the only difference between the two constructs is in the provision of an intron and its flanking sequences in the 3′ untranslated region of any transcripts initiating at the SV40 early promoter (see Figs 1 and 2(b)), we conclude that the presence of this intron and/or flanking regions is deleterious to expression of ovalbumin in oocytes.

The highest level of ovalbumin expression obtained with pSV_2OV+ over four experiments was 0·07%, measured as the incorporation of [^{35}S]methionine into ovalbumin relative to its incorporation into total protein. When ovalbumin expression from pTK_2OV+ DNA was directly compared with that from pSV_2OV+ DNA, over sevenfold more protein was reproducibly found using pTK_2OV+ DNA (Fig. 3(b)). In recent experiments, as much as 0·42% of the [^{35}S]methionine incorporation into protein was found in ovalbumin after pTK_2OV+ injection. This represents a 42-fold increase in the synthesis of ovalbumin as compared to that found after injection of ovalbumin genomal DNA (Wickens *et al.*, 1980).

(b) *Variant ovalbumin expression in oocytes*

On analysis of the oocytes injected with pSV_2OV- (Figs 3(a) and 4(a)), we were surprised to note both the appearance of immunospecific bands and their abundance in comparison to products from pSV_2OV+ injected oocytes. No such bands were found after pSV_1OV- injection (Fig. 3(a)). Several bands are evident in both cytosol and membrane fractions of pSV_2OV- injected oocytes. For the clarity of the following discussion, we have labelled some of these bands a to f (Fig. 4(a)). The membrane-associated bands (Fig. 4(a), bands a and b) have different mobilities from the cytosol forms (Fig. 4(a), bands c and d), and only those bands with similar mobility to bands a and b appear in the medium. The apparent molecular weight of the most abundant membrane form synthesized after pSV_2OV- DNA injection (band a) is about 3000 greater than the glycosylated, wild type form found in mRNA (band f) or pSV_2OV+ (band e)

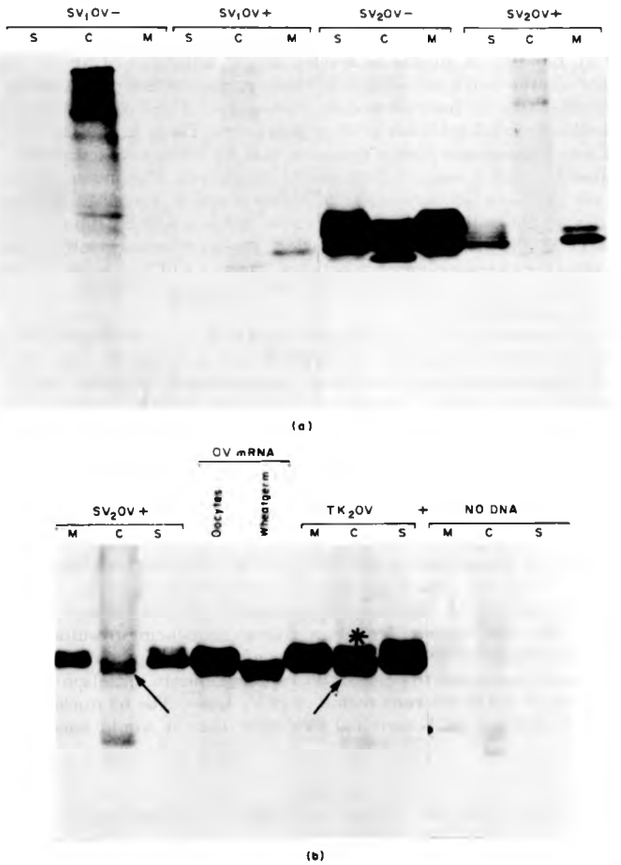


FIG. 3. Synthesis of wild type ovalbumin after DNA injections. Oocytes were injected with the indicated DNAs and incubated in [^{35}S]methionine as indicated in Materials and Methods. Oocytes were then fractionated into membrane (M) and cytosol (C) components. Incubation media (S) and fractions M and C were then immunoprecipitated with anti-ovalbumin antibody and electrophoresed as described in Materials and Methods. Oviduct mRNA was also translated by either injection into oocytes or in a wheat germ cell free system. mRNA-injected oocytes were homogenized directly in resuspension buffer (see Materials and Methods) before immunoprecipitation, whereas samples from the wheat germ assay were diluted 10-fold into resuspension buffer before immunoprecipitation. The arrows point to the "miscompartmented" normal length ovalbumin (see the text and Fig. 4(b)), whereas the asterisk indicates a novel extended ovalbumin. (a) and (b) are separate experiments with different batches of oocytes.

injected oocyte membranes. However, it is notable that this latter "wild type" ovalbumin band (Fig. 4(a), band e) is also present in the pSV_2OV- oocytes (Fig. 4(a), band b). A similar molecular weight difference of about 3000 is also seen between the two most abundant bands in the cytosol of pSV_2OV- injected oocytes (Fig. 4(a), cf. bands c and d). Incubation of pSV_2OV- injected oocytes with tunicamycin (an inhibitor of *N*-glycosylation; Tkacz & Lampen, 1975) results in the major membrane forms (bands a and b) comigrating with the cytosolic forms (bands c and d, respectively; results not shown). This indicates that bands a and b are glycosylated derivatives of bands c and d, respectively, but that the molecular weight difference of 3000 between bands a and b cannot be ascribed to differences in the extent of glycosylation. As we discuss below, a more likely explanation for the size difference between bands a and b or between bands c and d is the presence of an approximately 20 amino acid extension on the higher molecular weight proteins.

The only difference in the vectors pSV_1 and pSV_2 is in the region of SV40 DNA between the *Eco*RI site and the *Hin*dIII site (see Fig. 1). The vector pSV_1 contains, between the translation stop oligonucleotide adjacent to the *Hin*dIII site and the transcription terminator region, an additional piece of the SV40 genome encoding the small t intron. Thus in the constructs pSV_1OV- and pSV_2OV- (Fig. 2(b)), the SV40 sequences immediately upstream from the initiation ATG of the inserted ovalbumin gene are different. We therefore hypothesize that the large ovalbumin species, seen in pSV_2OV- injected oocytes, results from initiation of translation at an AUG (encoded by vector sequence) that is upstream from, and in phase with, the ovalbumin coding sequence, and is present only in transcripts of the pSV_2 DNA. The small amount of wild type ovalbumin (Fig. 4(a), band b) seen in pSV_2OV- injected oocytes would presumably arise from initiation of translation of the normal initiator AUG encoded in the cDNA insert. Maxam & Gilbert sequencing (results not shown) of the relevant region of pSV_2OV- DNA confirmed the presence of an in phase ATG, 63 bases upstream from the wild type ovalbumin initiation ATG. No such ATG is present in the relevant region of pSV_1 DNA. The 63 nucleotide sequence and the 21 amino acid N-terminal extension that it would encode are shown below.

−21										
ATG	GCT	GAT	TAT	GAT	CCG	GTC	AAT	CAA	TCA	GCA
Met	Ala	Asp	Tyr	Asp	Pro	Val	Asn	Gln	Ser	Ala
										+1
AGC	TTG	CCG	AAA	GAC	AAC	TCA	GAG	TTC	ACC	ATG
Ser	Leu	Pro	Lys	Asp	Asn	Ser	Glu	Phe	Thr	Met

The pattern of ovalbumin bands seen after pTK_2OV- injection into oocytes is identical to that produced by pSV_2OV- injection (Fig. 4(a)). Since the sequences upstream from the wild type ovalbumin initiator ATG are the same in these two vectors, a similar argument to that above could account for the appearance of

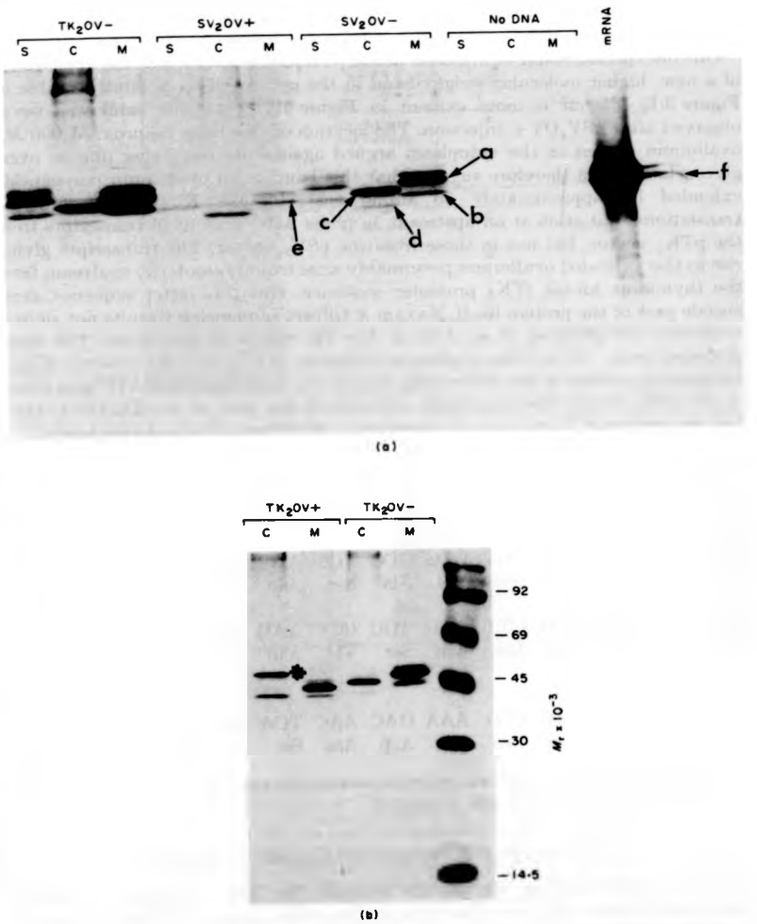


FIG. 4. Synthesis of variant ovalbumins after DNA injection. Oocytes were injected with the indicated DNAs and incubated in [^{35}S]methionine as indicated in Materials and Methods. Oocytes were then processed as described in Fig. 3. Oviduct mRNA injected oocytes were also processed as described in Fig. 3. The experiments shown in (a) and (b) were conducted with different batches of oocytes.

extended ovalbumins after pTK_2OV- injection. Corroborative evidence for the nature of this extension is provided in section (c), below.

One unexpected result of injection of the pTK_2OV+ vector was the appearance of a new, higher molecular weight band in the cytosol. This is faintly visible in Figure 3(b) (*) but is more evident in Figure 4(b) (*). This band was never observed after pSV_2OV+ injection. The location of this large (approx. 51,000 M_r) ovalbumin species in the cytoplasm argued against its size being due to over-glycosylation. We therefore suggest that this band is an ovalbumin polypeptide extended by approximately 50 amino acids at the N terminus, due to translational initiation at an upstream, in phase AUG present in transcripts from the pTK_2 vector, but not in those from the pSV_2 vector. The transcripts giving rise to this extended ovalbumin presumably arise from promoter(s) upstream from the thymidine kinase (TK) promoter sequence, and this latter sequence must encode part of the protein itself. Maxam & Gilbert sequencing (results not shown) confirmed the presence of an ATG in the TK region of the vector, 153 bases upstream from the normal ovalbumin initiation ATG, and the absence of any termination codons in the intervening region. No such upstream ATG was found in the pSV_2 vector. The nucleotide sequence of this part of the TK_2OV+ DNA and the predicted 51 amino acid extension it would encode are shown below.

−51												
ATG	CAG	TCG	GGG	CGG	CGC	GGT	CCG	AGG	TCC	ACT	TCG	CAT
Met	Gln	Ser	Gly	Arg	Arg	Gly	Pro	Arg	Ser	Thr	Ser	His
ATT	AAG	GTG	ACG	CGT	GTG	GCC	TCG	AAC	ACC	GAG	CGA	CCC
Ile	Lys	Val	Thr	Arg	Val	Ala	Ser	Asn	Thr	Glu	Arg	Pro
TGC	AGC	GAC	CCG	CTT	AAC	AGC	GTC	AAC	AGC	GTG	CCG	CAG
Cys	Ser	Asp	Pro	Leu	Asn	Ser	Val	Asn	Ser	Val	Pro	Gln
												+1
ATC	GCA	AGC	TTG	CCG	AAA	GAC	AAC	TCA	GAG	TTC	ACC	ATG
Ile	Ala	Ser	Leu	Pro	Lys	Asp	Asn	Ser	Glu	Phe	Thr	Met

Further evidence corroborating the nature of this and the previously discussed 21 amino acid extension is now discussed.

(c) *Partial peptide analysis of ovalbumin polypeptides*

In the above sections we provide evidence for the synthesis of extended ovalbuumin; however, the evidence for the N-terminal location of these extensions is indirect. More direct evidence is provided by partial peptide mapping of the various electrophoresed gel bands using *N*-chlorosuccinimide. This reagent specifically cleaves polypeptide chains at tryptophan residues (Lischwe & Ochs, 1982), and the predicted extension sequence contains no tryptophan residues. Since ovalbumin contains three tryptophan residues, partial cleavage should generate nine fragments (see Fig. 5(a)); however, some would be too small to

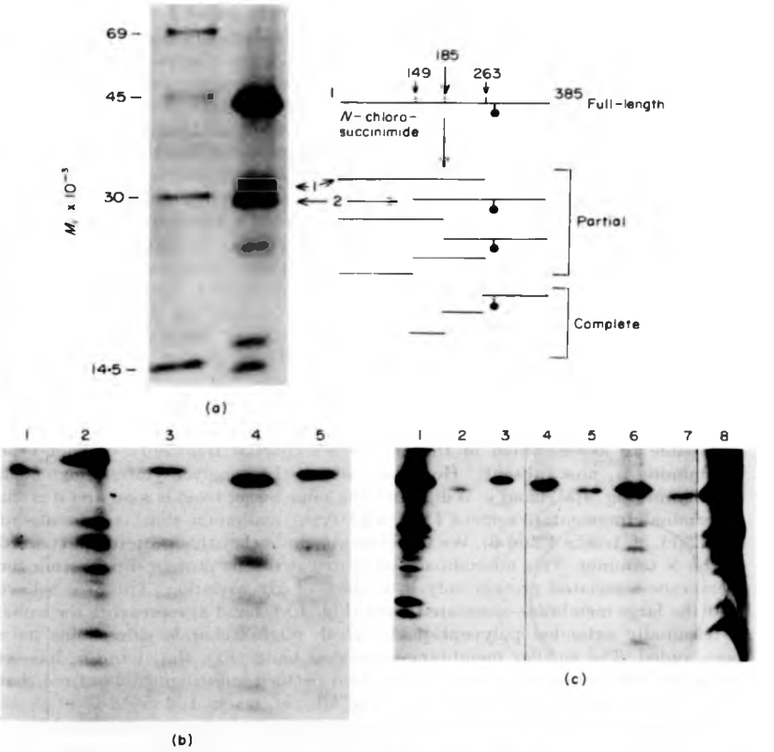


FIG. 5. Partial peptide mapping of immunoprecipitated ovalbumins. (a) Oviduct mRNA was translated in the wheat germ cell free system and the sample immunoprecipitated with anti-ovalbumin antibody before electrophoresis on a 12·5% polyacrylamide gel. The major product corresponding to full-length ovalbumin was excised from the dried fluorogram and processed with *N*-chlorosuccinimide as described in Materials and Methods. Processed samples were then electrophoresed on an 18% polyacrylamide gel. The cleavage positions (i.e. locations of tryptophan residues) in full-length ovalbumin are indicated by arrows and the possible complete and partial digestion products are shown. The symbol (♦) indicates the position (Asn293) of the oligosaccharide side-chain in glycosylated ovalbumin. The wheat germ product will not have this side-chain. This fluorograph was exposed for 3 days. (b) Tracks 1 to 5 display the results of partial digestion of various electrophoresed ovalbumins. Track 1, Fig. 4(a), band f; track 2, Fig. 4(a), band a; track 3, Fig. 4(a), band e; track 4, unglycosylated ovalbumin synthesized in wheat germ assay as in (a); track 5, Fig. 4(a), band b. This fluorograph was exposed for 6 weeks. (c) Tracks 1 to 8 display the results of partial digestion of various electrophoresed ovalbumins. Track 1, Fig. 4(a), duplicate of band f; track 2, Fig. 4(a), duplicate of band b; track 3, Fig. 4(b), asterisked band; track 4, Fig. 4(a), duplicate of band a; track 5, Fig. 4(a), duplicate of band e; track 6, Fig. 4(a), band c; track 7, Fig. 4(a), band d; track 8, unglycosylated ovalbumin synthesized in wheat germ, as in (a). In track 3 the extended digestion product is indicated by the arrowhead. This fluorograph was exposed for 8 weeks.

resolve on the polyacrylamide gels used. Figure 5(a) shows a typical digestion pattern resulting from the partial cleavage of the unmodified ovalbumin produced after mRNA translation in the wheat germ translation system. The two bands labelled 1 and 2 correspond to the two largest partial fragments and, on the basis of tryptophan distribution in ovalbumin, the larger fragment (fragment 1) is an N-terminal fragment, whilst fragment 2 derives from the C terminus. When the major glycosylated proteins produced in oocytes after mRNA (Fig. 4(a), band f) or pSV_2OV+ (Fig. 4(a), band e) injection are digested, the digestion patterns are similar (Fig. 5(b), cf. tracks 1 and 3), though as expected, different from the pattern from the unglycosylated polypeptide (Fig. 5(b), track 4). Unfortunately, as a consequence of the oligosaccharide side-chain at asparagine 293 in ovalbumin, the mobility of fragment 2 (and other C-terminal fragments) is reduced and this fragment co-migrates with fragment 1. This glycosylation complicates analysis of the extended ovalbumin which, on digestion, should produce a new, higher molecular weight fragment at the expense of either fragment 1 (for an N-terminal extension) or fragment 2 (for a C-terminal extension). A larger band (Fig. 5(b), track 2) is indeed found when the glycosylated, extended ovalbumin from pSV_2OV- or pTK_2OV- injected oocytes (Fig. 4(a), band a) is digested but, for the reasons given above, it is impossible to assess which of the two largest partial fragments of wild type ovalbumin is now absent. However, when the unglycosylated, extended ovalbumin (Fig. 4(a), band c) is digested, the same larger band is seen and it is the N-terminal fragment (fragment 1) of wild type ovalbumin that is now absent (Fig. 5(c), cf. tracks 4 and 6). We therefore conclude that this protein is extended at the N terminus. This miscompartmentalized cytosolic protein differs from the membrane-associated protein only in its lack of glycosylation. Thus we believe that the large membrane-associated band (Fig. 4(a), band a) represents the same, N-terminally extended polypeptide, to which oligosaccharide side-chains have been added. The smaller membrane-associated band (Fig. 4(a), band b) has, as predicted (see section (a), above), a digestion pattern indistinguishable from that of full-length, glycosylated ovalbumin (Fig. 5(b), cf. tracks 1, 3 and 5).

This logic also applies to the digestion analysis of the large cytosolic band found after pTK_2OV+ injection (discussed above). The predicted 51 amino acid extension lacks any tryptophan residues so, on digestion, fragment 1 should be replaced by a much larger band if the extension is N-terminal. This larger band can be seen in Figure 5(c), track 3.

(d) *Analysis of transcripts from injected pTK_2OV+ and pTK_2OV- DNAs in oocytes*

(i) *Initiation at the TK promoter*

McKnight & Kingsbury (1982) have demonstrated, using both S_1 mapping and primer extension methods, that the thymidine kinase promoter is efficiently and accurately transcribed in *Xenopus* oocytes. Figure 6 confirms that the TK promoter in our expression vectors is accurately transcribed, although we have not measured the efficiency. Using probe 1, described in the legend to Figure 6, we

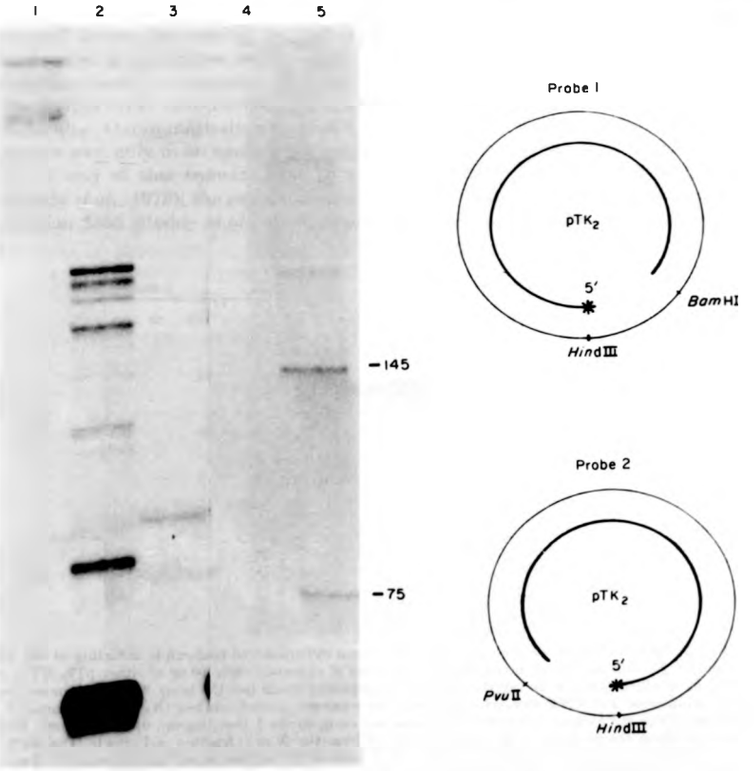


FIG. 6. Predominant transcripts from the TK and cryptic SV40 promoters. Oocytes were injected with either pTK_2OV+ or pTK_2OV- DNA and cultured for 48 h before RNA extraction and S_1 analysis. The labelled DNA probes used for the analysis are described in Materials and Methods and shown above. Track 1, unhybridized probe; track 2, probe 1 × pTK_2OV+ RNA; track 3, probe 2 × pTK_2OV- RNA; track 4, probe 1 × uninjected oocyte RNA; track 5, markers, end-labelled *Hin*fI digest of pAT153.

would anticipate protection of a 63 base fragment if initiation is occurring at the correct position. In fact, the major protected species was 61 bases long (see also Fig. 7); however, this reduction in size could be due to "nibbling" by the S_1 nuclease. As can be seen in Figure 7, most of these transcripts are present in the oocyte cytoplasm under the steady-state conditions used in this study.

Several higher molecular weight, S_1-protected bands are also visible in Figure 6. Presumably, the transcript that gives rise to the 51 amino acid-extended

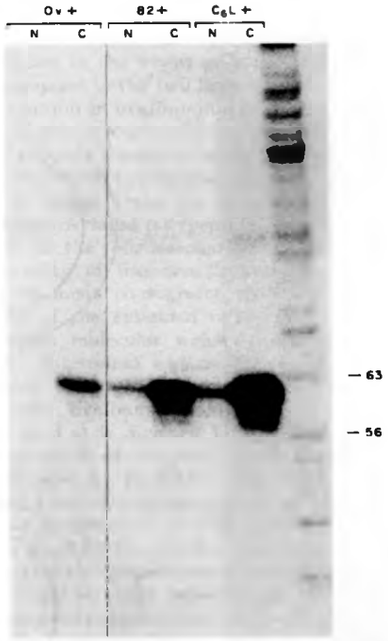


FIG. 7. Distribution between the oocyte nucleus and cytoplasm of transcripts initiating at the TK promoter. Several oocytes received nuclear injections of approximately 10 ng of either pTK_2OV+ or pTK_282+ which contains an insert encoding for preprochymosin (see the text). 48 h later the oocytes were enucleated and RNA was extracted from the separate, pooled, nuclear (N) and cytoplasmic (C) fractions. The RNAs were subjected to S_1 analysis using probe 1 (see diagram 9) as described. Each track contains the products of S_1 analysis of RNA from the N or C fraction of 1 oocyte. The marker track contains a dideoxy A-sequencing track of histone DNA calibrated and provided by P. Turner (University of Warwick). A third DNA shown above, pTK_2C_6L+, contains the N-terminal 24 amino acids of preprochymosin and the C terminal 122 amino acids of lysozyme. Expression from this is included only to preserve gel clarity.

ovalbumin on injection of pTK_2OV+ (section (c), above) is responsible for one of these bands since, in this case, the TK promoter region of the DNA encodes part of the amino acid sequence of the extended peptide. Therefore, the transcript must originate from another promoter upstream from the TK promoter.

(ii) *Relative translational efficiencies of pTK_2OV+ transcripts and ovalbumin mRNA*

Ovalbumin transcripts from the injected DNAs will differ from ovalbumin mRNA molecules in two ways: firstly, transcripts from DNA will have a 5′ viral

leader sequence and secondly, some 5′ ovalbumin mRNA leader sequence has been excised during the cloning. In the experiment illustrated by Figure 8 we have attempted to quantitate steady-state levels of full-length transcripts after DNA injection and to compare the amount of ovalbumin produced from such transcripts with that formed after injection of equivalent numbers of mRNA molecules. Our quantitation is based on the amount of a 1150 base protected species seen only in an exclusively cytoplasmic region (vegetal half) of the oocyte; the 3′ end of this transcript maps to position 2540±20 of the SV40 genome (Reddy *et al.*, 1978); the natural 3′ end of mature SV40 early transcripts occurs at position 2536 (Reddy *et al.*, 1978). Our results indicate that 48 hours after DNA

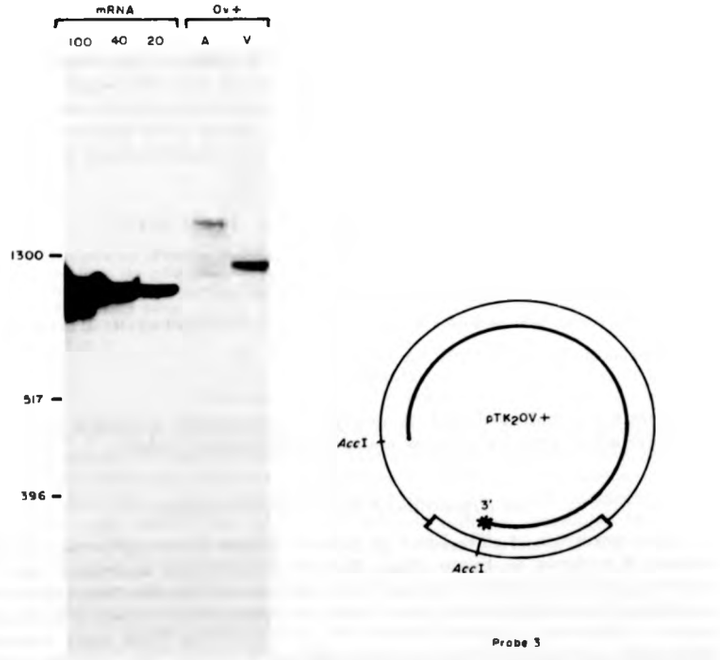


FIG. 8. Quantitation of 3′ ends of ovalbumin-coding transcripts. Several oocytes each received a nuclear injection of 10 ng of pTK_2OV+ DNA. 48 h later the oocytes were rapidly frozen at −70°C and bisected along the animal–vegetal boundary with a razor blade. RNA was extracted from the pooled animal pole (A) and vegetal pole (V) fractions, and subjected to S_1 analysis using probe 3. Each track contains the products of analysis of 2 half oocytes-worth of RNA. The probe was also hybridized to 100 ng, 40 ng and 20 ng samples of hen oviduct poly(A)$^+$ RNA.

injection, the number of ovalbumin transcripts per oocyte was equivalent to the ovalbumin mRNA content from approximately 5 ng of polyadenylated oviduct mRNA. However, when comparisons were made of the amount of ovalbumin produced after injection of DNA or 5 ng of mRNA, it was found that over 25-fold more protein was present after mRNA injection. We conclude that the injected mRNA is translated more efficiently than the transcripts produced in the oocyte.

(iii) *Mapping of a new SV40 "promoter"*

In sections (a), (b) and (c), above, we described the synthesis of protein after injection of constructs where the TK promoter and protein coding sequences are of opposing polarities. To account for these proteins, transcripts of opposite polarity to TK transcripts must be synthesized. Figure 6 (track 3) shows the results of an S_1 mapping procedure designed to map the 5′ ends of such transcripts. Only one major site of initiation of transcription in the opposite orientation to that of transcription from the TK promoter was identified at a position approximately 92 bases 5′ to the *Hin*dIII site. The sequence of this region is known (Reddy *et al.*, 1978), and the mapped initiation site corresponds to position 2620 of the SV40 genome. As shown below (with sequence hyphens omitted for clarity), various promoter-like sequences (underlined) are present 5′ to the initiation site:

```
5′ CCATTATAAG CTGCAATAAA CAAGTTAACA ACAACAATTG
                        Approximate transcription start
                                  ↓
   CATTCATTTT ATGTTTCAGG TTCAGGGGGA GGTGTGGGAG
   GTTTTTTAAA GCAAGTAAAA CCTCTACAAA TGTGGTATGG CTGATTATGA
   TCCGGTCAAT CAATCAGCAA GCTTGCCG . . 3′
                        HindIII
```

Figure 6 also indicates that at steady-state, transcripts beginning at the TK promoter are in great excess over transcripts of opposite polarity.

(e) *Expression of chick lysozyme in oocytes*

The construction of pTK_2Lys+ is shown in Figure 9 and expression from this construct is shown in Figure 10(a). Most of the lysozyme is secreted into the medium, a result consistent with previous data from mRNA injection experiments (Colman *et al.*, 1981; Cutler *et al.*, 1981). In Figure 10(b) we have also assayed levels of lysozyme produced under the control of the SV40 early promoter (pSV_2Lys+) or a retroviral promoter derived from the long terminal repeat (LTR) region of a murine sarcoma virus ($pLTR_2Lys+$, see Fig. 1). The LTR promoter is known to be highly active in mammalian cells (reviewed by Weiss *et al.*, 1984). However, in oocytes it is evident that, whilst use of the TK promoter leads to better expression of lysozyme than with the SV40 promoter, no further boost to expression is achieved with the LTR promoter.

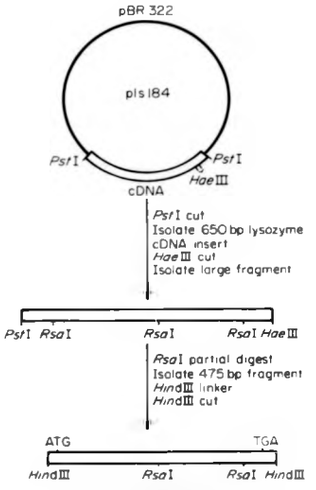


FIG. 9. Manipulation of lysozyme cDNA for insertion into expression vectors. Chick lysozyme cDNA was obtained from the pBR322 clone pls184, by excision of the *Pst*I insert carrying the lysozyme sequence. To remove the ends of the insert bearing G-C tails, a 475 base-pair *Rsa*I-*Hae*III fragment was isolated as shown using a partial *Rsa*I digest. This fragment, which contained the entire coding region, was *Hin*dIII linkered and then cut with *Hin*dIII for insertion into the expression vectors described in Fig. 1.

(f) *Expression of calf prochymosin*

Injection of calf preprochymosin mRNA into oocytes elicits the production of calf prochymosin, which is secreted (Fig. 11(a)). The prochymosin can be partially cleaved into chymosin (an aspartyl proteinase) by acidification of oocyte or media extracts. Normally, prochymosin is secreted into the fourth stomach of the unweaned calf, where the acid environment induces cleavage. We have constructed various recombinants containing cloned chymosin DNA (Fig. 11). pTK_282+ contains the complete preprochymosin sequence, pTK_270+ contains the prochymosin sequence with an added initiator methionine residue, whilst pTK_286+ contains the chymosin sequence, with an added initiator methionine residue.

We would expect expression of pTK_282+ to lead to a membrane-associated prochymosin that is secreted, whereas expression of pTK_270+ should lead to a cytosol-located prochymosin, since the N-terminal signal sequence present in the preprochymosin DNA of pTK_282+ is absent from pTK_270+. The results of such experiments, shown in Figure 11, confirm these predictions; however, only in the experiments where expression from pTK_282+ was very high was expression from

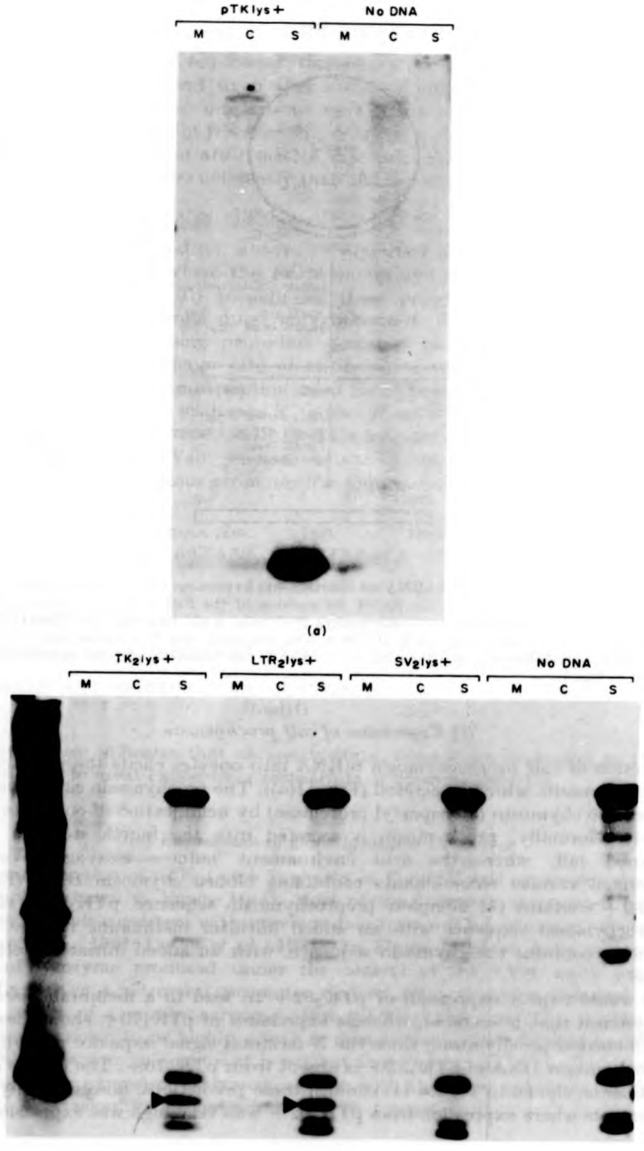


Fig. 10

pTK_270+ detected. Since the low expression of pTK_270+ might be attributable to a poor stability of prochymosin in the cytosolic compartment, we attempted to "rescue" this miscompartmentalized protein by injection of rabbit anti-prochymosin into the oocytes. However, this procedure resulted in no prochymosin being detected (Fig. 11(c)). No expression from pTK_286+ was ever detected. In the best experiments, expression from pTK_282+ represented 0·57% of the methionine incorporation into total oocyte protein. At the level of transcription, reproducibly greater amounts of transcript were found after pTK_282+ injection than after pTK_2OV+ injection (Fig. 7).

(g) *Expression of ovalbumin, lysozyme and prochymosin in cultured cells*

Cultured baby hamster kidney (BHK) cells have been used to express recombinant DNAs after their introduction by microinjection (Garoff *et al.*, 1983) or transfection (Rose & Bergman, 1982). The flattened appearance of these cells facilitates immunofluorescent analysis of protein location. The various DNA constructs as well as the appropriate mRNA preparations were microinjected into BHK cells, and the cells cultured for eight hours before fixing and study by immunofluorescence. The results are shown in Figure 12, and allow several conclusions: first, expression of all the inserted DNAs was observed when the inserted DNA was in the + orientation. However, expression from the ovalbumin constructs was very poor as judged by the immunofluorescence comparison between mRNA-injected and DNA-injected cells. The combination of poor expression and high background probably obscures the reticular fluorescence that is so characteristic of stained secretory networks (Garoff *et al.*, 1983), and which is seen in the case of mRNA and prochymosin and lysozyme DNA injections. Second, similar levels of fluorescence were observed between pSV_1OV+ (i.e. plus intron)-injected and pSV_2OV+ (no intron)-injected cells. Finally, no difference has been observed in levels of protein expression between constructs containing SV40, TK or LTR promoters, although immunofluorescence comparisons are highly inaccurate.

4. Discussion

(a) *Choice of vector and promoter*

When cloned rabbit globin genes were first expressed in cultured mammalian cells, it was reported that the presence of an intron in the transcribed region of the gene was essential to the stability of the transcript, although the exact position of this intron is unimportant (Hamer & Leder, 1979). The generality of this finding has been challenged, as further genes have been expressed in cultured cells. For

Fig. 10. Synthesis of lysozyme after DNA injection. Oocytes were injected with (a) pTK_1Lys+ or (b) pTK_2Lys+, $pLTR_2Lys+$ and pSV_2Lys+ and cultured and processed as described in Fig. 3. Lys indicates the [^{14}C]lysozyme band present in the ^{14}C-labelled marker proteins. The arrowhead marks a band of similar mobility present only in the DNA injected oocytes. The additional bands seen in the S tracks in (b) are probably due to microbial contamination of the media.

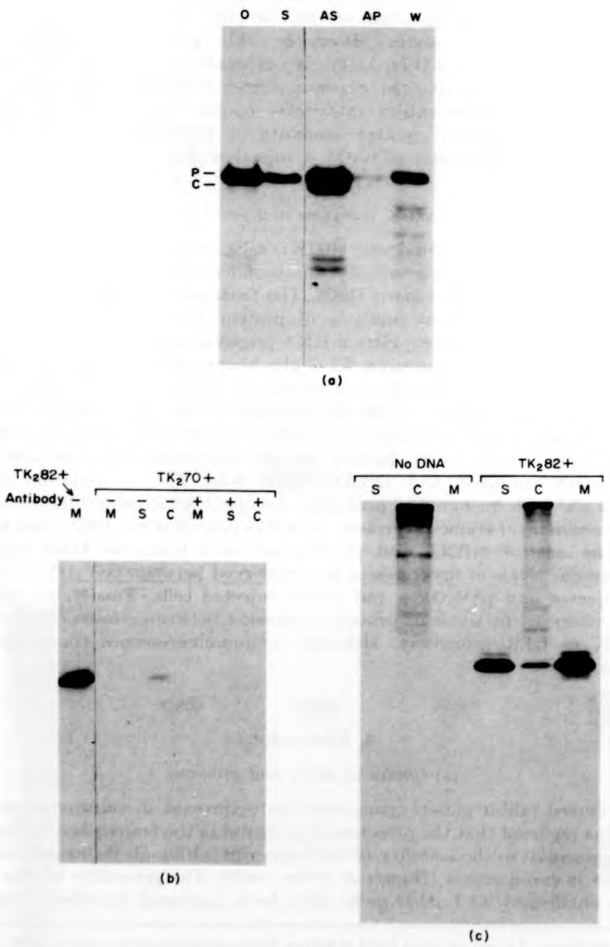


FIG. 11.

several genes, the presence of an intron is unnecessary (Gething & Sambrook, 1982) or can lead to lower or aberrant expression at the protein level (Kondor-Koch *et al.*, 1983). We have expressed chicken ovalbumin cDNA in two SV40/pBR "shuttle" vectors (Mulligan & Berg, 1980), which differ in the provision in only one of them of an intron and associated flanking sequences in the region encoding the 3′ untranslated region of the transcripts initiating at the SV40 early promoter. From our results (Fig. 3), it was clear that the presence of the intron and/or its flanking regions led to reduced protein expression. We do not know the reason for this. Whilst the efficiency of splicing foreign gene transcripts is somewhat variable in oocytes (D. A. Melton, personal communication), Wickens & Gurdon (1983) have reported that 50% of stable SV40 late transcripts are correctly spliced in oocytes; a 50% reduction in "correct" transcripts would not explain the differences in expression we observed. However, in the absence of appropriate pulse–chase experiments, the possibility cannot be excluded that many unspliced or incorrectly spliced transcripts are highly unstable in the Wickens & Gurdon (1983) experiments.

The replacement of the SV40 early promoter with the *Herpes simplex* thymidine kinase promoter resulted in sevenfold higher ovalbumin expression in oocytes. This substitution of promoters had a similar effect on lysozyme expression, although the use of the murine sarcoma virus long terminal repeat promoter in place of the TK promoter did not further boost lysozyme protein production. Since no quantitative comparisons of transcript abundance were made following protein expression with the different promoters, the relatively inferior performance of the SV40 promoter in oocytes remains unclear.

The occurrence of transcripts from a promoter of reverse polarity to the TK (or SV40) promoter was unexpected, since this promoter had not been reported before and probably does not function in cultured mammalian cells (A. Colman, unpublished results). We do not know whether use of this promoter reflects some special characteristic of amphibian oocyte nuclei or one of amphibian cell nuclei in general. Clearly, steady-state levels of transcripts initiating at this promoter are

FIG. 11. Synthesis of prochymosin after mRNA or DNA injection. (a) mRNA injection. Oocytes were injected with preprochymosin mRNA (1 mg/ml) or distilled water and cultured in media containing [^{35}S]methionine as described in Materials and Methods. Oocytes were then homogenized in resuspension buffer (see Materials and Methods) and homogenates (O) and incubation media (S) were immunoprecipitated with anti-prochymosin antibody and electrophoresed: 100 µl of medium was also acidified by addition of HCl to pH 2·0. The acidified fraction was left for 25 min at 0°C before neutralization to pH 7·0 by addition of 2 M-Tris·HCl (pH 7·4). The sample was then centrifuged (10,000 *g*, 10 min) and the supernatant (AS) and the pellet, solubilized in resuspension buffer (AP), were immunoprecipitated. The final track shows the immunoprecipitated products from a wheat germ translation of preprochymosin mRNA (W). Abbreviations: P, position of Coomassie blue-stained prochymosin; C, position of Coomassie blue-stained chymosin. (b) and (c) DNA injection. *Bcl*I fragments containing preprochymosin, chymosin (Mellor *et al.*, 1983) and prochymosin (Emtage *et al.*, 1983), were end-filled and then ligated to *Hin*dIII linkers. After *Hin*dIII digestion, the fragments were ligated into the *Hin*dIII site of the pTK$_2$ expression vector (Fig. 1). Constructs pTK$_2$82+ (preprochymosin), pTK$_2$70+ (met-prochymosin) and pTK$_2$86+ (met-chymosin) were injected into oocytes and cultured and processed as in Fig. 3 with the following changes. First, anti-prochymosin or anti-chymosin antibodies were used. Secondly, in some experiments, 50 nl of rabbit antiserum containing anti-prochymosin antibodies were injected into oocytes before their incubation in radioactive media.

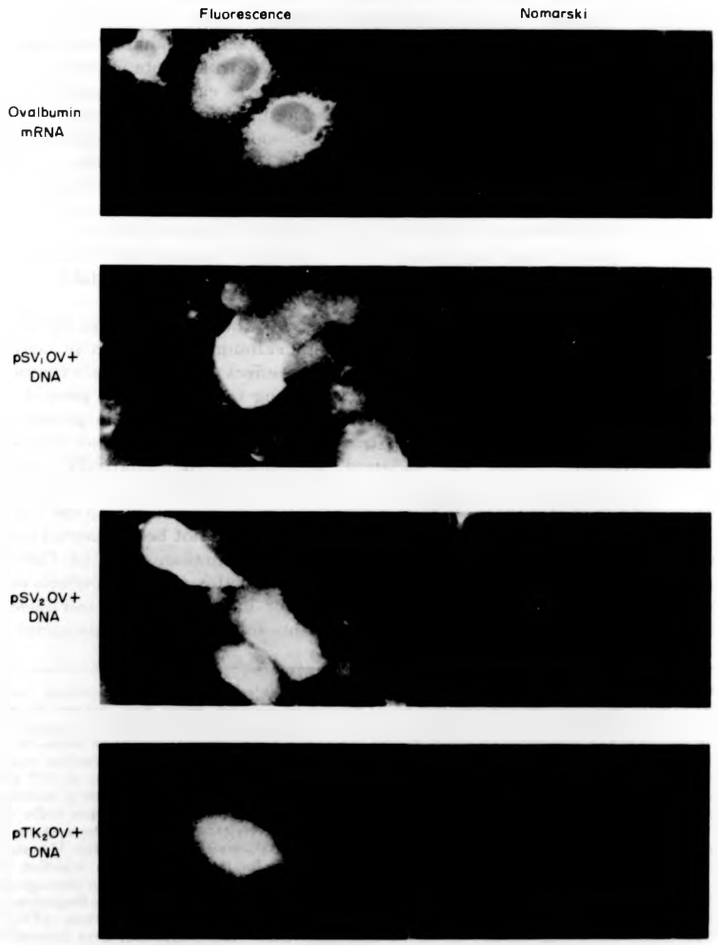


FIG. 12.

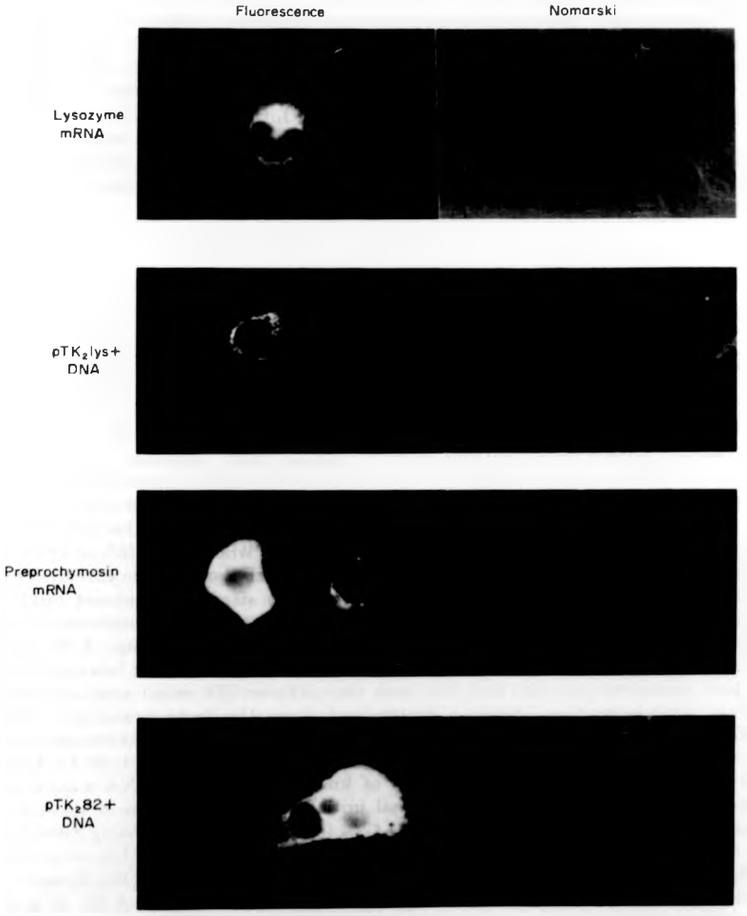


FIG. 12. Microinjection of BHK cells with mRNA or cloned DNA. Cultured BHK cells were injected with mRNA or cloned DNA (both at 1 mg/ml) as indicated and cultured for 8 h before fixing and immunofluorescent staining as described in Materials and Methods. Fluorescence micrographs are shown on the left; the corresponding Nomarski or, in one case (pTK$_2$Lys+), phase micrographs are shown on the right.

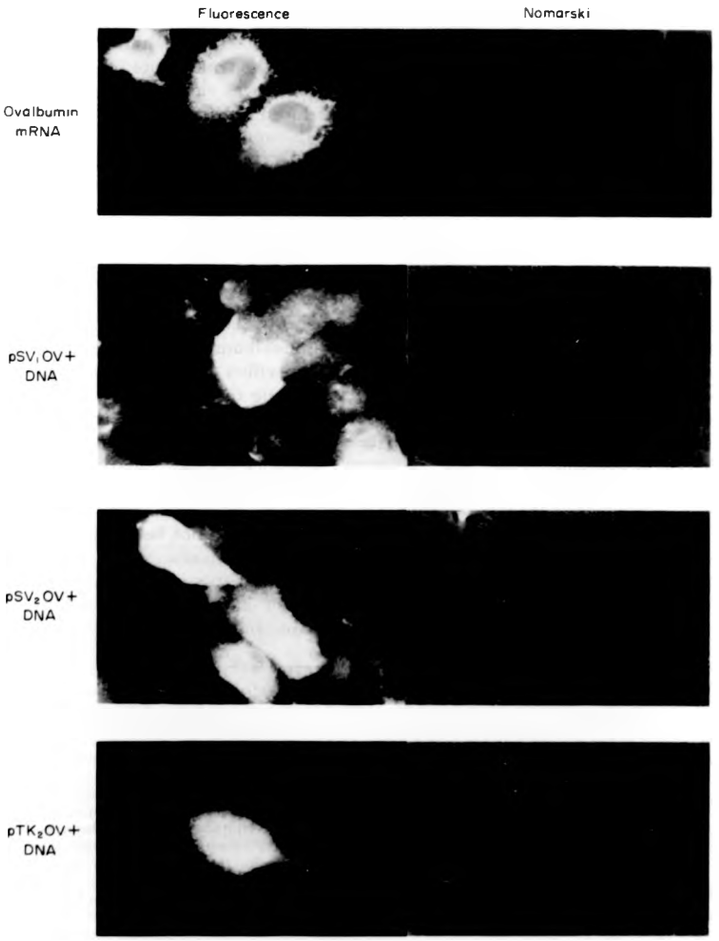


FIG. 12

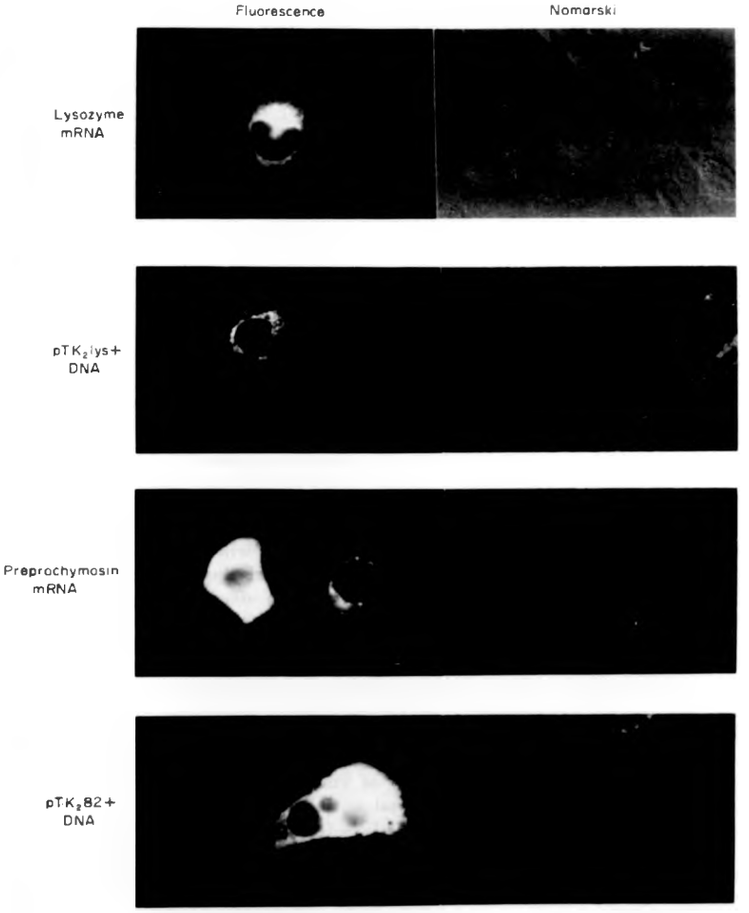


FIG. 12. Microinjection of BHK cells with mRNA or cloned DNA. Cultured BHK cells were injected with mRNA or cloned DNA (both at 1 mg/ml) as indicated and cultured for 8 h before fixing and immunofluorescent staining as described in Materials and Methods. Fluorescence micrographs are shown on the left; the corresponding Nomarski or, in one case (pTK$_2$Lys+), phase micrographs are shown on the right.

much lower than those from the TK promoter (Fig. 6 and our unpublished results), making it difficult to account for the larger amount of protein seen when the injected DNA contains ovalbumin coding sequences inserted in the opposite polarity to the TK (or SV40) promoter (see Figs 3(a) and 4(a)). It is also difficult to explain why certain injected DNAs, e.g. preprochymosin cDNA ($pTK_2 82+$), always express well at the level of protein synthesis, whereas protein expression from others, e.g. prelysozyme cDNA ($pTK_2 Lys+$), is highly variable, whilst the relative levels of transcription from each construct remain similar from experiment to experiment. Variation in protein expression from injected DNA between different batches of oocytes has been noted by others (Asselbergs *et al.*, 1983; Jones *et al.*, 1984).

(b) *Expression of specific cDNA inserts*

(i) *Ovalbumin*

From the results described in Results, section (d)(ii), it is clear that a discrepancy exists between the amounts of ovalbumin produced after DNA or mRNA injection. This could arise from differences in the 5′ untranslated region of injected mRNA and transcripts from the injected recombinant DNA. McReynolds *et al.* (1978) noted the presence of a palindromic sequence at the 5′ end of ovalbumin mRNA that could give rise to a striking hairpin structure. These bases were excised in the preparation of the ovalbumin inserts used in this work (Fig. 2). It is conceivable that their presence enhances translational efficiency.

Apart from the above quantitative differences in translation, the fate of the synthesized wild type ovalbumin proteins are similar whether mRNA or DNA is injected. Two additional ovalbumin species, which are shown to be extended by 21 or 51 amino acids at the N terminus, were seen after DNA but not mRNA injection. The smaller, extended protein was secreted, whilst the larger remained in the cytosol. The significance of the different compartmentation of the two proteins is discussed in the accompanying paper (Tabe *et al.*, 1984); however, it is worth commenting on the fact that both the extended (21 amino acid) and wild type ovalbumins are probably synthesized from the same transcript after $pTK_2 OV-$ or $pSV_2 OV-$ injection. The 5′ sequences flanking the extended and normal AUG initiation codons are 5′-G-G-U-A-U-G-G-G-3′ and 5′-A-C-C-A-U-G-G-3′, respectively. From a survey of known eukaryotic mRNA sequences, Kozak (1983) concluded that functional initiation codons occur in a restricted sequence context with 5′-A-N-N-A-U-G-G-3′ (normal ovalbumin) being favoured over 5′-G-N-N-A-U-G-G-3′ (21 amino acid extension). Kozak (1984) has shown for the expression of preproinsulin in cultured cells that a change of the A residue three residues upstream of the initiator AUG (5′-A-U-U-A-U-G-A-3′), to a G (5′-G-U-U-A-U-G-A-3′) or C residue (5′-C-U-U-A-U-G-A-3′) led to a 3 or 15-fold reduction of proinsulin formation, respectively. However, it is also clear from Kozak (1984) that the presence of an AUG 54 bases upstream of the wild type AUG led to preferential translation from the upstream AUG, even though the flanking sequence (5′-C-U-U-A-U-G-A-3′) is not optimal for preproinsulin expression (see above). This apparent dominance of "position" effect over sequence context would explain the predominant synthesis of the 21 amino acid extended ovalbumin over the wild type and might also account for the appearance of the 51 amino acid extended ovalbumin, since the surrounding sequence here, 5′-C-N-N-A-U-G-C-3′ was only noted in only one out of over 180 eukaryotic mRNAs (Kozak, 1983).

(ii) *Lysozyme*

We demonstrated previously that lysozyme is secreted from oocytes over 12 times faster than ovalbumin after mRNA injection (Cutler *et al.*, 1981). Similar results were obtained after DNA injection (cf. Figs 3 and 10).

In Figure 12, the distribution of lysozyme and ovalbumin are displayed following mRNA and DNA injection into cultured cells. Because of background fluorescence problems, it is difficult to assess ovalbumin distribution following DNA injection; however, the distribution of lysozyme and ovalbumin following RNA injection appears different, with lysozyme fluorescence concentrated in a perinuclear region, which is probably the Golgi apparatus; in contrast, ovalbumin has a reticular distribution, indicating a predominant localization in the endoplasmic reticulum. These locations are consistent with a faster intracellular transport time for lysozyme than ovalbumin in cultured cells as well as oocytes.

(iii) *Preprochymosin and prochymosin*

Injection of preprochymosin DNA ($pTK_2 82+$) elicits the production of prochymosin within the oocyte. This protein is segregated within oocyte membranes and secreted (Fig. 11(b)). When prochymosin DNA ($pTK_2 70+$) was injected, only a small amount of prochymosin was detected and, as expected for a protein lacking its signal sequence, this protein was located in the cytosol compartment. Initiation of transcription from the TK promoter in $pTK_2 82+$ and $pTK_2 70+$ appeared similar (not shown). It is tempting to speculate that prochymosin in the cytosol is unstable. However, the regions immediately flanking the initiator ATG in $pTK_2 82+$ and $pTK_2 70+$ are different (Mellor *et al.*, 1983) and this could affect translational efficiency (see above). Interestingly, when these same cDNA inserts were expressed in yeast cells, 20-fold more protein was detected from the prochymosin cDNA (Mellor *et al.* 1983).

5. Conclusions

This paper describes the expression, in frog oocytes, of genes for three secretory proteins, in a variety of vector/promoter combinations. Based on the results of these experiments we have chosen, for use in further studies, an optimal vector containing the *Herpes simplex* thymidine kinase gene promoter and lacking an SV40 intron. In the accompanying paper (Tabe *et al.*, 1984), we demonstrate how this expression system was used to assess the effect of *in vitro* mutagenesis on the secretion of chicken ovalbumin from injected oocytes.

The authors are grateful to the Medical Research Council and Cancer Research Campaign for support. Particular thanks are due to Beate Timm (EMBL, Heidelberg) for assistance with tissue cell microinjections, and to Henrik Garoff (EMBL, Heidelberg) for advice during the initial stages.

REFERENCES

Ash, J. F., Louvard, D. & Singer, S. J. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5584–5588.
Asselbergs, F. A. M. (1979). *Mol. Biol. Rep.* **5**, 199–208.
Asselbergs, F. A. M., Smart, J. E. & Mathews, M. B. (1983). *J. Mol. Biol.* **163**, 209–238.
Bonner, W. M. & Laskey, R. A. (1974). *Eur. J. Biochem.* **46**, 83–88.
Colman, A. (1982). *Trends Biochem. Sci.* **7**, 435–437.
Colman, A. (1984). In *Transcription and Translation—A Practical Approach* (Hames, D. & Higgens, S., eds), pp. 271–302. IRL Press, Oxford, Washington, DC.
Colman, A. & Morser, J. (1979). *Cell*, **17**, 517–526.
Colman, A., Lane, C. D., Craig, R., Boulton, A., Mohun, T. & Morser, J. (1981). *Eur. J. Biochem.* **113**, 339–348.
Cutler, D., Lane, C. D. & Colman, A. (1981). *J. Mol. Biol.* **153**, 917–931.
Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 3937–3941.
Emtage, J. S., Angal, S., Doel, M. T., Harris, T. J. R., Jenkins, B., Lilley, G. & Lowe, P. A. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **80**, 3671–3675.
Fraser, T. H. & Bruce, B. J. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 5936–5940.
Garoff, H., Kondor-Koch, C., Pettersson, R. & Burke, B. (1983). *J. Cell Biol.* **97**, 652–658.
Gething, M. J. & Sambrook, J. (1982). *Nature (London)*, **300**, 598–603.
Graessmann, A., Graessmann, M. & Mueller, C. (1980). *Methods Enzymol.* **65**, 816–825.
Gurdon, J. B. & Melton, D. A. (1981). *Annu. Rev. Genet.* **15**, 189–218.
Hamer, D. H. & Leder, P. (1979). *Cell*, **18**, 1299–1302.
Harris, T. J., Lowe, P. A., Lyons, A., Thomas, P. G., Eaton, M. A., Millican, T. A., Patel, T. P., Bose, C. C., Carey, N. H. & Doel, M. T. (1982). *Nucl. Acids Res.* **10**, 2177–2187.
Jones, N. C., Richter, J. D., Weeks, D. L. & Smith, L. D. (1984). *Mol. Cell Biol.* **3**, 2131–2142.
Kondor-Koch, C., Reidel, H., Soderberg, K. & Garoff, H. (1982). *Proc. Nat. Acad. Sci., U.S.A.* **79**, 4525–4529.
Kondor-Koch, C., Burke, B. & Garoff, H. (1983). *J. Cell Biol.* **97**, 644–651.
Kozak, M. (1983). *Microbiol. Rev.* **47**, 1–45.
Kozak, M. (1984). *Nature (London)*, **308**, 241–246.
Kressmann, A., Clarkson, S. G., Pirrotta, V. & Birnsteil, M. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 1176–1180.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. (1981). *Nucl. Acids Res.* **9**, 2251–2266.
Lane, C. D. (1981). *Cell*, **24**, 281–282.
Lane, C. D. (1983). *Curr. Top. Develop. Biol.* **18**, 89–116.
Lane, C. D. & Knowland, J. S. (1975). In *Biochemistry of Animal Development* (Weber, R. A., ed.), vol. 3, pp. 145–181. Academic Press, New York.
Lane, C. D., Colman, A., Mohun, T., Morser, J., Champion, J., Kourides, I., Craig, R., Higgins, S., James, T. C., Appelbaum, S. W., Ohlsson, R. I., Paucha, E., Houghton, M., Matthews, J. & Miflin, B. J. (1980). *Eur. J. Biochem.* **111**, 225–235.
Lischwe, M. A. & Ochs, D. (1982). *Anal. Biochem.* **127**, 453–457.
Maniatis, T., Frisch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
Marbaix, G. & Huez, G. (1980). In *Transfer of Cell Constituents into Eukaryotic Cells. NATO Advanced Study Institutes. Series A: Life Sciences* (Celis, J., Graessmann, A. & Loyter, A., eds), pp. 347–382. Plenum Press, New York.

McKnight, S. & Kingsbury, R. (1982). *Science*, **217**, 316–324.
McReynolds, L. A., Catterall, J. F. & O'Malley, B. W. (1977). *Gene*, **2**, 217–231.
McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M. & Brownlee, G. G. (1978). *Nature (London)*, **273**, 723–728.
Mellor, J., Dobson, M. J., Roberts, N. A., Tuite, M. F., Emtage, J. S., White, S., Lowe, P. A., Patel, T., Kingsman, A. J. & Kingsman, S. M. (1983). *Gene*, **24**, 1–14.
Mulligan, R. C. & Berg, P. (1980). *Science*, **209**, 1422–1427.
Old, R. W., Woodland, H. R., Ballantine, J. E. M., Aldridge, T. C., Newton, C. A., Bains, W. A. & Turner, P. C. (1982). *Nucl. Acids Res.* **10**, 7561–7580.
Probst, E., Kressmann, A. & Birnstiel, M. L. (1979). *J. Mol. Biol.* **135**, 709–732.
Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. & Weissman, S. M. (1978). *Science*, **200**, 494–502.
Roberts, B. E. & Paterson, B. M. (1973). *Proc. Nat. Acad. Sci., U.S.A.* **70**, 2330–2334.
Rose, J. K. & Bergman, J. E. (1982). *Cell*, **30**, 753–762.
Rose, J. K. & Bergman, J. E. (1983). *Cell*, **34**, 513–524.
Rungger, D. & Turler, H. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 6073–6077.
Sabatini, D. D., Kreibach, G., Monimoto, T. & Adesnik, M. (1982). *J. Cell. Biol.* **92**, 1–22.
Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.
Soreq, H. (1984). *Crit. Rev. Biochem.* In the press.
Sveda, M. M., Markoff, L. J. & Lai, C.-J. (1982). *Cell*, **30**, 649–656.
Tabe, L., Krieg, P., Strachan, R., Jackson, D., Wallis, E. & Colman, A. (1984). *J. Mol. Biol.* **180**, 645–666.
Talmadge, K., Stahl, S. & Gilbert, W. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 3369–3373.
Tkacz, J. & Lampen, J. (1975). *Biochem. Biophys. Res. Commun.* **65**, 248–257.
Valle, G., Besley, J., Williamson, A. R., Mosmann, T. R. & Colman, A. (1983). *Eur. J. Biochem.* **132**, 131–138.
Weiss, R. A., Teich, N. M., Varmes, H. & Coffin, J. M. (1984). In *Molecular Biology of Tumor Viruses*, part 3, 2nd edit., Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
Wickens, M. P. & Gurdon, J. B. (1983). *J. Mol. Biol.* **163**, 1–26.
Wickens, M. P., Wood, S., O'Malley, B. W. & Gurdon, J. B. (1980). *Nature (London)*, **285**, 628–634.
Wilkie, N. M., Eglin, R. P., Sanders, P. G. & Clements, J. B. (1980). *Proc. Roy. Soc. ser. B*, **210**, 411–421.

Edited by W. Franke

Reprinted from *J. Mol. Biol.* (1984) **180**, 645–666

Segregation of Mutant Ovalbumins and Ovalbumin–Globin Fusion Proteins in *Xenopus* Oocytes

Identification of an Ovalbumin Signal Sequence

L. Tabe, P. Krieg, R. Strachan, D. Jackson, E. Wallis and A. Colman

J. Mol. Biol. (1984) **180**, 645–666

Segregation of Mutant Ovalbumins and Ovalbumin–Globin Fusion Proteins in *Xenopus* Oocytes

Identification of an Ovalbumin Signal Sequence

L. Tabe, P. Krieg†, R. Strachan, D. Jackson, E. Wallis and A. Colman

Medical Research Council Developmental Biology Group
Department of Biological Sciences, University of Warwick
Coventry CV4 7AL, U.K.

(Received 5 June 1984, and in revised form 3 September 1984)

The intramolecular signals for chicken ovalbumin secretion were examined by producing mutant proteins in *Xenopus* oocytes. An ovalbumin complementary DNA clone was manipulated *in vitro*, and constructs containing altered protein-coding sequences and either the simian virus 40 (SV40) early promoter or *Herpes simplex* thymidine kinase promoter, were microinjected into *Xenopus laevis* oocytes. The removal of the eight extreme N-terminal amino acids of ovalbumin had no effect on the segregation of ovalbumin with oocyte membranes nor on its secretion. A protein lacking amino acids 2 to 21 was sequestered in the endoplasmic reticulum but remained strongly associated with the oocyte membranes rather than being secreted. Removal of amino acids 231 to 279, a region previously reported to have membrane-insertion function, resulted in a protein that also entered the endoplasmic reticulum but was not secreted. Hybrid proteins containing at their N terminus amino acids 9 to 41 or 22 to 41 of ovalbumin fused to the complete chimpanzee α-globin polypeptide were also sequestered by oocyte membranes. We conclude that the ovalbumin "signal" sequence is internally located within amino acids 22 to 41, and we speculate that amino acids 9 to 21 could be important for the completion of ovalbumin translocation through membranes.

1. Introduction

The original signal hypothesis (Blobel & Dobberstein, 1975*a*) envisaged that the co-translational transfer through the endoplasmic reticulum membrane of the nascent chains of secretory proteins occurred in a threadlike manner. Observations on the final disposition in the membrane of various transmembrane proteins has necessitated a change in the original model (Blobel, 1980) as well as prompting the formulation of alternative models (Inouye *et al.*, 1977; von Heijne & Blomberg, 1979; Wickner, 1980; Engelman & Steitz, 1981) to accommodate the insertion of looped polypeptides into and through the ER‡ membrane. In some of these models, a looped mode of insertion has been evoked for membrane and

† Present address: Department of Biochemistry and Molecular Biology, 7 Divinity Avenue, Harvard University, Cambridge, Mass. 02138, U.S.A.

‡ Abbreviations used: ER, endoplasmic reticulum; cDNA, complementary DNA; bp, base-pair(s); kb, 10^3 bp.

0022-2836/84/350645-22 $03.00/0

secretory proteins on the basis of a postulated interaction between the N-terminal amino acids of these proteins and the cytoplasmic side of the ER membrane (Inouye *et al.*, 1977; von Heijne & Blomberg, 1979) or as a means of generating favourable free energies for polypeptide transfer through the membrane (Engelman & Steitz, 1981). In other models, the looped insertion mechanism is extended from membrane proteins to secretory proteins essentially for the sake of consistency (Blobel, 1980). However, in the case of one secretory protein, chicken ovalbumin, a looped insertion model has seemed unavoidable due to the position of its signal sequence. This protein has been shown to lack a cleavable signal peptide (Palmiter *et al.*, 1978), although studies *in vitro* have demonstrated the presence of the functional equivalent of a signal sequence (Lingappa *et al.*, 1978). On the basis of competition studies *in vitro*, Lingappa *et al.* (1979) ascribed signal function to a region of ovalbumin between amino acids 229 and 279; however, subsequent studies involving different strategies *in vitro* have located the signal within the N-terminal 150 (Braell & Lodish, 1982) or 70 (Meek *et al.*, 1982) amino acids of the protein. Examination of the amino acid sequence of ovalbumin reveals that the first 25 amino acids bear no resemblance in terms of hydrophobicity index to any identified signal sequence (for a review, see Chan & Bradley, 1982), although there is a suitably hydrophobic region between amino acids 28 and 46 (McReynolds *et al.*, 1978). This evidence suggests the presence of a signal sequence in the N-terminal region of ovalbumin, although not extending to the extreme N terminus.

The function ascribed to signal sequences has undergone some revision recently. Originally, the hydrophobic signal peptide was thought to interact directly with some membrane component. The discovery of the signal recognition particle (Walter & Blobel, 1980) and recent insight into its interaction with the signal sequence (Gilmore & Blobel, 1983) and with docking protein (Meyer & Dobberstein, 1983) during only the initial stages of translocation of nascent secretory polypeptides make it unclear as to whether the signal sequence has a further role during this translocation process. If there is an interaction between the signal and a permanent membrane component, then for ovalbumin transfer this interaction must be transient, since the signal sequence is eventually secreted as part of the mature protein. For all other secretory or membrane proteins, the signal is either cleaved off or remains as part of the membrane-spanning region (Markoff *et al.*, 1984). It has been suggested that phylogenetically, secretory proteins arose from membrane proteins by proteolytic release of their externalized domains (Sabatini *et al.*, 1982). If movement of the signal peptide out of a membrane is problematical for secretory proteins, and is normally solved by cleavage, then in the case of ovalbumin other mechanisms may be involved in its release from the ER membrane.

In this paper, we investigate the contribution of various regions of the ovalbumin polypeptide to its sequestration by, and subsequent secretion from oocyte membranes. The expression of normal and variant ovalbumins was achieved by the injection into *Xenopus* oocyte nuclei of plasmid DNA containing cloned ovalbumin complementary DNA as described in the accompanying paper (Krieg *et al.*, 1984).

2. Materials and Methods

(a) *Materials*

Sources of chemicals, biochemicals and radiochemicals are exactly as reported in the accompanying paper (Krieg *et al.*, 1984), except for rabbit anti-human globin, which was purchased from Miles-Yeda.

(b) *Construction of mutant ovalbumins*

Unless otherwise stated, the genetic manipulations described below were made using the methods described by Maniatis *et al.* (1982). As explained by Krieg *et al.* (1984), the symbols + and −, used in construct descriptions, refer to the polarity of the coding region of the insert in relation to the simian virus 40 (SV40) early promoter or the *Herpes simplex* thymidine kinase promoter present in the vectors pSV_2 or pTK_2, respectively.

(c) $pSV_2OV_{Sst}+$

$pSV_2OV_{Sst}+$, lacking bases 82 to 460 of the *Hin*dIII ovalbumin cDNA insert (Fig. 1) was made by cutting pSV_2OV+ with *Sst*I and gel-purifying the larger fragment (the vector contains no *Sst*I site). The larger fragment was ligated and transformation performed. The deleted region corresponds to amino acids 20 to 145 of ovalbumin.

(d) $pTK_2OV_{Sau}+$

$pTK_2OV_{Sau}+$, lacking bases 717 to 864 of the *Hin*dIII-ovalbumin cDNA insert (Fig. 1), was constructed as follows: $pSV_2OV_{Sst}+$ (above) was cut with *Hin*dIII and the 926 bp ovalbumin-coding insert purified. This fragment was digested with *Sau*3A, giving fragments of 331 bp, 147 bp and 451 bp. The 331 bp and 451 bp fragments, after purification from a 3% agarose gel, were ligated to form mixed concatamers, which were then digested with *Hin*dIII yielding 3 types of dimers of the original fragments: 2 × 331 (662 bp), 2 × 451 (902 bp) and 331 + 451 (782 bp). This latter, mixed dimer was isolated from a 3% agarose gel and ligated with *Hin*dIII-cut pTK_2 vector (see Krieg *et al.*, 1984), which had been treated with calf intestinal alkaline phosphatase.

A transformant containing the 782 bp insert in the + orientation was grown up and plasmid DNA prepared. The DNA was linearized by digestion with *Sst*I and treated with calf intestinal alkaline phosphatase before ligation with a 378 bp *Sst* fragment purified from a *Sst*I digest of pTK_2OV+. Transformants containing a 1160 bp *Hin*dIII insert were then isolated and the orientation of the inserted *Sst*I fragment checked. The final product of this procedure ($pTK_2OV_{Sau}+$) differed from pTK_2OV+ in the absence of a 147 bp fragment encoding amino acids 231 to 279 of ovalbumin (also see Fig. 2).

(e) $pTK_2OV_{43}+$

pTK_2OV+ DNA was digested with *Hin*dIII and the 1318 bp ovalbumin coding fragment was purified: 5 μg of this fragment were digested at 30°C with 0·5 unit of *Bal*31

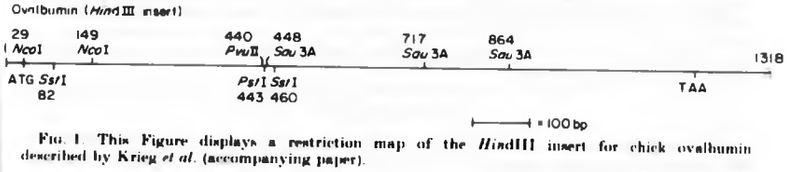


FIG. 1. This Figure displays a restriction map of the *Hin*dIII insert for chick ovalbumin described by Krieg *et al.* (accompanying paper).

exonuclease for 90 s. The *Bal*31-digested DNA was ligated with *Hin*dIII linkers and then cut with *Hin*dIII and *Pst*I. DNA molecules between 300 bp and 400 bp long were isolated from a 3% agarose gel and inserted between the *Hin*dIII and *Pst*I sites in the M13-mp8 vector. Using dideoxy sequencing (Sanger *et al.*, 1977), a transformant was identified in which 42 nucleotides had been removed from the terminus of the ovalbumin insert which had contained the wild-type ATG codon. This DNA fragment was excised with *Hin*dIII and *Pst*I from a double-stranded replicative form of the M13 clone, and ligated with the 875 bp *Pst*I-*Hin*dIII fragment prepared from the 1318 bp ovalbumin insert (nucleotides 443 to 1318, see Fig. 1) and the *Hin*dIII-linearized pTK_2 vector. Transformants were screened, and $pTK_2OV_{43}+$, a construct differing from pTK_2OV+ by the absence of the first 42 bp of the *Hin*dIII-ovalbumin insert (Fig. 1, and see Fig. 2) was isolated.

(f) $pTK_2OV_{Nco}+$

pTK_2OV+ was restricted with *Nco*I producing 4·75 kb and 120 bp fragments. The larger fragment was gel-purified, ligated and transformation performed. The final product ($pTK_2OV_{Nco}+$) differs from pTK_2OV+ by the absence of the 120 nucleotides encoding the first 40 amino acids of ovalbumin (see Fig. 2).

(g) $pTK_2OV_{His}+$

A recombinant DNA ($pTK_2OV_{92}+$) containing all but the first 91 nucleotides of the *Hin*dIII-ovalbumin insert (Fig. 1) was prepared exactly as described for $pTK_2OV_{43}+$, above. The ovalbumin insert was excised and the overhanging 5′ ends filled in using *Escherichia coli* DNA polymerase (Klenow fragment). The DNA was then cut with *Pvu*II and the 349 bp *Hin*dIII-*Pvu*II fragment (nucleotides 92 to 440, Fig. 1) was isolated and treated with calf intestinal phosphatase (fragment A).

pTK_2OV+ was cut with *Nco*I and the 4·75 kb fragment (see section (f), above) was end-filled before restriction with *Acc*I. The 471 bp *Acc*I-*Nco*I fragment, which contains the TK promoter, was gel-purified (fragment B).

Finally, pTK_2OV+ was restricted with *Pvu*II and *Acc*I yielding several small fragments and a 4·0 kb fragment. The largest (4·0 kb) fragment was gel-purified (fragment C).

Fragments A, B and C were ligated, then transformants were screened for plasmids containing one of each fragment in the appropriate order and orientation. The final construct, $pTK_2OV_{His}+$, resembles pTK_2OV+ in that the 5′ leader sequence and initiator ATG are present. It differs in that the nucleotides encoding amino acids 2 to 21 of ovalbumin are absent, although 3 new amino acids are introduced by the *Hin*dIII linker used in the construction (Fig. 2).

(h) $pTK_2OV_{43}G+$

$pTK_2OV_{43}+$ DNA was cut with *Stu*I and *Nco*I yielding fragments of 4·2 kb and 634 bp, the larger of which was gel-purified (fragment A). The plasmid $pGOV_2$ (a gift from V. Lingappa, UCSF, San Francisco, U.S.A.) was cut with *Stu*I and *Nco*I yielding a large fragment and a fragment of approximately 990 bp containing the entire protein-coding region of chimpanzee α_1-globin cDNA, including the normal terminator codon. The 990 bp fragment was gel-purified and ligated with fragment A (see above) and transformants were screened for a construct containing one of each fragment in the correct orientation. The expected protein product of this construct ($pTK_2OV_{43}G+$) would contain residues 9 to 41 of ovalbumin fused to the N terminus of chimpanzee α-globin. No new amino acids are added by these manipulations.

(i) $pTK_2OV_{His}G+$

pTK_2OV_{His} DNA was cut with *Nco*I and *Stu*I yielding fragments of 4·2 kb and 634 bp, the larger of which was gel-purified and ligated with the 990 bp *Nco*I-*Stu*I fragment

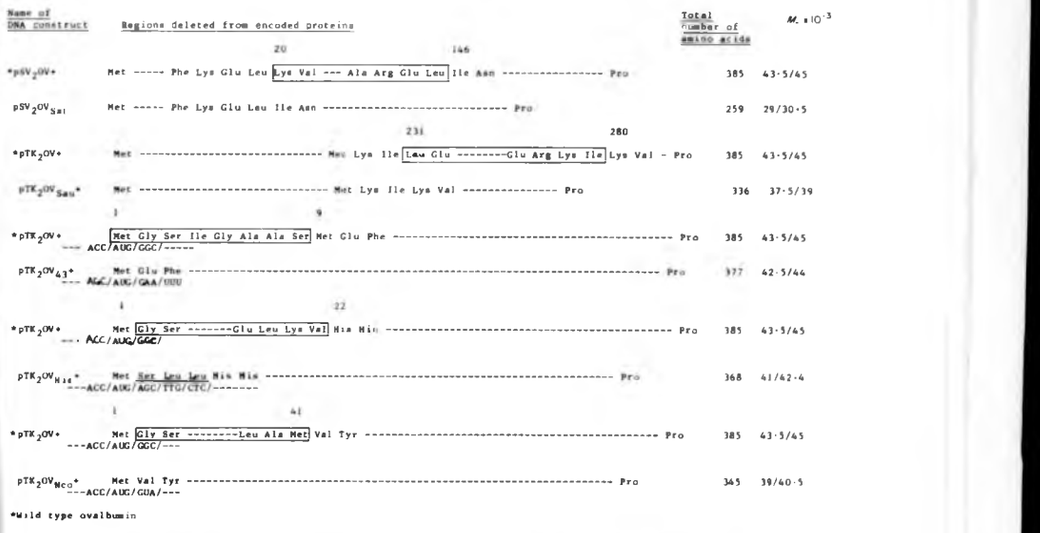


FIG. 2. Amino acid modifications in mutant ovalbumins. The altered coding sequences in each of the mutant constructs described in Materials and Methods are displayed underneath the wild type sequence for the same region. Boxed areas indicate deleted sequences, whilst underlined amino acids ($pTK_2OV_{His}+$) indicate new amino acids introduced in the cloning. Numbers above the amino acids designate the position of the residue in wild type ovalbumin. The nucleotide sequences flanking the initiation codon are also shown. The columns list the total number of amino acids in each protein, and the expected molecular weight of an unglycosylated (first figure) or glycosylated (second figure) wild type or mutant ovalbumin.

encoding chimpanzee globin, described in section (h), above. The expected protein product of this construct would contain the amino acid sequence Met-Ser-Leu-Leu followed by amino acids 22 to 41 of ovalbumin followed by chimpanzee α-globin. Thus, the amino acids introduced at the N terminus of the OV_{His} protein (see section (g), above) are retained in the OV_{His} globin protein, although no new amino acids are introduced at the junction between the ovalbumin and globin.

(j) *Microinjection of oocytes*

Microinjection and culture of *Xenopus* oocytes were as described by Krieg *et al.* (1984).

(k) *Oocyte fractionation*

In most experiments, oocytes were homogenized and fractionated on sucrose step gradients in the presence of phenylmethylsulphonyl fluoride as described by Krieg *et al.* (1984). The supernatants containing cytosolic components were stored at −20°C, whilst the pellets containing membranes were either resuspended in Triton X-100-containing buffer (Krieg *et al.*, 1984) or in T buffer (50 mM-NaCl, 10 mM-magnesium acetate, 20 mM-

Tris · HCl, pH 7·5), modified by the addition of 100 mM-NaCl and 10% (w/v) sucrose, before freezing to −20°C.

In some experiments, injected oocytes (20 to 25) were homogenized in 0·5 ml of modified T buffer containing 1% Triton, before layering on a 10-ml 5% (w/v) to 20% (w/v) linear sucrose gradient in T buffer, with 0·1% (v/v) Triton and centrifuging at 39,000 revs/min for 20 h in a SW40 Beckman rotor at 4°C. Fractions (1 ml) were collected, diluted with 1 ml of 2 × immunoprecipitation buffer and precipitated as described by Krieg *et al.* (1984).

(l) *Carbonate extraction of membranes*

Samples of oocyte membranes (100 μl) resuspended in T buffer were diluted with 100 μl of 200 mM-sodium carbonate (pH 11·0) and left at 0°C for 30 min, before centrifugation at 2·8 kg/cm² in a Beckman airfuge, for 5 min at 4°C. The supernatants were removed and neutralized with 20 μl of 1 M-HCl. The pellet was either resuspended in 200 μl of 100 mM-sodium carbonate and the treatment repeated, or dissolved in 100 μl of homogenization buffer (5 mM-$MgCl_2$, 100 mM-NaCl, 50 mM-Tris · HCl (pH 7·6), 1% (v/v) Triton X-100). Neutralized supernatants and solubilized pellets were then immunoprecipitated.

(m) *Immunoprecipitation and electrophoresis*

Immunoprecipitation and electrophoresis of oocyte media, cytosol or membrane fractions on reducing, 12·5% (w/v) polyacrylamide gels were performed as described by Krieg *et al.* (1984).

3. Results

(a) *Effect of deletions on secretion*

(i) *Removal of amino acids 20 to 145*

Baty *et al.* (1981) have demonstrated that the removal, with *Sst*I, of a section of the ovalbumin gene encoding amino acids 20 to 145 inclusive, prevents the segregation of a β-galactosidase–ovalbumin fusion protein through the inner membrane of the bacterium *Escherichia coli*. We have constructed a similar deletion in the discrete ovalbumin gene contained in the pSV_2OV+ vector (Krieg *et al.*, 1984). As can be seen from Figure 2, no new amino acids are introduced by this manipulation. After injecting this deletion construct ($pSV_2OV_{Sst}+$) into oocytes, a protein migrating at the molecular weight of 29,000, as predicted for an unglycosylated, deleted ovalbumin, is detected only in the cytosol fraction (Fig. 3(a)). We conclude, therefore, that the deleted region contains either all or an essential part of the information necessary to initiate segregation of ovalbumin. Surprisingly, although we have reported the appearance of an extended ovalbumin when the polarity of the complete ovalbumin insert is reversed (i.e. pSV_2OV-, see the accompanying paper), no product at all was detected after injection of $pSV_2OV_{Sst}-$ (results not shown).

(ii) *Removal of amino acids 231 to 279*

Lingappa *et al.* (1979) described data that they interpreted as localizing part or all of the ovalbumin signal sequence between amino acids 229 and 279. These results were obtained using *in vitro* translation and segregation systems, which only examine the information necessary for the protein to gain access to the

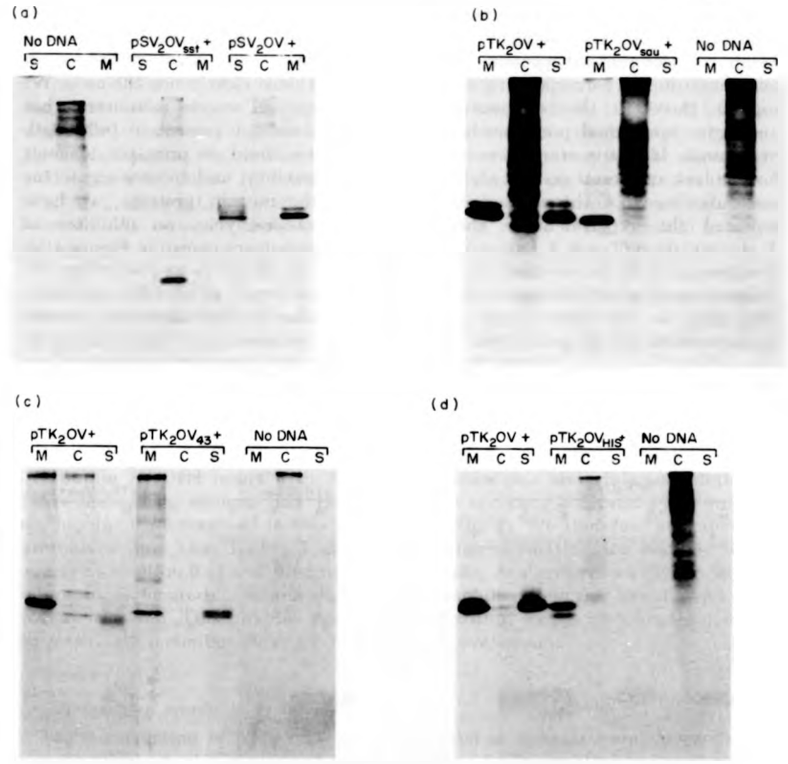


FIG. 3. Effects of mutation on the segregation and secretion of ovalbumin. Oocytes were injected with the various constructs, as indicated, and incubated for 24 h before the addition of [^{35}S]methionine to the medium. After a further 24 h, oocytes were fractionated as described in Materials and Methods into membrane (M) and cytosol (C) fractions. These fractions, and incubation media (S), were immunoprecipitated with anti-ovalbumin antisera and electrophoresed. (a), (b), (c) and (d) are experiments conducted with different batches of oocytes; 2·5 oocyte equivalents of the M and C fractions, and 5 equivalents of the S fraction were loaded on each track.

lumen of the endoplasmic reticulum and not that for any subsequent steps in the secretory process. Using the restriction enzyme *Sau*3A, we removed the nucleotides encoding amino acids 231 to 279 inclusive, from the ovalbumin DNA insert (see Materials and Methods, and Fig. 2). No new amino acids are introduced into the encoded protein. After injection of this construct into oocytes, a major product of apparent molecular weight 43,000 was found segregated in the

membranes, but was not secreted (Fig. 3(b)), in contrast to the full-length wild type protein. As a consequence of the deletion, we anticipated the molecular weight of the mutant protein to be 6000 M_r lower than that of full-length ovalbumin in the corresponding oocyte fraction. This is clearly not the case. We suggest, therefore, that the mutant protein present in oocyte membranes has undergone additional post-translational modifications not present in full-length ovalbumin. If this is true, these extra modifications could, in principle, account for the lack of secretion. In order to test this possibility, and to investigate the molecular basis of the anomalous mobility of the mutant proteins, we have repeated the experiment in the presence of tunicamycin, an inhibitor of *N*-glycosylation (Tzack & Lampen, 1975). In the experiment shown in Figure 4(b), the treatment with tunicamycin was only partially effective, as judged by the prevention of glycosylation of about 50% of wild type ovalbumin molecules (Fig. 4(a)). In the case of the mutant, in addition to the 43,000 M_r species, products of 42,000 M_r and 40,000 M_r were found after treatment with tunicamycin (Fig. 4(b)), the latter (40,000 M_r) species having the expected mobility of an unglycosylated deleted ovalbumin. A product of similar mobility is also seen in

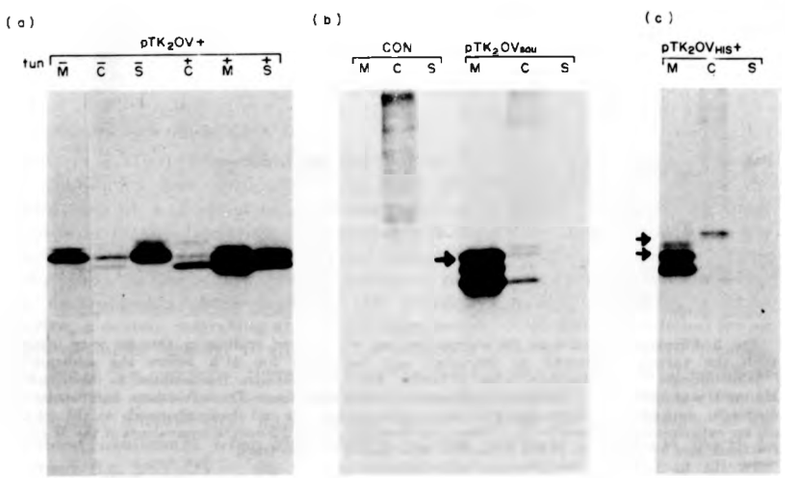


FIG. 4. Effect of tunicamycin on secretion of wild type and mutant ovalbumins. Oocytes were injected with various DNAs as indicated. Incubation and processing were as described for Fig. 3. Tunicamycin (tun) at 40 μg/ml in distilled water was injected into the cytoplasm of the oocytes, about 6 h after the DNA injection, and tunicamycin at 2 μg/ml was added to the media surrounding tunicamycin-injected oocytes. The arrows in (b) and (c) indicate the positions of the major glycosylated forms of the mutant proteins, as seen in Fig. 3, ((b) and (d)). (a), (b) and (c) represent one experiment, using one batch of oocytes. CON, control. All samples in (b) and (c) were from oocytes treated with tunicamycin.

the cytosol. Presumably, this cytosol product arises from miscompartmentation of the molecule during synthesis, a phenomenon we have noted for full-length ovalbumin (Colman *et al.*, 1981; Fig. 4(a)). It would appear, therefore, that the extra modification that causes reduced mobility of the OV_{Sau} protein is an *N*-glycosylation event. The appearance of an intermediate (42,000 M_r) species under the conditions of partial inhibition by tunicamycin might imply that two glycosylation sites are used on the deleted protein rather than the one site used in wild type ovalbumin. In contrast to the unglycosylated wild type ovalbumin (Fig. 4(a); also see Colman *et al.*, 1981), the unglycosylated mutant protein is not secreted. We conclude that the lack of secretion of the OV_{Sau} protein is not a consequence of its glycosylation status.

Expression was also obtained after injection of a construct in which the polarity of the *Hin*dIII OV_{Sau} insert was reversed ($pTK_2OV_{Sau}-$). The mutant protein, with its additional 21 amino acid N-terminal extension (see Krieg *et al.*, 1984), had an apparent molecular weight of 46,000 and was segregated in oocyte membranes, but not secreted (results not shown).

(iii) *Removal of amino acids 1 to 8*

The construct $pTK_2OV_{43}+$, was made by progressive digestion of the ovalbumin *Hin*dIII insert with *Bal*31 exonuclease (see Materials and Methods). DNA sequencing showed that the nucleotides encoding the wild type initiator methionine were removed in this process (see Fig. 2). We therefore anticipated a protein product that initiated at the next internal methionine residue, which occurs at position 9 of the full-length polypeptide. As displayed in Figure 3(c), an immunospecific product with a slightly faster mobility than full-length ovalbumin was synthesized. This product was segregated within oocyte membranes and was secreted with a similar efficiency to full-length ovalbumin.

(iv) *Removal of amino acids 2 to 21*

The construction of $pTK_2OV_{His}+$ is described in Materials and Methods and the new N-terminal region is displayed in Figure 2. Unlike all the above deletion mutants, some new amino acids have been introduced into this construct, giving an expected product 365 amino acids long, in which the first 21 amino acids of ovalbumin have been replaced by the sequence Met-Ser-Leu-Leu (see Fig. 2). Expression of this, and the wild type ovalbumin construct are shown in Figure 3(d). Two immunospecific products that segregate with oocyte membranes are seen after injection of the mutant construct. Neither OV_{His} product is secreted, and the electrophoretic mobilities of the proteins are slower than expected for a deleted ovalbumin. It was, therefore, suspected that the main OV_{His} species represented an abnormally post-translationally modified, deleted protein. Treatment with tunicamycin resulted in a major product of apparent molecular weight 41,000 (Fig. 4(c)), the predicted molecular weight for an unglycosylated ovalbumin containing 365 amino acids (Fig. 2). This unglycosylated product segregated with membranes, but was not secreted.

(v) *Removal of amino acids 1 to 40*

A construct ($pTK_2OV_{Nco}+$) lacking the nucleotide sequence encoding amino acids 1 to 40 of ovalbumin (see Materials and Methods, and Fig. 2) was injected into oocytes. No protein product was detected after injection of this DNA into oocyte nuclei, although transcripts were synthesized and exported to the oocyte cytoplasm (results not shown).

(b) *Sucrose gradient fractionation of OV_{Sau} and OV_{His} proteins*

In the above sections we have demonstrated that the OV_{His} and OV_{Sau} products segregate, during centrifugation, with the oocyte membranes, but are not secreted. One possible explanation for this lack of secretion could be that a changed polypeptide conformation, resulting from mutation, may lead to aggregation of the proteins within the endoplasmic reticulum. In order to assess the aggregation status of the proteins *in vivo*, we have extracted injected oocytes with a buffer containing Triton X-100 and 150 mM-NaCl, after which the homogenates were sedimented on sucrose gradients (Fig. 5). Oocytes injected with ovalbumin mRNA were also processed as a control. It is clear that both mutant proteins and full-length ovalbumin sediment at a similar rate. We conclude that no Triton X-100-insoluble aggregates are formed within the oocytes. Since the relatively low salt concentration in the gradients (150 mM) would not be expected to disrupt ionic associations between proteins, it can also be said that the OV_{Sau} and OV_{His} proteins do not form ionically bonded aggregates *in vivo*.

(c) *Carbonate extraction of isolated membranes containing OV_{His} and OV_{Sau} protein*

An alternative explanation for the absence of OV_{His} and OV_{Sau} secretion is that these proteins are either peripherally or integrally associated with the oocyte membranes. Peripheral and integral membrane association can be distinguished by extraction of membranes with sodium carbonate at pH 11 (Fujiki *et al.*, 1982). When membranes prepared from oocytes injected with pTK_2OV+, $pTK_2OV_{Sau}+$ or $pTK_2OV_{His}+$ were extracted with sodium carbonate, as described in Materials and Methods, the results shown in Figure 6 were obtained. Over 40% of full-length ovalbumin and OV_{Sau} proteins were released by incubation in sodium carbonate (Fig. 6). A second extraction of the membrane pellet caused a further (approx. 40%) release of the residual ovalbumin and OV_{Sau} (results not shown). In contrast, none of the OV_{His} product was released from the membranes (Fig. 6). The co-sedimentation of the OV_{His} product with membranes after extraction is not due to carbonate-induced aggregation of released protein since, on centrifugation of the carbonate-extracted membranes after addition of Triton X-100 (to 1%, v/v), all the OV_{His} protein is found in the supernatant (results not shown). These results indicate that, within the oocyte, the OV_{His} product may be integrated in oocyte membranes, whereas any association of the OV_{Sau} protein with the membranes can be no more than peripheral.

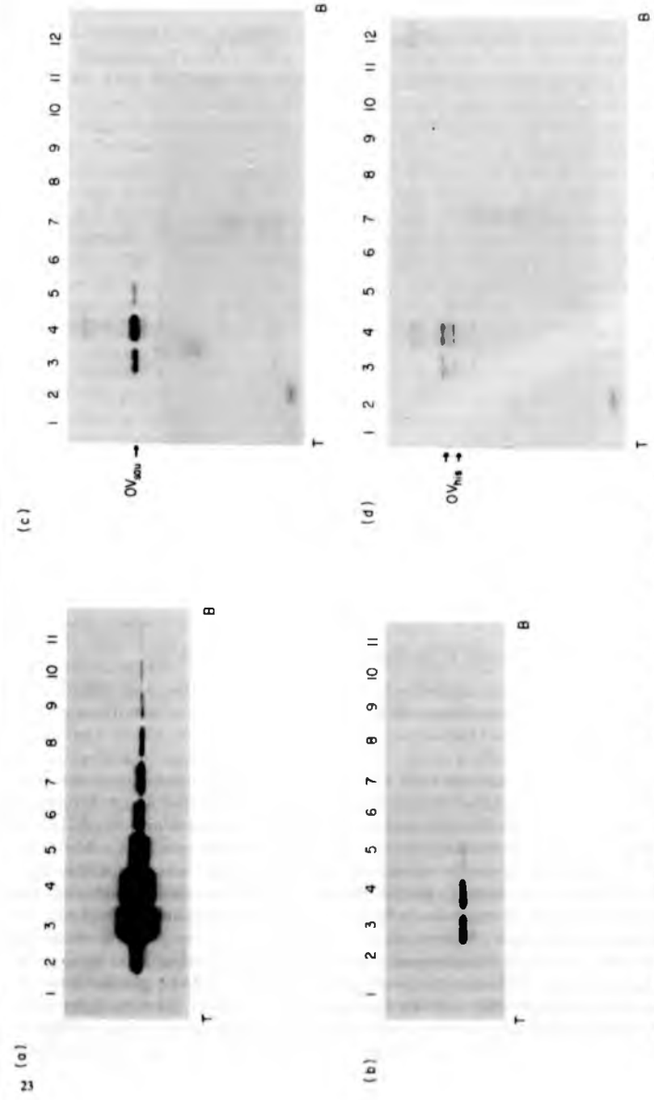


FIG. 5. Sucrose-gradient fractionation of injected oocytes. Oocytes injected with (a) ovalbumin mRNA, (b) and (c) $pTK_2OV_{Sau}+$ DNA or (d) $pTK_2OV_{His}+$ DNA were labeled in media containing ^{35}S]methionine as described for Fig. 3. Oocytes were then homogenized and the homogenate layered onto 10-ml, 5% to 20% (w/v) sucrose gradients as described in Materials and Methods. Fractions (1 ml) were collected from the gradients, after centrifugation, and immunoprecipitated. T, top of gradient; B, bottom of gradient. (a) and (b), and (c) and (d) are separate experiments performed with different batches of oocytes. Abbreviations OV_{Sau} and OV_{His} indicate the positions of the mutant ovalbumins.

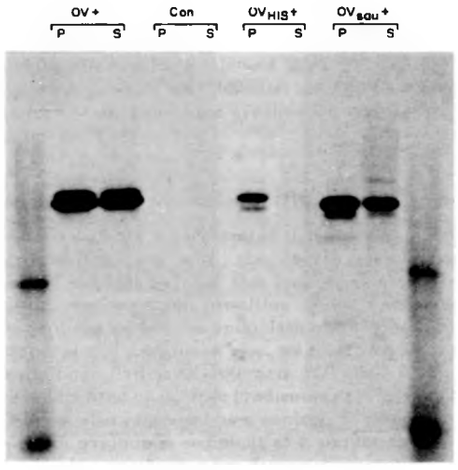


FIG. 6. Carbonate extraction of oocyte membranes. Oocytes were injected with constructs as indicated and processed as described for Fig. 3, with the exception that oocyte membrane fractions were resuspended and stored in modified T buffer. Membranes were then extracted with sodium carbonate (pH 11·0) before centrifugation (see Materials and Methods). Triton X-100 was added to the neutralized supernatants to 1% (v/v), whilst pellets were resuspended in modified T buffer containing 1% (v/v) Triton X-100. The samples were then immunoprecipitated and electrophoresed as described. P, pellets; S, supernatants; Con, control.

(d) *Expression of ovalbumin–globin fusion proteins*

The deletion strategy used above makes the assumption that the failure of a protein to segregate with membranes following the removal of a specific region of the polypeptide, implicates that region as containing part or all of the signal sequence. This strategy suffers from the criticism that gross changes in the conformation of proteins mutated by large deletions may have a non-specific, negative effect on their transmembrane translocation. A positive assay for signal function in a part of the ovalbumin molecule involves the construction of fusions between various regions of ovalbumin and a protein that ordinarily would not be segregated. The successful translocation of the fusion product would indicate that the ovalbumin fragment present in the fusion had "signal" capacity. Although the primary sequence of some cytosolic proteins might itself impose restrictions on transmembrane segregation (Moreno *et al.*, 1980), it has been shown that the presence of the bacterial β-lactamase signal sequence at the N terminus of chimpanzee α-globin is sufficient to ensure the segregation of the globin into the lumen of isolated dog pancreas vesicles (Lingappa *et al.*, 1984). Globins have been shown to exhibit no affinity for membranes either *in vitro* (Meek *et al.*, 1982) or *in vivo*, in *Xenopus* oocytes (Zehavi-Willner & Lane, 1977). The data in the above

sections implicate, as a signal sequence, an internal section of ovalbumin located in the N-terminal half of the protein. We, therefore, constructed two gene fusions in which two N-terminally located regions of ovalbumin coding sequence were fused to the complete protein-coding region of a chimpanzee α-globin cDNA clone. pTK$_2$OV$_{43}$ G+ encodes amino acids 9 to 41 of ovalbumin fused to amino acids 1 to 142 of chimpanzee α-globin, whilst pTK$_2$OV$_{His}$ G+ contains the information for amino acids 22 to 41 of ovalbumin fused to the globin protein (see Fig. 7(a)). Figure 7(b) shows the results of an experiment where these two constructs were expressed in oocytes. After injection of each construct into oocytes, a protein of the expected molecular weight is precipitated by anti-human haemoglobin antibody. The fusion proteins are located predominantly in the membrane fraction, although in the case of the OV$_{His}$ globin protein, a small amount of the full-length protein, and a possible degradation product, are present in the cytosol fraction. No secreted protein is detectable in either sample. These fusion proteins were resistant to digestion of oocyte membranes with trypsin and chymotrypsin in the absence of added Triton X-100. However, in the presence of 1% Triton X-100, the proteins were degraded (see Fig. 7(c)), indicating that the proteins are segregated into the lumen of the oocyte membranes. We conclude that there is sufficient information in the region of ovalbumin comprising amino acids 22 to 41 to direct the insertion and translocation of a cytosolic protein, chimpanzee α-globin, through oocyte membranes.

4. Discussion

(a) *Effects of deletions on ovalbumin secretion*

In this paper, we describe the most recent of several published attempts to identify the uncleaved signal sequence of chicken ovalbumin. It is currently thought that the signal sequence must reside close to the N-terminus of the protein (Meek *et al.*, 1982; Braell & Lodish, 1982) although an earlier report states that signal function may be encompassed by amino acids 229 to 279 of ovalbumin (Lingappa *et al.*, 1979). We therefore concentrated our search in these regions of the protein's primary structure.

It has been noted by Baty *et al.* (1981) that a hybrid protein consisting of the first seven amino acids of *E. coli* β-galactosidase fused to chick ovalbumin lacking amino acids 1 to 5 and 20 to 145, was not inserted through the plasma membrane of *E. coli*. This contrasts with the observed secretion across the bacterial inner membrane of a fusion protein containing the seven N-terminal amino acids of β-galactosidase followed by ovalbumin lacking the first five amino acids (Baty *et al.*, 1981). Taken together, these results indicate that the amino acids 6 to 145 of ovalbumin contain all, or an essential part of the information necessary for the passage of ovalbumin across a membrane, in prokaryotes. We find that deletion of amino acids 20 to 145, from an otherwise unaltered ovalbumin protein, prevents the segregation of the mutant protein inside the ER of *Xenopus* oocytes, thus implying that all or part of the information specifying segregation of ovalbumin in a eukaryotic cell, is also contained within this region.

This conclusion agrees with the work of others who have used synchronized *in vitro* translation systems to estimate that nascent ovalbumin peptides must be at least 50 to 60 amino acids long (Meek *et al.*, 1982), but no longer than 150 amino acids (Braell & Lodish, 1982), in order to associate functionally with microsomal membranes. These studies contradict an earlier report (Lingappa *et al.*, 1979) that a tryptic fragment relatively near the C terminus of ovalbumin, and containing amino acids 229 to 279, competitively inhibited the translocation of nascent rat pre-prolactin into the lumen of dog pancreatic microsomes *in vitro*. On the basis of this competition it was suggested that signal function could be ascribed to a block of hydrophobic amino acids (234 to 253) that shared some amino acid sequence homology with known, N-terminal signal peptides. The results of Meek *et al.* (1982) and Braell & Lodish (1982), argue against this interpretation without precluding the possibility that the "internal signal" region has another role in ovalbumin secretion. For example, rather than effecting the insertion of ovalbumin into the ER membrane, the C-terminal region may release the completed protein into the ER lumen by displacing a more N-terminally located signal peptide from the membrane, in lieu of proteolytic processing.

In order to test the role of the "internal signal" region in ovalbumin secretion, we have created a mutated cDNA, $pTK_2OV_{Sau}+$, lacking the bases encoding amino acids 231 to 279 of the wild-type protein. Expression of this construct in oocytes gave rise to a shortened protein (OV_{Sau}) that co-fractionated with oocyte membrane vesicles, but was not secreted into the culture medium. The protein is over-*N*-glycosylated compared to wild type ovalbumin, as judged by the large increase in electrophoretic mobility of the mutant protein after treatment with tunicamycin. Since *N*-glycosylation is known to occur on the lumenal side of the ER membrane (Hanover & Lennarz, 1981), we conclude that OV_{Sau} gains access to the ER lumen.

Abnormal glycosylation (discussed later) was ruled out as a possible explanation for the lack of secretion of the OV_{Sau} protein, as unglycosylated protein synthesized in the presence of tunicamycin was also segregated in

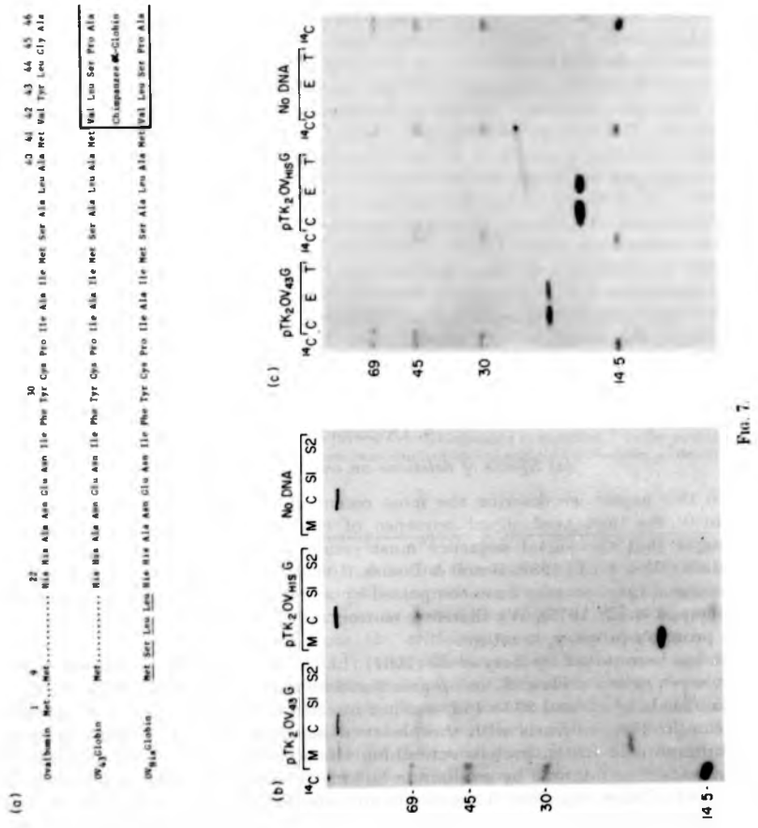


FIG. 7.

FIG. 7. Segregation of ovalbumin–globin fusion proteins. (a) The N-terminal amino acid sequences of wild type ovalbumin and the 2 ovalbumin–globin fusion proteins is shown. Globin amino acids are shown in the boxes. The new amino acids introduced into the N terminus of OV_{His} during its construction are shown underlined. (b) Oocytes were injected with $pTK_2OV_{43}G+$ and $pTK_2OV_{His}G+$ DNA and then cultured as described for Fig. 3, except that oocytes were incubated for only 9 h in radioactive media. At this time, some oocytes were fractionated into membrane (M) and cytosol (C) fractions, whilst the remainder were incubated for a further 8 h in unlabelled media containing 10 mM-methionine. Media from the 0 to 9 h (S1) and 9 to 17 h (S2) oocytes were analysed separately. All samples were immunoprecipitated with rabbit anti-human globin and electrophoresed on 12·5% polyacrylamide gels. In contrast to the previous experiments, 4 oocyte equivalents of the V and C fractions, and 3 equivalents of the S fractions were loaded onto each track. (c) Membrane fractions were prepared from oocytes as in (b), except that phenylmethylsulphonyl fluoride was omitted from the buffers. The membranes were resuspended in 300 μl of modified T buffer containing 10% (w/v) sucrose. Samples (100 μl) were then diluted into 200 μl of T buffer alone (C) or T buffer containing 100 μg/ml of each of trypsin (Sigma) and chymotrypsin (Sigma) fluoride (E) or containing the enzymes and 1% (v/v) Triton X-100 (T). After 90 min at 15°C, sodium dodecyl sulphate and phenylmethylsulphonyl fluoride were added to 1% (v/v) and 5 mM, respectively. Samples were immunoprecipitated with anti-haemoglobin and electrophoresed.

membranes, but not secreted from the oocytes. These results were consistent with the hypothesis that the deletion of amino acids 231 to 279 may impair the release of the completed mutant protein from the ER membrane. However, the release of OV_{Sau} from oocyte membrane vesicles after carbonate treatment (Fig. 6) indicates that the protein is *not* an integral part of the ER membrane, but is free in the lumen of secretory membranes.

Whilst no strong association seemed to exist between OV_{Sau} and membranes, it remained possible that the mutant protein could aggregate with itself or other proteins in the ER lumen. Aggregation has been cited as the main reason for the lack of movement to the cell surface of several vesicular stomatitis virus membrane proteins containing temperature-sensitive mutations in their lumenal domains (Leavitt *et al.*, 1977*a*,*b*; Rice & Strauss, 1982); however, in this paper, no sign of aggregation of the OV_{Sau} protein was detected (Fig. 5).

Another possibility is that the deletion of amino acids 231 to 279 of ovalbumin may directly or indirectly (*via* conformational perturbation) remove some hypothetical sorting sequence necessary for transport of the protein beyond the ER. It has been reported that mutation can block the movement of proteins to the cell surface. For example, Mosmann *et al.* (1979) have described a variant MOPC 315 immunoglobulin light chain that entered the ER but failed to be secreted from myeloma cells, probably due to conformational changes arising from a single amino acid change. This same phenotype was apparent when this protein was produced in *Xenopus* oocytes (Valle *et al.*, 1983).

Similarly, removal of the transmembrane "anchors" and C-terminal "tails" of various integral membrane proteins, including vesicular stomatitis virus G-protein (Rose & Bergman, 1983) or the Semliki Forest virus E_2 membrane protein (D. Cutler, personal communication), leads to retention of the deleted proteins within the ER, although truncation of influenza haemagglutinin results in complete secretion of the shortened proteins (Sveda *et al.*, 1981; Gething & Sambrook, 1982).

Thus, our data rule out firm attachment to membranes, aggregation, and abnormal glycosylation as reasons for the lack of complete secretion of the OV_{Sau} protein. The remaining possibilities are that OV_{Sau} remains within the ER lumen due to the loss of a specific signal that is either encoded by the deletion region or is masked by conformational changes resulting from this deletion, or that the mutant protein sticks to the inner surface of secretory membranes by virtue of a peripheral hydrophobic interaction that prevents or greatly retards protein movement out of the cell, but is disrupted by carbonate treatment of oocyte membrane vesicles. Nevertheless, despite the non-secretory phenotype of OV_{Sau}, the segregation of this protein within the ER lumen, as judged by its glycosylation and its release from membranes by carbonate extraction, argues against an essential role for the deleted sequence (residues 231 to 279) in the process of translocation of ovalbumin through the ER membrane.

The involvement of the extreme N terminus of ovalbumin in secretion was addressed by deleting the first eight amino acids of the protein, a manipulation that had no detectable effect on either the segregation or secretion of the mutant product (OV_{43}). It has been demonstrated that the removal of similar regions of

N-terminal, cleavable signal sequences prevents membrane-insertion of the protein (Talmadge *et al.*, 1981; Gething & Sambrook, 1982; Sekikawa & Lai, 1983), thus we conclude that the extreme N terminus of ovalbumin plays no role in the segregation or secretory process, indicating that the signal sequence of the protein is "internal", to the extent of at least eight amino acids. In this respect, ovalbumin resembles influenza neuraminidase, which appears to possess, near, but not at the N terminus, a hydrophobic region that acts as an uncleaved signal sequence; although in the case of neuraminidase, the hydrophobic domain seems to serve a dual function as both signal sequence and membrane anchor (Davis *et al.*, 1983; Markoff *et al.*, 1984).

When 20 N-terminal amino acids of ovalbumin were removed (OV_{His}), the resulting protein was segregated in oocyte membranes, but was not secreted into the culture medium. Like OV_{Sau}, the OV_{His} protein was heavily *N*-glycosylated, indicating its intralumenal location. Meek *et al.* (1982) have speculated that the role of amino acids 1 to 25 in ovalbumin is to participate in the formation of a hairpin loop with amino acids 25 to 45, and that this loop structure is the signal element for insertion into the ER membrane. Our finding that ovalbumin segregation can occur in the absence of amino acids 2 to 21, excludes this hypothesis. The absence of any secretion of OV_{His}, however, may implicate the deleted region in the secretory process.

As in the case of OV_{Sau}, abnormal glycosylation (discussed later) of the OV_{His} protein was ruled out as an explanation for its non-secreted phenotype. However, unlike OV_{Sau}, OV_{His} was not released from oocyte membrane vesicles by treatment with carbonate, indicating that OV_{His} remains anchored in the membranes, thus implicating the deleted amino acids (2 to 21) in the normal release of ovalbumin into the lumen of the ER. It must be remembered that, in the construction of this mutant, four amino acids, including two leucine residues, were introduced at the amino terminus of the resultant protein. We cannot exclude the possibility that the apparent association of the OV_{His} protein with oocyte membranes is an artefactual function of these added amino acids. With this qualification, we hypothesize that the OV_{His} mutant protein remains anchored in the ER membrane by the hydrophobic block of amino acids near the N terminus (amino acids 28 to 46). Since the OV_{43} protein, lacking amino acids 1 to 8, *was* secreted, we suggest that the normal role of the N-terminal amino acids in ovalbumin, specifically amino acids 9 to 21, is to displace the uncleaved, hydrophobic signal sequence from the membrane.

Hortin & Boime (1981) have reported that uncleaved pre-prolactin containing the threonine analogue, β-hydroxynorvaline, accumulated in the ER of isolated rat pituitaries, and was not secreted. The absence of processing by signal peptidase was attributed to the incorporation of β-hydroxynorvaline at position 29 in the pre-protein, adjacent to the signal cleavage site. Both sonication and low concentrations of detergent released the pre-protein from isolated microsomes, indicating that the uncleaved protein had been translocated completely through the ER membrane and was free in the intra-cisternal space. These results would seem to disprove the idea that signal cleavage is necessary for passage of a secretory protein out of the ER membrane although, in the absence of comparison

with the behaviour of a known membrane protein under these experimental conditions, the data presented by Hortin & Boime (1981) cannot unequivocally define the disposition of the pre-prolactin polypeptides within the ER. In addition, the possible incorporation of β-hydroxynorvaline at position 12 in the signal peptide, as well as at position 29, may affect the interaction of this region of the pre-protein with the ER membrane.

It should be pointed out that, in this paper, comparison with an appropriate integral membrane protein "control" was not possible in the case of the OV_{Sau} and OV_{His} mutants. It can only be said that, after carbonate treatment of oocyte membrane vesicles, the disposition of the OV_{Sau} protein, resembled that of wild type ovalbumin, which is known to be secreted from the oocytes. On the other hand, the OV_{His} protein, which lacked amino acids 2 to 21 but retained the hydrophobic region between amino acids 28 and 46, showed a much stronger association with oocyte membranes. Julius *et al.* (1984) have suggested that, in yeast, the initially uncleaved, hydrophobic N terminus of the pre-proα-mating factor may direct the correct secretion of the precursor by maintaining contact between the protein and organelle membranes until relatively late in the secretory pathway. It is possibly this sort of protein–membrane interaction that is perverted by the deletion of 21 amino acids from the N terminus of ovalbumin, although pre-proα-factor differs from ovalbumin in being extensively proteolytically processed just before release from the cell.

If amino acids 28 to 46 do constitute all or part of the ovalbumin signal sequence (see below), the deletion of this region should prevent translocation. We failed to detect any immunospecific proteins after injection of $pTK_2OV_{Nco}+$, a construct lacking information for amino acids 1 to 40, although the mutated DNA was transcribed with an efficiency comparable to that of the wild type construct, pTK_2OV+ (result not shown). The predicted 5′ leader sequences of the transcripts from these two constructs are identical, thus the expected translation efficiency of each species would be the same. Consequently, the most likely explanation for the absence of OV_{Nco} protein is that the protein is made, but rapidly degraded, most probably in the cytosol compartment of the oocyte. Consistent with this interpretation are the results of other workers, who note much reduced levels (Gething & Sambrook, 1982) or no protein at all (Davis *et al.*, 1983) when signal-minus mutant proteins were expressed in cultured cells. Stability in the oocyte cytoplasm presumably depends on the conformation of a particular protein as miscompartmented wild type ovalbumin and the OV_{Sal} mutant were clearly visible in cytoplasmic fractions, whereas the OV_{Nco} protein as well as a double deletion mutant lacking amino acids 20 to 145 and 231 to 279, or an N-terminally extended version of OV_{Sal} (from $pTK_2OV_{Sal}-$) were never detected (results not shown).

(b) *Synthesis of ovalbumin–globin fusion proteins*

Lingappa *et al.* (1984) have shown that the complete chimpanzee α-globin polypeptide can be translocated into the lumen of dog pancreatic vesicles *in vitro* if it is fused to the bacterial β-lactamase signal sequence. In this paper, we demonstrate that amino acids 9 to 41 or 22 to 41 of ovalbumin are sufficient to transfer this same globin polypeptide into the endoplasmic reticulum of oocytes. These results indicate that sufficient information for the insertion and translocation of a polypeptide resides within amino acids 22 to 41 of ovalbumin, although at this time we cannot rule out a contribution from either the four "new" amino acids introduced into the N terminus of OV_{His} globin (see Fig. 7(a)), or from valine 42 of ovalbumin, which is replaced by valine 2 of globin. We could not detect the secretion of either fusion product, but further information is required before the molecular basis of the non-secretory phenotype can be determined (for a detailed discussion, see above).

(c) *Fate of extended ovalbumins*

Analysis of the deletion mutants described here has demonstrated that the signal sequence of ovalbumin begins at least 21 amino acids in from the N terminus of the proteins. In the accompanying paper, we describe how ovalbumin containing an additional 21 amino acids at the N terminus is segregated and secreted, whereas ovalbumin containing a 51 amino acid extension remains in the cytosol. The primary structures of these extensions are completely different but neither contains extensive hydrophobic regions. We interpret these results as indicating that the presence of the 51 but not the 21 amino acid extension of the nascent chain of ovalbumin can interfere with the recognition of the signal sequence during translation. Interference of this kind is thought to explain why translocation of neither ovalbumin nor secretory proteins containing N-terminal signal sequences will occur unless microsomes are present soon after translation initiates, *in vitro* (Blobel & Dobberstein, 1975*b*). Thus, although the internal position of the signal sequence of ovalbumin suggests a looped mode of polypeptide insertion into membranes, there would seem to be a limit, *in vivo*, to the permitted length of the N-terminal portion of the loop.

(d) *Glycosylation of mutant ovalbumins*

We have commented above that OV_{Sau} and OV_{His} proteins are over-glycosylated with respect to wild type ovalbumin. The inhibition of glycosylation observed in the presence of tunicamycin indicates that the oligosaccharide side chains are N-linked. Ovalbumin has only two potential *N*-glycosylation sites, at asparagines 293 and 312, and *in vivo* only asparagine 293 becomes glycosylated (Huang *et al.*, 1970), although asparagine 312 in unfolded ovalbumin can be glycosylated by an *in vitro* glycosylation system (Pless & Lennarz, 1977; Glabe *et al.*, 1980). Under conditions where the tunicamycin treatment was only partially effective, two additional bands of higher mobility were prominent after OV_{Sau} and OV_{His} expression. This observation, together with the fact that OV_{His} and OV_{Sau} appear to be over-glycosylated, suggests that asparagine 312 is utilized in these two proteins, in addition to asparagine 293. In the case of OV_{His}, the presence of a small amount of the "intermediate" band in the absence of tunicamycin indicates that the second site is not always utilized. The cause of the over-glycosylation

remains uncertain, although it would probably result from the mutant proteins possessing a different conformation to the wild type protein. Alternatively, over-glycosylation may occur because the proteins are not secreted. We have demonstrated that a non-secretory mouse myeloma immunoglobulin light chain becomes *N*-glycosylated post-translationally in both oocytes and myeloma cells (Valle *et al.*, 1983), whereas the wild type light chain, which is secreted, is never glycosylated.

5. Conclusions

The behaviour of deletion mutants reveals that the eight extreme N-terminal amino acids of ovalbumin evidently play no role in protein secretion, whereas the deletion of amino acids 2 to 21 prevents the secretion of the shortened protein, possibly by preventing the completion but not the initiation of transmembrane translocation into the ER. Removal of amino acids 1 to 40 results in no detectable protein product in oocytes, and we speculate that this is a consequence of the miscompartmentation and subsequent degradation of the mutant protein in the cytosol. This miscompartmentation would presumably itself be a consequence of the loss of information specifying insertion into the ER, thereby implicating a region beginning between amino acids 22 and 40 as having signal function. This explanation is supported, and the carboxy-terminal limit of the signal region is defined, by the demonstration that amino acids 22 to 41 of ovalbumin are sufficient to specify transport of the cytosolic protein, α-globin, through the oocyte ER membrane, although we do not know whether it leaves the membrane (cf. OV_{His}). We therefore conclude that ovalbumin contains an internal signal sequence encompassed by amino acids 22 to 41, although amino acids 9 to 21 may be essential for the successful completion of trans-membrane translocation. This conclusion is in agreement with a suggestion first made by McReynolds *et al.* (1978), that the signal sequence of ovalbumin is likely to include the block of hydrophobic amino acids at position 28 to 46. Since we have not systematically scanned the ovalbumin polypeptide for more regions that can confer translocation ability on globin, it remains possible that the "physiological" signal sequence is more internally located. However, this is unlikely since the "signal" sequence identified in this paper is that nearest the N terminus, and would be expected to mediate in the initiation of translocation as soon as it emerges from the ribosome. Our results exclude the involvement of the previously described internal signal sequence (Lingappa *et al.*, 1979) in initiation of ovalbumin translocation; however, this region *may* have a role in intralumenal transport of the protein beyond the ER.

We have deduced the transmembrane orientation of the non-secreted mutant ovalbumins on the basis of their fractionation, glycosylation status, and response to carbonate extraction. Unfortunately, the resistance of the mutant ovalbumins to all but very severe protease treatments has precluded precise investigation of the disposition of the mutant proteins within oocyte membranes, although the localization of the globin fusion proteins was determined successfully by the use of standard protease-protection analysis. We are currently examining both the insertion of the OV_{His} mutant into dog pancreatic microsomes, and the behaviour of the OV_{Nco} protein in *in vitro* translation systems programmed with *in vitro*-synthesized mRNAs.

The authors are grateful to the Medical Research Council (U.K.) for support.

REFERENCES

Baty, D., Mercereau-Puijalon, O., Perrin, D., Kourilsky, P. & Lazdunski, C. (1981). *Gene*, **16**, 79–87.
Blobel, G. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 1496–1500.
Blobel, G. & Dobberstein, B. (1975*a*). *J. Cell Biol.* **67**, 835–851.
Blobel, G. & Dobberstein, B. (1975*b*). *J. Cell Biol.* **67**, 852–862.
Braell, W. A. & Lodish, H. F. (1982). *J. Biol. Chem.* **257**, 4578–4582.
Chan, L. & Bradley, W. A. (1982). In *Cellular Regulation of Secretion and Release* (Conn, P. M., ed.), pp. 301–321. Academic Press, New York.
Colman, A., Lane, C. D., Craig, R., Boulton, A., Mohun, T. & Morser, J. (1981). *Eur. J. Biochem.* **113**, 339–348.
Davis, A. R., Bos, T. J. & Nayak, D. B. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **80**, 3976–3980.
Engelman, D. M. & Steitz, T. A. (1981). *Cell*, **23**, 411–422.
Fujiki, Y., Hubbard, A. L., Fowler, S. & Lazorow, P. B. (1982). *J. Cell Biol.* **93**, 97–102.
Gething, M. J. & Sambrook, J. (1982). *Nature (London)*, **300**, 598–603.
Gilmore, P. & Blobel, G. (1983). *Cell*, **35**, 677–685.
Glabe, C. G., Hanover, J. A. & Lennarz, W. J. (1980). *J. Biol. Chem.* **255**, 9236–9242.
Hanover, J. A. & Lennarz, W. J. (1981). *Arch. Biochem. Biophys.* **211**, 1–19.
Hortin, G. & Boime, I. (1981). *J. Biol. Chem.* **256**, 1491–1494.
Huang, C. C., Mayer, H. E. Jr & Montgomery, R. (1970). *Carbohydr. Res.* **13**, 127–137.
Inouye, S., Wang, S., Sekiezawa, J., Halegowa, S. & Inouye, M. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 1004–1008.
Julius, D., Schekman, R. & Thorner, J. (1984). *Cell*, **36**, 309–318.
Krieg, P., Strachan, R., Wallis, E., Tabe, L. & Colman, A. (1984). *J. Mol. Biol.* **180**, 615–643.
Leavitt, R., Schlessinger, S. & Kornfeld, S. (1977*a*). *J. Virol.* **21**, 375–385.
Leavitt, R., Schlessinger, S. & Kornfeld, S. (1977*b*). *J. Biol. Chem.* **252**, 9018–9023.
Lingappa, V. R., Shields, D., Woo, S. L. C. & Blobel, G. (1978). *J. Cell Biol.* **79**, 567–572.
Lingappa, V. R., Lingappa, J. R. & Blobel, G. (1979). *Nature (London)*, **281**, 117–121.
Lingappa, V. R., Chaidex, J., Yost, C. S. & Hedgpeth, J. (1984). *Proc. Nat. Acad. Sci., U.S.A.* **81**, 456–460.
Maniatis, T., Frisch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
Markoff, L., Lin, B.-C., Sveda, M. M. & Lai, C.-J. (1984). *Mol. Cell Biol.* **4**, 8–16.
McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M. & Brownlee, G. G. (1978). *Nature (London)*, **273**, 723–728.
Meek, R. L., Walsh, K. A. & Palmiter, R. D. (1982). *J. Biol. Chem.* **257**, 12245–12251.
Meyer, D. & Dobberstein, B. (1983). *J. Cell Biol.* **87**, 498–502.
Moreno, F., Fowler, A. V., Hall, M., Silhavy, T. J., Zabin, I. & Schwart, M. (1980). *Nature (London)*, **286**, 356–359.
Mosmann, T. R., Baunial, R. & Williamson, A. R. (1979). *Eur. J. Immunol.* **9**, 511–516.
Palmiter, R. D., Gagnon, J. & Walsh, K. A. C. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 94–98.
Pless, D. D. & Lennarz, W. J. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 134–138.
Rice, C. M. & Strauss, J. H. (1982). *J. Mol. Biol.* **154**, 325–348.

Rose, J. K. & Bergman, J. E. (1983). *Cell*, **30**, 753–762.
Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982). *J. Cell Biol.* **92**, 1–22.
Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.
Sekikawa, K. & Lai, C.-J. (1983). *Proc. Nat. Acad. Sci., U.S.A.* **80**, 3563–3567.
Sveda, M. M., Markoff, L. J. & Lai, C.-J. (1981). *Cell*, **30**, 649–656.
Talmadge, K., Brosias, J. & Gilbert, W. (1981). *Nature (London)*, **294**, 176–178.
Tzack, J. & Lampen, J. (1975). *Biochem. Biophys. Res. Commun.* **65**, 248–257.
Valle, G., Besley, J., Williamson, A. R., Mosmann, T. R. & Colman, A. (1983). *Eur. J. Biochem.* **132**, 131–138.
von Heijne, G. & Blomberg, C. (1979). *Eur. J. Biochem.* **97**, 175–181.
Walter, P. & Blobel, G. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 7112–7116.
Wickner, W. (1980). *Science*, **210**, 861–868.
Zehavi-Willner, T. & Lane, C. D. (1977). *Cell*, **11**, 683–693.

Edited by W. Franke

INT