Analysis of the novel C-terminal secretion signal of the Escherichia coli Haemolysin protein

Thesis submitted for the degree of Doctor of Philosophy
to the University of Leicester

by

Brendan Kenny, B.A (Mod) Genetics
Trinity College Dublin

Submitted December 1990

Dept. of Genetics
University of Leicester

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

Analysis of the novel C-terminal secretion signal of the *Escherichia coli* Haemolysin protein.

Brendan Kenny, Dept. of Genetics, Leicester University, U.K.

Abstract

The release of a haemolytic toxin from some urinopathogenic strains of *E. coli* is unique in that it is the only reported polypeptide to be truly secreted from this Gram-negative organism. This secretion system comprises four genes on a contiguous, approximately 7.5Kb DNA fragment, encoding the toxin itself (HlyA, 107KD), two membrane localised export proteins (HlyB,D) and the cytoplasmic HlyC protein, whose only apparent function is to post-translationally activate the toxin. Recently, another unlinked gene, *tolC*, encoding a minor outer membrane protein has also been reported to be required for the export process together with HlyB,D. A novel feature of this system is the presence of a C-terminal targeting signal to direct HlyA from the cell.

In this study I show that the HlyA C-terminal targeting signal can be harnessed to secrete the majority of both the mammalian prochymosin and the *E. coli* cytoplasmic, LacZ, proteins in an HlyB,D dependent manner. I have also shown that the efficiency of secretion dramatically decreases when either the C-terminal domain is reduced in size, from 23KD to 4KD, or as the "passenger" domain increases in size. These results suggest that the HlyA signal domain is composed of sequences required for both efficiency of secretion and targeting and that the secretion process is inhibited by heterologous passenger domains, the effect increasing with size presumably due to the adoption of more stable folded conformations.

This study has also been concerned with the investigation into the nature of the novel *hlyA* targeting signal by deploying a series of *in vitro* mutagenesis methods (hydroxylamine, site directed and "saturation"), to introduce point mutations. This work has generated a bank of mutants, the analysis of which, has highlighted several residues essential for efficient secretion and also indicated a minimal region containing the signal motif. However, this information together with comparisons with other molecules carrying similar targeting signals have not yet identified a common signal motif.

Acknowledgements

I would like, first of all, to thank my supervisor, Prof I.B. Holland, not only for his advise and the freedom he allowed me during the course of my research work, but also for providing me with the opportunity to undertake this study.

Thanks must be given to those members of the Genetics department, both past and present (they know who they are), who helped make my stay in Leicester both a good social, sporting as well as educational experience.

I wish to express my gratitude to Dr Geoff Yarranton, for his help and encouragement during my visits to Celltech (Slough) and for introducing me to some very good scientific techniques. I also acknowledge SERC and Celltech Ltd (Slough) for their joint financial support of this project.

Finally, I wish to thank my parents for their support and for those refreshing breaks away from Leicester to rejuvenate my "batteries".

Contents

CHAPTER

1

General Introduction

1.1 Introduction:

The cell is a distinct entity due to the presence of an enveloping lipid bilayer membrane system separating the cellular components from the surrounding milieu. This "barrier" is common to all cells and is called the cytoplasmic membrane. In bacteria, this membrane is surrounded by a rigid layer of peptidoglycan, a heteropolymer of 2 different acetamidosugars and of from 4 to 8 different amino acid residues, which lends structural support and shape to the cell. Certain bacteria - Gram negative - also have a second enclosing membrane and the space between the two membrane systems containing the peptidoglycan layer is called the periplasmic space. Here I shall deal, primarily, with the structure of the cell envelope of Gram negative bacteria as this is most relevant to this thesis.

1.1.1 The cytoplasmic membrane:

The cytoplasmic membrane in all organisms is composed of a bilayer of amphipathic phospholipids with proteins both integrated and associated with it as described in the fluid mosaic model of Singer and Nicholson (1972). The membrane is an asymmetric bilayer of the constituent phospholipid forms (Bretscher, 1972) mainly phosphatidylglycerol, its amino acyl derivatives and phosphatidylethanolamine, with a flip-flop mechanism proposed for lipid transmembrane migration (Kornberg and McConnell, 1971). At least 120 proteins have been resolved by 2D electrophoresis to be associated with this membrane (Sato *et al.*, 1977). These are involved mostly with cellular respiration, ATP synthesis and transport of molecules and proteins (Cronan *et al.*,1987), with an asymmetric location for some of these proteins within the membrane having been described (Owen and Salton, 1975). There is some evidence for the existence (if only transitory) of non-bilayer lipid domains (Burnell *et al.*, 1980; de Kruiff *et al.*, 1980), which could be important in such mechanisms as lipid flip-flop and protein translocation.

In relation to protein translocation evidence for a role for the anionic phospholipid (phosphatidylglycerol, the major acidic phospholipid in *E. coli*) has been indicated by using mutants defective in its synthesis (de Vrie *et al.*, 1988). Also, *in vitro* experiments investigating the interaction of signal peptides with membrane lipids suggests that a negatively-charged lipid-specific insertion of the signal peptide into the membrane, results in changes in both signal peptide structure and lipid organisation (Batenburg *et al.*, 1988) suggesting a possible role in the translocation process.

1.1.2 Periplasmic Space:

The rigidity and shape of bacterial cells is derived from their enveloping peptidoglycan (murein) layer. In Gram negative bacteria this structure is sandwiched between the cytoplasmic and the additional outer membrane in what is termed the periplasmic space. Cryoelectron microscopy has shown the periplasmic space to be of a uniform thickness of approximately 7.5nm (as are both the cytoplasmic and outer membranes), but appears to be composed of two distinct layers (Oliver, 1987). The upper (outer) region of the periplasm (approximately 2nm thick) is associated with the tightly cross-linked peptidoglycan layer, whereas the lower region contains a less cross-linked peptidoglycan structure. The network spacing of the peptidoglycan molecules, together with the rest of the periplasmic space are thought to have a density similar to that of the cytoplasm (helping to buffer the cell against osmotic/ionic changes) facilitating, perhaps, the diffusion of solutes and proteins. This dynamic structure has been referred to as the periplasmic gel (Hobot *et al.*, 1984), and is thought to contain at least 100 distinct proteins involved mostly in nutrient uptake and catabolism (Pugsley and Schwartz, 1985).

1.1.3 The Outer Membrane:

An outer membrane is present in Gram negative but not in Gram positive bacteria and as the outermost barrier has an important role in the interaction with the environment. It is composed of phospholipids, lipopolysaccharides and proteins, being physically linked to the peptioglycan layer by the Braun lipoprotein (Braun, 1975). Lipopolysaccharide is unique to the outer membrane, and is found exclusively in the outer leaflet of the membrane. Lipopolysaccharides (LPS) are composed of a proximal hydrophobic lipid A region, a core polysaccaride and a distal hydrophilic O antigen polysaccaride region. The lipid A moiety forms the outer leaflet of the membrane with phospholipids forming the inner leaflet. The LPS molecules have strong lateral interactions lowering the permeability to solutes and acting as a further barrier. Mutants in the assembly of LPS are lethal, indicating its important role in growth and viability, and may be involved in the assembly of a functional outer membrane. Regarding the biogenesis of LPS there is evidence that in Salmonella the newly synthesised LPS molecules are translocated from the cytoplasm to the outer membrane through zones of adhesion between inner and outer membranes (Bayer like junctions; Bayer 1979) prior to lateral diffusion (Muhlradt *et al.*, 1973).

The outer membrane of Gram negative bacteria is dominated by a few major proteins (Sato *et al.*, 1977) chiefly involved in transport or having a structural role. The Braun lipoprotein is one of these major proteins and as indicated above links the outer membrane to the peptidoglycan layer (Braun, 1975). Other proteins (see Lugtenberg and van Alphen, 1983) include the general porin molecules (such as OmpF/C and PhoE) which allow the diffusion of many small molecules, and porins involved in the import of specific molecules such as FhuA (Fe^{++} ferrichrome uptake) and LamB (maltose, maltodextrin uptake).

Finally, additional surface layers (protein matrices, capsules) and surface appendages (fimbriae, pili and flagella) may also be present in both Gram positive and negative organisms, although I shall not describe them as they are not pertinent to this discussion.

1.1.4 Extracellular environment:

The environment in which bacteria usually find themselves is often a medium low in nutrient concentration and other essential requirements. For this reason cells often secrete proteins or other molecules to aid the recovery and processing of substrates for entry across the membrane system as well as factors required for virulence. In Gram negative bacteria this release is inhibited by the outer protective layer and indeed, to date, relatively few proteins are known to be specifically secreted. These organisms can, however, export a large number of proteins to the periplasmic space where they are trapped. In contrast, in Gram positive organisms where no other barriers exist, passage of proteins across the cytoplasmic membrane effectively represents secretion. However, in this thesis I shall refer to the translocation of proteins across the inner, cytoplasmic membrane of Gram negative bacteria as export, whereas the translocation of proteins across both inner and outer membranes, resulting in specific release, will be designated secretion. Before I describe some of the mechanisms elucidated so far, for secretion of polypeptides across the two membrane systems of Gram negative bacteria, it is first relevant to describe the general mechanism employed for the translocation of most exported proteins across the ubiquitous primary barrier, the cytoplasmic membrane, in bacteria or its equivalent for this purpose the endoplasmic reticulum in eukaryotes.

1.2 Protein translocation signals:

1.2.1 Existence of N-terminal targeting signals for exported proteins:

A question which has been addressed vigorously over the last few decades is how does the cellular machinery determine which proteins are to be translocated, and when identified, to which compartment. Although it is technically easier to approach these questions by studying the prokaryotic system, due to the limited compartmentalisation of such organisms, one of the first insights into the problem was achieved using an eukaryotic based *in vitro* system. Milstein *et al.*, (1972) using such systems inferred the presence of a transient N-terminal extension found only on the pre-exported protein being investigated (immunoglobulin light chain), the existence of which was later confirmed by partial amino acid sequencing (Schechter *et al.*, 1975). The idea of specific "export" signals had previously been inferred from observations in eukaryotic cells that secreted proteins were almost exclusively synthesised on membrane associated ribosomes, whereas cytoplasmic proteins were associated with free ribosomes. Earlier investigations into the export process, again using *in vitro* systems, revealed that secretory proteins were vectorially exported into microsomes during synthesis indicating the existence of specific translocation sites (Redman and Sabatini, 1966; Blobel and Dobberstein, 1975). Moreover, the removal of the transient extension signal peptide was observed as being dependent upon the presence of a membrane bound component.

Subsequently, the translocation of most prokaryotic exported proteins was shown, using in this case *in vivo* experiments (see below), to be dependent on the presence of a similar, usually transient, N-terminal signal peptide. As in eukaryotic cells, membrane associated proteins were identified which removed the signal peptide during the translocation process. Thus, signal peptidase I (Wolfe *et al.*, 1982; 1983) was required for the cleavage of most exported proteins while signal peptidase II (Yu *et al.*, 1984) was required for those proteins carrying lipid modifications at their cleavage sites, for example, lipoprotein (Tokunaga *et al.*, 1982).

1.2.2 Analysis of the prokaryotic signal sequence using export mutants:

The analysis of export defective mutants has provided the basis in bacteria for the analysis of targeting signals, the identification of the cellular components involved in the translocation process and give an insight into the molecular mechanism of the translocation process. The

advent of recombinant DNA technology provided the means to generate defective export proteins, the analysis of which helped to identify the particular residues/features required for protein targeting and translocation. The organism chosen for the *in vivo* studies was *E. coli* as it was one of the best characterised organisms also being relatively easy to manipulate. The most successful procedure adopted for isolating targeting mutants was the generation of fusions of the N-terminal region of the exported protein of interest to the *E. coli* cytoplasmic protein beta-galactosidase (LacZ). The proteins chiefly used for the construction of such fusions were LamB (lambda receptor, outer membrane protein), MalE and PhoA (periplasmic maltose binding protein and alkaline phosphatase) (Bassford and Beckwith, 1979; Bedouelle *et al.*, 1980; Emr *et al.*, 1978; Emr *et al.*, 1980b; Emr and Silhavy, 1980 and Michaelis *et al.*, 1983). These hybrid proteins were generally found to be incapable of being exported, apparently due to the interaction with and blocking of the translocation machinery by the LacZ moiety, indicated by the accumulation of the precursor form of other exported proteins. This characteristic was taken to indicate the existence of a finite number of discrete export channels (Emr *et al.*, 1980a). The activity of the LacZ moiety of these fusions was in some cases found to be reduced, presumably due to interaction with the cytoplasmic membrane and interference with the conformation or ability of LacZ to tetramerise. High level expression of such hybrid proteins resulted in cell death demonstrating the essential nature of the process being inhibited. These phenotypes were exploited in order to isolate export defective mutants by selecting cells that survived high level expression of these hybrid molecules and or exhibited higher levels of LacZ activity. These latter properties indicated the cytoplasmic location of the hybrid, presumably due to its inability to be recognized now by the export machinery. The selection for the *lacZ*$^+$ phenotype also ensured that the survival of the cells was not due purely to non synthesis of the hybrid protein (Bassford and Beckwith, 1979; Bedouelle *et al.*, 1980; Emr *et al.*, 1978; Emr *et al.*, 1980b; Emr and Silhavy 1980 and Michaelis *et al.*, 1983). This type of study led to the identification of a large number of export defective mutants, the analysis of which revealed the specific location of the N-terminal signal sequence, highlighting the important residues and structures involved.

1.2.3 Characterisation of the bacterial signal sequence:

Comparison of the primary amino acid sequence comprising the N-terminal extension region of exported proteins showed no obvious homologies except for N-terminal location and overall hydrophobicity (Watson, 1984; von Heijne, 1986a). A preponderance for small neutral residues was also identified at the signal peptidase cleavage site (von Heijne, 1983), suggesting a degenerate recognition pattern presumably involving some loose secondary/tertiary structure. Characterisation of export defective mutants, together with the results of site specific alterations, revealed that the majority of export defective mutants involved alteration of the hydrophobic nature of the signal peptide. However, the analysis of N-terminal mutants led to the division of the targeting region into 3 structurally and functionally distinct regions (von Heijne, 1986a). Firstly, there is a positively charged N-terminal region, followed by a 7-15 residue unbroken hydrophobic core and lastly a 5-6 more polar C-terminal region containing the site for signal peptide cleavage by signal peptidase.

The role of the charged N-terminal region was investigated by making specific changes in the lipoprotein signal sequence. Thus, studies by Inouye *et al.*, (1982) and Vlasuk *et al.*, (1983) demonstrated that changing the overall charge of the N-terminus of the *E. coli* lipoprotein from 2^+ to 1^+ or 0 had little effect, whereas a charge of 1^- or 2^- severely reduced the rate of synthesis and processing. This indicated an important role in the efficiency of translocation, for example, by accelerating the interaction of the export protein with the negatively charged membrane, as supported by later work carried out by Batenburg *et al.*, (1988) and deVrie *et al.*, (1988) -see section 1.1.1. The importance of the positioning of the signal sequence was probed by the addition of an extra N-terminal extension onto preproinsulin (Talmadge *et al.*, 1981). Talmadge and co-workers demonstrated that the N-terminal addition of 18 amino acids onto preproinsulin had no observable effect on export, whereas the addition of 51 amino acids did prevent secretion in Xenopus oocytes, implying that with certain limitations, signal sequences need not be amino terminal. In contrast, chicken ovalbumin is reported to have an internal signal sequence located somewhere between residue 22 and 41, which is not processed (Tabe *et al.*, 1984).

The importance of the hydrophobic core of the signal sequence for export was indicated by the observation that the majority of selected export defective mutants were altered in this region. In the case of the export of

such proteins as LamB (Emr and Silhavy, 1982; Emr *et al.*, 1980b) and the periplasmic, maltose binding protein (Bedouelle *et al.*, 1980), substitutions to charged residues or deletions of a subset of hydrophobic residues were obtained. The results indicated that a certain critical subset of hydrophobic residues may be functionally more important for signal recognition, for example, through the adoption of a specific alpha-helix structure, than adjacent hydrophobic residues. The requirement for an alpha-helix structure was also supported by the isolation of intragenic suppressors of export defective mutants of LamB which were predicted to re-establish the putative alpha-helical structure (Emr and Silhavy, 1982). Such structures might be capable of mediating interaction of the signal peptide with membrane lipid or to facilitate protein-protein interactions. Studies with wild-type and mutant signal peptides have demonstrated a specific interaction with the membrane which results in the formation of non-lipid bilayer membrane structures. The hydrophobic nature of the signal peptide is apparently essential for this process, with the charged region being required only for increased efficiency of translocation (Killian *et al.*, 1990). These non-bilayer structures have been postulated to facilitate the translocation of export proteins, with other protein components possibly being required to stabilise or regulate these structures (Killian *et al.*, 1990). However, whether these are formed by signal peptides and/or required for translocation *in vivo* is not known.

Mutations found to affect processing of exported proteins apparently do not affect normal translocation, assembly or activity of the protein (Kuhn and Wickner, 1985; Randall *et al.*, 1987). Thus, it is possible that the role of this cleavage in such cases, although not essential for translocation, might be to aid stability by allowing more rapid assumption of the native conformation of the protein following transport.

1.2.4 Protein compartmentalisation and signal sequences in bacteria:

Several studies in bacteria have been concerned with the possibility that the N-terminal signal sequence is sufficient only for translocation across the cytoplasmic membrane and that other signals might be required for final localisation in the *E. coli* envelope. *In vitro* techniques have been used to exchange the N-terminal signal sequences of either beta-lactamase, a periplasmic protein (Tommassen and Lutenberg, 1984) or the inner membrane protein, Penicillin binding protein PBP5 (Jackson *et al.*, 1985), for that of outer membrane proteins. The results showed that the proteins were still apparently localised as normal to the outer

membrane indicating that other topological signals were present in the mature protein. Examination of the location of other hybrid proteins, consisting of beta-galactosidase fused to the C-terminus of varying portions of maltose binding protein and alkaline phosphatase, indicated that increasing the portion of the export protein domain eventually resulted in export of some of the hybrid as far as the outer membrane (Bassford *et al.*, 1979; Michaelis *et al.*, 1983) suggesting the presence of distinct localisation signals. However, doubts as to the validity of hybrid localisation due to problems with fractionation techniques of such aberrant proteins were suggested and have now been substantiated (Voorhout *et al.*, 1988). On the other hand analysis of deletions of OmpA (Freudl *et al.*, 1990) and PhoE (Bosch *et al.*, 1986; 1988) have indicated that it is the conformation of the mature protein and not some specific sorting signal that specifies their localisation in either the periplasm or outer membranes once translocated (also see Randall *et al.*, 1987). In the case of inner membrane proteins it is thought that they are localised by the presence of a hydrophobic "anchoring" sequence or in the case of PBP5 an amphiphilic helix (Jackson and Pratt, 1987)

Experiments with some signal peptides from Gram positive organisms (often considerably longer than Gram negative signals) promoted translocation in *E. coli* and visa versa, while certain eukaryotic secreted proteins, such as rat preproinsulin (Talmadge *et al.*, 1980a, 1980b), are exported and correctly processed in *E. coli*. The fact that eukaryotic signal sequences directing proteins to the endoplasmic reticulum can be recognized in bacteria is not too surprising, as their overall structure and nature of their cleavage sites is quiet similar. The only noticeably different feature is the existence, on average, of only one positive charge at the N-terminus while bacteria often have two (von Heijne, 1984). However, bacterial signal sequences when fused to eukaryotic exported proteins are relatively inefficient in effecting export in *E. coli* (see Beckwith and Ferro-Novick, 1986). These results suggest that in bacteria, where proteins are exported only across the cytoplasmic membrane, a single, common, signal motif might be utilised, whereas in eukaryotes where different proteins have to be directed to different compartments (such as the endoplasmic reticulum, mitochondria and peroxisomes), other signals may be required to aid targeting and translocation.

1.3 Models for protein translocation:

Although it is clear that specific targeting signals ensure the delivery of export proteins to the site of translocation, how do the proteins (containing both hydrophobic and hydrophilic residues) cross the membrane barrier ?. Initial evidence supported the idea that other protein components were required for translocation across this barrier. However, several workers have proposed on theoretical grounds that translocation could proceed directly across the phospholipid bilayer, without the requirement for any other proteins.

1.3.1 Models for protein mediated translocation pathways:

i) The Signal Hypothesis:

From the work of Milstein *et al.*, (1972) and other groups who re-affirmed the existence of transient signal peptides, it was envisaged that the signal peptide targeted the export protein to a specific export channel, presumably composed of other proteins, which would secure its specific translocation. This work led to the elaboration of the signal hypothesis which proposed that the signal peptide extension interacts with a complex protein pore, within the membrane, upon emergence from the ribosome, and continued synthesis results in the linear translocation of the polypeptide, followed by removal of the signal peptide (Blobel and Sabatini, 1971; Blobel and Dobberstein, 1975). Indeed, proteins required for the translocation process have been isolated from eukaryotic cells. A major component, the signal recognition particle (SRP - now identified as a complex composed of six different polypeptides and a 7S RNA molecule) was found to specifically interact with secretory protein synthesizing polysomes, mediating interaction with the membrane, but inhibiting translation in a membrane free environment (Walter and Blobel, 1981; and 1982). The second element to be identified was membrane bound and termed the SRP-receptor or docking protein (Gilmore *et al.*, 1982; Meyer *et al.*, 1982), which appeared to relieve the arrest of translation induced by SRP. A revised signal hypothesis now envisaged the SRP complex binding the signal peptide, arresting translation until the new complex interacts with the docking protein (releasing the SRP complex), allowing the polypeptide to be vectorially translocated through the membrane during its further synthesis. This model necessitates that translocation proceeds co-translationally as was initially thought for the eukaryotic system. However, more recent evidence suggests that translocation can indeed occur post-translationally (Perara *et al.*, 1986; Rothblatt and Meyer, 1986). Moreover, it was suggested that translation arrest, coupled with subsequent co-translational export, might be an artifact due to

disulphide bond formation in the oxidised environment of *in vitro* systems (Maher and Singer, 1986) although this remains controversial. Export in bacterial systems may also occur post-translationally (Wickner, 1979; Date and Wickner, 1981; Koshland and Botstein, 1982; Randall, 1983). This led to the suggestion that translocation was dependent upon the presence of cytoplasmic factors (Collier *et al.*, 1988) required to maintain the export proteins in a translocational competent state (see Randall *et al.*, 1987). Nevertheless, considerable evidence indicates that transport might frequently occur co-translationally *in vivo* (Smith *et al.*, 1977; Josefesson and Randall, 1981). In addition, the existence of a similar general proteinaceous export pathway in bacteria, including membrane proteins, was supported by the ability to select and isolate extragenic suppressor mutations for export proteins carrying defective signal sequences (see Section 1.4.3).

ii) Singer Model of Translocation:

This model proposed by Singer *et al.*, (1987) is the only model to confront the possibility of protein channels in the translocation process. This model was proposed to explain the transfer of hydrophilic polypeptides across membranes which would involve thermodynamic high energy costs in inserting ionic residues in the non-ionic membrane environment. Thus, this model insists on the requirement that the ionic groups of polypeptides must largely remain in contact with water throughout the process of translocation minimising the energy costs. Indeed, the requirement for polypeptides to remain in contact with water during transfer has been supported by experiments by Gilmore and Blobel (1985). The Singer model proposed the generation of a transmembrane aqueous channel due to the aggregration of a number of homologous but not identical integral membrane proteins (translocator proteins, TPs) allowing the ionic residues of the export protein to face the aqueous channel and the nonpolar residues to remain away from the aqueous environment within the interface of the translocator proteins during translocation. It is proposed that polypeptides are translocated domain by domain and that the initial energy of translocation could be obtained from energy-induced conformational changes in the translocator subunit complex, aided by the energy associated with the folding of the translocated domains.

1.3.2 Models for protein independent-translocation:

i) Membrane Trigger Hypothesis:

Bretscher (1973), puzzled by the problem of how many proteins are able to insert into the cytoplasmic membrane (most lacking classical signal sequences), proposed that the nature of the protein itself would ensure its correct assembly into the membrane after diffusion to and insertion into the cytoplasmic membrane. This model was extended by Wickner (1979) to explain the observed post-translational insertion of some export proteins into the cytoplasmic membrane (such as M13 coat protein). This was developed further as the Membrane Trigger hypothesis (Wickner, 1979), in order to explain complete translocation. The role of the signal peptide was envisaged as maintaining the protein in the correct conformation compatible for a particular environment, exposing hydrophilic residues in aqueous and hydrophobic residues in lipid surroundings. The hydrophobic nature of the signal peptide and, perhaps, other regions of the mature protein would therefore interact with the membrane element (protein or lipid) leading to the folding of the protein into a conformation that exposes hydrophobic residues to the lipid bilayer. This model proposed that newly synthesised or even nascent incomplete polypeptides can associate with and insert into the membrane, and removal of the signal peptide renders the insertion irreversible. The model was principally proposed in order to explain post-translational assembly into the membrane but complete translocation could be explained through the adoption of a conformation, following transfer through the membrane, compatible with the periplasmic environment, induced again by the leader peptide, with signal cleavage finally rendering the transfer irreversible.

ii) The Direct Transfer Model:

The direct transfer model (von Heijne and Blomberg, 1979) proposes that the hydrophobic signal peptide of a nascent polypeptide adopts an alpha-helical conformation which interacts with the membrane bilayer allowing the ribosomes also to attach to the membrane surface. Subsequent polypeptide chain elongation then drives the transfer of nascent polypeptide through the membrane. The now numerous demonstrations that at least under certain conditions protein translocation can be uncoupled from peptide bond formation renders this model untenable (Randall *et al.*, 1987).

iii) The Helical Hairpin Model:

The helical hairpin model of Engelman and Steitz (1981) envisaged that the signal sequence forms an alpha-helical hairpin which facilitates insertion into and passage across the membrane by its hydrophobicity and the exploitation of favourable free energies. Continuous synthesis of the nascent polypeptide, it is suggested, is only required to prevent the folding of domains. Integral membrane proteins in this model are proposed to be formed due to the presence of non-polar residues preventing complete translocation.

iv) Loop Model:

The loop model of Inouye and colleagues (Inouye *et al.*, 1982), proposed that the positive character of the N-terminus of the signal peptide that interacts with the negatively charged lipids of the cytoplasmic membrane, with peptide elongation leading to the insertion of the hydrophobic portion of the leader peptide, generating a loop structure. Exposure of this loop at the periplasmic side of the membrane results in peptide cleavage, elongation of the polypeptide leading to the threading of the chain through the membrane. Indeed, altering the relative positive charge at the extreme N-terminus of export proteins does affect the rate of synthesis and/or efficiency of export (Inouye *et al.*, 1982; and Vlasuk *et al.*, 1983) suggesting an important role in targeting. This model in complete contrast to that of Blobel and Dobberstein predicting that the cleaved signal peptide should be found in the membrane. This indeed was observed in the case of lipoprotein (Hussain *et al.*, 1982). Although this model also depends upon the coupling of peptide bond formation and translocation, which is not necessarily the case, it is generally accepted that the signal peptide does in some way insert finally into the membrane in the proposed loop structure in the case of many export proteins.

Evidence may be presented for and against all of these models (for review see Duffaud *et al.*, 1985) but the isolation of specific protein components required for the translocation of most exported proteins in both pro- and eu-karyotes now suggests that the vast majority of exported proteins do carry an N-terminal signal sequence and that this mediates specific translocation via a general protein- dependent export pathway. However, the existence of alternative pathways/ mechanisms for a subset of proteins in some cells or organisms cannot be dismissed and indeed have been shown to exist for the secretion of atleast some proteins from *E. coli* (haemolysin and colicin, see Section 1.5.3) and other bacteria. Some of

the evidence for the existence of accessory protein components required for the translocation process is described below.

1.4 Protein components involved in the general translocation pathway:

1.4.1 Eukaryotic proteinaceous translocation system:

As mentioned above the eukaryotic *in vitro* translocation systems have been used successfully to investigate the nature of the protein translocation process across the endoplasmic reticulum and has led to the identification and purification of some of the cellular components involved. So far, a signal recognition particle (SRP) has been identified, which is composed of a 7S RNA particle and at least 6 different polypeptides. This complex apparently binds to the signal sequence as it emerges from the ribosome, leading to the arrest of the synthesis of the nascent export protein chain, in the wheat germ although not in the HeLa or reticulocyte *in vitro* system (Meyer, 1985). One of the polypeptide components of the SRP, a 54KD protein, binds to the signal sequence as it emerges from the ribosome (Krieg *et al.*, 1986). The gene encoding this 54KD polypeptide has recently been sequenced (Romisch *et al.*, 1989; Bernstein *et al.*, 1989), revealing significant homology to a eukaryotic docking protein (see below) and to two *E. coli* proteins of unknown function. One of the *E. coli* proteins, FtsY, is encoded by a gene of a cell division operon while the other is encoded by an open reading frame upstream of the *trmD* operon. Both proteins contain a conserved consensus GTP binding site. The SRP-nascent polypeptide complex then binds to a 72KD membrane protein, the SRP-receptor or docking protein (Meyer *et al.*, 1982; Gilmore *et al.*, 1982) releasing the SRP (and any translational arrest), with the signal sequence then interacting with another membrane protein. This approximately 35KD integral membrane glycoprotein, the signal sequence receptor (SSR), was identified by cross-linking studies (Wiedmann *et al.*, 1987). The next stages in translocation are as yet unknown except that most proteins in such *in vitro* systems are apparently simultaneously elongated and translocated through the membrane site, the transient signal peptide being cleaved by a membrane bound peptidase.

1.4.2 Protein components implicated in translocation across mitochondrial membranes:

Mitochondria (and chloroplasts) are considered to be endosymbiotic eukaryotic organelles of prokaryotic origin (Schwartz and Dayhoff, 1978), and like Gram negative bacteria have two enveloping membrane systems.

Most of the protein components are encoded by the 'host' cellular genome and have to be **import**ed into these organelles. Most of the work to date on the transport processes of such organelles has been carried out in mitochondria. Consequently, I shall limit myself to this example, although it is thought that similar mechanisms are involved in chloroplast protein import. Protein import into mitochondria also involves the presence of (usually transient) N-terminal targeting signals. These differ from bacterial and endoplasmic reticulum signals in that they are usually 10 to 70 residues in length, rich in positively charged and hydroxylated residues and generally lack acidic amino acids (Viebrook *et al.*, 1982; Kaput *et al.*, 1982; Hartl and Neupert, 1990). These signals have been proposed to adopt amphiphilic alpha-helices (von Heijne, 1986b) or beta-sheets (conceivably essential for protein-protein or protein-lipid interaction) and have been successfully used to direct non-mitochondrial passenger proteins to the matrix (Hurt *et al.*, 1984; Horwich *et al.*, 1985; Hurt and van Loon, 1986). A variety of techniques have been successfully used to identify translocation receptor proteins, including protease studies, identification of proteinaceous components interacting with signal sequences using crosslinking agents and anti-idiotype antibodies (Pfaller *et al.*, 1988; Sollner *et al.*, 1989; Vestweber *et al.*, 1989; Pain *et al.*, 1990; Hartl and Neupert; 1990). Indeed specific polypeptides of this kind are distributed over the whole surface of the outer membrane, but appear to be enriched at sites of contact between the inner and outer membranes (Pain *et al.*, 1990; Hines *et al.*, 1990; Hartl and Neupert; 1990). Such studies in fact indicate that a number of different receptors may be involved in the import of different subsets of precursor proteins (Pfaller *et al.*, 1988; Sollner *et al.*, 1989; Pain *et al.*, 1990; Hines *et al.*, 1990). Protein import is post-translational and is thought to occur at these contact sites since particular chimaeric proteins, blocked in the process of importation, have been found to span both membranes, the presequence having been cleaved by the matrix located peptidase (Schleyer and Neupert, 1985; Schwaiger *et al.*, 1987; Rassow *et al.*, 1989; Vestweber and Schatz, 1988a). The import of most proteins seems to require ATP or GTP (possibly to maintain or acquire a loosely folded conformation - see Section 1.4.5) as well as an electrochemical potential for import across the inner membrane (Pfanner *et al.*, 1987). Protein **export** also occurs in mitochondria for some imported proteins destined for the intermembrane space. These proteins carry longer targeting signals, apparently bipartate in structure, the N-terminal portion being reminiscent of a mitochondrial presequence (see Hartl and Neupert, 1990), the C-terminal portion resembling the features of a bacterial signal

sequence (von Heijne, 1983). These proteins are apparently imported to the matrix where they are cleaved by the matrix peptidase and then redirected across the inner membrane by the bacterial-like signal peptide which is subsequently cleaved by a membrane associated peptidase (outer surface of inner membrane). This export process is thought to require the hydrolysis of ATP, postulated to be involved in the release of a heat shock matrix protein (structurally and functionally related to the *E. coli* GroEL protein), which combines with the export protein probably to maintain the latter in a conformationally competent state (see Hartl and Neupert, 1990 for review).

1.4.3 Identification of proteinaceous components involved in the general protein export system of *E. coli*:

Several protein components required for the export of proteins from bacteria have been identified, mainly using *in vivo* selection procedures in *E. coli*. For example, genes encoding a number of these proteins were identified indirectly by exploiting the LacZ$^-$ phenotype of cells expressing hybrid proteins composed of the N-terminal signal region of an export protein fused to beta-galactosidase. Such fusions appear to interact with and jam the general export pathway. Lac$^+$ (ability to utilise lactose) temperature sensitive (for growth) mutants were selected, in which the hybrid protein is presumably cytoplasmic in location and can no longer engage the translocation machinery. Some of these mutants were found to exhibit a pleiotropic export defective phenotype, that is they accumulated envelope protein precursors at 42^{o}C. This observed pleiotropic phenotype suggested the existence and alteration of a protein component involved in the export of many export proteins and the presence of a limited number of export channels. The first genes to be identified by such a method were designated *secA* (Oliver and Beckwith, 1981) and *secB* (Kumamoto and Beckwith, 1983). In addition, *prl* mutants were also isolated, for example *prlA*/*secY* were selected as being able to suppress a LamB signal sequence deletion mutant (Emr *et al.*, 1981; Emr and Bassford, 1982; Shultz *et al.*, 1982).

1.4.4 *E. coli* protein export components:

i) *secA* (*prlD*)

The defect exhibited by the *secA*^{ts} mutants was the accumulation of the precursors of at least 4 of the major outer membrane and most of the periplasmic proteins at the non-permissive temperature (Oliver and Beckwith, 1981; Liss and Oliver, 1986). The mutation also blocked the assembly of the cytoplasmic membrane protein, leader peptidase (Wolfe *et*

al., 1985), indicating an essential role in protein export. This effect on the export of all types of envelope proteins implied the existence of a general export pathway. Other studies have shown SecA to be a 102KD peripheral inner membrane protein and by the use of a *secA-lacZ* gene fusion SecA is estimated to be present at approximately 500-1000 copies per cell. The level of SecA was also shown to be regulated according to the requirements for SecA in protein export (Oliver and Beckwith, 1982). Reduced levels of synthesis of some export proteins in a *secA*(amber) strain (Oliver and Beckwith, 1982; Strauch *et al.*, 1986) also suggested that SecA might play a role similar to the eukaryotic SRP, coupling translation to transport. However, this effect was not observed after the addition of cAMP or growth in rich medium, suggesting that the lack of SecA was inducing catabolite repression of the synthesis of certain outer membrane proteins rather than a specific effect on translation (Strauch *et al.*, 1986; Baker *et al.*, 1987).

Selection for extragenic suppressor mutations of an MBP export defective protein also generated mutations (*prlD*; Bankaitis and Bassford, 1985) in the *secA* gene. *prlD* indeed seemed to increase the level of export of several LamB signal sequence mutants suggesting a strong interaction with the signal peptide of these two proteins. The observation that some *prlD* mutations had no major effect on wild-type protein export, could be construed to indicate that these mutations do not alter the normal major export pathway. However, *prlD* mediated suppression was remarkably allele specific suggesting a direct interaction with the signal peptide. Also, the observation that certain combinations of *prlA* (see below) and *prlD* double mutants had severe effects on cell growth, indicated that they were indeed components of the export pathway, with interactions between these proteins and the signal sequence.

Indeed, biochemical analyses have now demonstrated a functional role for SecA in protein export *in vitro* and show that the protein possesses ATPase activity in the presence of inner membrane vesicles and an envelope protein precursor such as OmpA (Cabelli *et al.*, 1988; Cunningham *et al.*, 1989; Lill *et al.*, 1989).

ii) *secB*

In contrast to *secA*, *secB* mutants only showed specific effects on a subclass of envelope proteins. Thus, the outer membrane protein OmpF and periplasmic maltose binding protein (MBP) were inhibited for export, while the periplasmic PhoA and ribose-binding proteins appear to be

translocated in a *secB* independent manner (Kumamoto and Beckwith, 1983). Indeed, *secB* null mutations were not lethal, indicating the non-essential nature of this gene product. Recent work (for review see Eilers and Schatz, 1988) has indicated that proteins for export have to be maintained in a translocation competent, partially unfolded conformation, and SecB has been implicated in preventing, for example, MBP from adopting a structure incompatible with translocation (Collier *et al.*, 1988). Other export proteins, apparently *secB* independent, may require other antifolding proteins (see section 1.4.5 below), such as a soluble factor described by Muller and Blobel (1984). Conversely, some export proteins might not adopt "tightly folded" tertiary structures in the cytoplasmic environment therefore not necessitating the use of anti-foldases.

iii) *secD*

This is a more recently identified gene thought to define another component involved in the general export pathway. This gene was identified by isolating extragenic suppressor mutants to PhoA-LacZ and LamB-LacZ export fusion proteins with altered Lac$^+$ phenotype (Gardel *et al.*, 1987). One mutant allele *secD*1 was found to be cold sensitive and at non-permissive temperatures accumulated the precursors of a number of exported proteins. Such conditions also led to the increased (de-repressed) synthesis of *secA*, which taken together suggest that it represents another gene involved in the general export pathway. Recently, the region encoding the *secD* gene has been cloned and sequenced (Gardel *et al.*, 1990) and another *sec* gene, *secF*, has also been identified as part of the same operon. Both SecD and SecF share sequence similarities and appear to be integral membrane proteins. This location is supported by *Tn-PhoA* analysis, which also suggests the presence of a large periplasmic domain. In addition, the inability to find suppressors in *secD/F* that can suppress signal sequence mutants also indicates that these proteins act at a late stage in the translocation process.

iv) *secE*

The observation that *secA* is regulated to the needs of the cell (Oliver and Beckwith, 1982) was used to select mutants affecting the regulation of *secA-lacZ* fusions. Indeed, this selection produced mutations in all genes which were known when inactivated to derepress *secA*. One new gene was also detected *secE*, by this method (Riggs *et al.*, 1988). Indeed, *secE* encodes for an integral membrane protein required for export (Schatz *et al.*, 1989) and genetic experiments by Silhavy's group (Bieker and Silhavy, 1989a) investigating the possible interaction of PrlA (SecY see

below) and PrlG (SecE) suppressors indicate that SecE and SecY interact directly and that SecE functions before SecY in the export pathway.

v) *secY/prlA*

Another approach to isolate extragenic suppressor mutants was employed by Emr *et al.*, (1981) which involved using proteins with export defective signals to select for suppressor mutants which could restore some degree of recognition and export of the defective protein. Many of the defined suppressor mutations were located within ribosomal genes but one mutant (*prlA*) was highly pleiotropic and suppressed mutations in the signal sequences of several exported proteins, such as LamB, MBP and PhoA, (Emr and Bassford, 1982; Michaelis *et al.*, 1983). However, this mutation did not affect the export of wild-type precursor proteins suggesting that an additional component of the export machinery may be required to mediate the interaction with the signal sequence (Randall *et al.*, 1987). However, Ito *et al.*, (1983) using localised mutagenesis *in vivo* in the region of *prlA* obtained a temperature sensitive mutant (*secY*) which was later demonstrated by genetical studies to be identical to *prlA*. The identification of SecY as an integral cytoplasmic membrane protein (Akiyama and Ito, 1985) and the observed degree of allelic specificity to various export defective mutants, also suggests that this mutation might act at the level of translocation. Evidence that SecY can, in contrast to SecA, function efficiently in a post-translational step is consistent with a role at a late step in export (Baker *et al.*, 1987). Biochemical analysis has confirmed a functional role for SecY in *in vitro* transport systems (Fandl and Tai, 1987) and the cellular position and detailed genetic studies suggest that it is part of the translocator (Bieker and Silhavy, 1989b).

Additional potential export proteins successfully identified by selecting for extragenic suppressors were PrlB/C (Emr and Bassford, 1982) and SecC (Ferro-Novick *et al.*, 1984). These mutations were not found to be pleiotropic, on the contrary they only affected the export of a single or a few specific proteins, and may represent minor components required only for the export of certain proteins or represent components of alternative minor export pathways .

vi) *prlB*

prlB was found to suppress only LamB defective signals, resulting in export without processing and has been shown to be, surprisingly, a C-terminal deletion of the periplasmic ribose-binding protein (Silhavy *et al.*, 1983).

vii) *prlC*

prlC suppresses LamB and MBP defective signal sequences (perhaps accompanied by aberrant processing) but does not have any affect on normal protein export, and might be associated with a minor export pathway.

viii) *secC*

An extragenic suppressor (*secC*) was isolated by selecting for revertants of a *secA*^{ts} strain and was found to be a cold sensitive mutant affecting synthesis of export but not of cytoplasmic proteins (Ferro-Novick *et al.*, 1984). However, inhibition of protein synthesis, by chloramphenicol at low concentrations, also reverses the effects of *secA*^{ts} mutants (Lee and Beckwith, 1986), suggesting that simply a reduction in the number of nascent proteins competing for the defective secretory machinery might allow it to function better.

Other attempts to identify additional genes involved in the export mechanism by isolating suppressor mutants of secretory defective proteins led, mostly, to the identification of elements of the protein synthesising machinery, such as ribosomal proteins. Perturbation of protein synthesis either directly or indirectly can affect the export of proteins (Lee and Beckwith, 1986; Shiba *et al.*, 1986) and as there are over 100 genes involved in protein synthesis the majority of suppressor mutations are likely to be located within such genes.

1.4.5 Other requirements for the bacterial translocation process:

The above information unquestionably demonstrates the absolute requirement for accessory proteins in the translocation process across the cytoplasmic membrane of *E. coli*. How this process is achieved and how many proteins are involved and indeed the precise mechanism involved is not yet known. The translocation process might require the accessory proteins to generate a specific proteinaceous pore (as initially postulated in the signal hypothesis - see section 1.3.1) or alter the protein or lipid bilayer to allow the protein to partition into and across the hydrophobic environment. Whatever form this process takes in bacteria, it is now known that this process is also dependent on the hydrolysis of ATP, a proton motive force also being required, though probably only for maximal rates (Chen and Tai, 1985; Geller *et al.*, 1986). Randall (1983) has shown, using protease studies on spheroplasts, that translocation in *E. coli* is not necessarily coupled to translation and that translocation

may only commence apparently after a critical length of the nascent protein has been synthesised. These results indicated the need for a separate energy source for translocation as the energy of chain elongation can no-longer be envisaged to force the export protein through the membrane.

Randall and Hardy (1986) have also shown that the export of the MBP precursor is dependent on the adoption and maintenance of a loose unfolded conformation. These authors also noticed that jamming of the export channel led the wild-type precursor to adopt a folded form incompetent for export. This effect occurred with similar kinetics to that displayed by a signal sequence export defective precursor protein (which did not interact with the export mechanism), suggesting that the signal sequence acts in concert with other cellular components to maintain an export competent conformation. A similar conclusion that proteins cannot be translocated in a tightly folded state has also been derived from work on the import of proteins into mitochondria. In their studies Eilers and Schatz (1986) have shown that dihydrofolate reductase (DHFR), fused to a mitochondrial presequence, cannot be imported into mitochondria in the presence of folate analogues which bind DHFR and presumably stabilise its tertiary structure. This idea was supported by the fact that presentation of the precursor protein in an unfolded form, either by urea denaturation or by the introduction of a destabilising mutation, led to accelerated import into mitochondria (Eilers *et al.*, 1988a; Vestweber and Schatz, 1988b). Thus, it would appear that *in vivo*, additional factors are required to maintain export proteins in a translocation competent form or to unfold such proteins prior to translocation. Indeed, such proteins have been identified in bacteria and eukaryotic systems and have been collectively called chaperones. In *E. coli*, SecB, trigger factor (TF) and GroEL have been identified in binding to the OmpA precursor, stabilising it as a substrate for translocation *in vitro* (Lecker *et al.*, 1989). Proteins with similar functions have been identified in eukaryotic systems, such as the Hsp70 family of proteins (heat shock proteins), which appear to maintain eukaryotic secretory and mitochondrial proteins in a loosely folded state prior to their translocation and may even have an unfoldase activity (Zimmermann *et al.*, 1988; Deshaies *et al.*, 1988; Chirico *et al.*, 1988). The post-translational import of proteins into mitochondria has been proposed to require ATP-dependent unfoldases (Rothman and Kornberg, 1986) but it is not certain whether ATP is directly required for the unfolding that appears to take place at the surface of mitochondria. The yeast mitochondrial matrix Hsp60 protein, highly homologous to the

E. coli GroEL protein, is also required to maintain imported proteins competent for insertion into or re-translocation across the inner membrane. Hsp60 is also proposed to aid protein assembly by binding unfolded polypeptides. The ATPase activity of this protein may therefore be coupled to conformational changes, allowing controlled domain by domain folding of the bound protein, with another activity (NEM-sensitive) required for Hsp60 release from the folded polypeptide (Ostermann *et al.*, 1989). A similar role have been suggested for the Hsp related endoplasmic reticulum protein, BIP (Normington *et al.*, 1989).

It is evident from the above information that the bacterial translocation process necessitates the use of an energy source (ATP), cytoplasmic factors as well as specific *sec* gene products, some of which have been defined above. Are these components also required for the true secretion mechanism(s) evolved in Gram negative bacteria for the transport of polypeptides from the site of synthesis (the cytoplasm), through two membrane systems and the intermembrane space (periplasm) to the external medium ?. For example, do these organisms use the general export machinery to export the protein across the first barrier and then translocate the secretory protein across the outer membrane by a different mechanism, or conversely, have unique systems evolved to translocate proteins directly from the cytoplasm across both membranes without using the general export pathway.

1.5 Translocation of proteins across bi-membrane systems:

In Gram positive bacteria the translocation of many export proteins across the cytoplasmic membrane follows a similar pathway to that described above for *E. coli* corresponding in effect to secretion. However, in Gram negative bacteria an extra complication arises for the secretion of polypeptides to the medium as a second membrane system, the outer membrane, has also to be traversed. Until recently it was thought that only a few proteins could be specifically secreted from Gram negative bacteria but now the literature describes the secretion of a wide number of proteins, mainly virulence factors, a number of which have been extensively investigated. From work to date (see below) it does not appear as if there is a single mechanism for secretion across the double membrane system. Nevertheless, the mechanisms elucidated so far can be subdivided into separate groups which are either dependent or independent of the general export pathway used to translocate proteins across the cytoplasmic membrane. Before describing the strategies evolved to traverse the double membrane cellular envelope it is first

relevant to briefly comment on the translocation of outer membrane proteins in Gram negative bacteria.

1.5.1 Export of proteins to the outer membrane in bacteria:

Like all periplasmic proteins most outer membrane proteins carry signal sequences which presumably ensure interaction with the general export machinery. It has been proposed that the assembly of outer membrane proteins is either a two step process, export to the periplasm followed by assembly into the outer membrane, or a single translocation step from the inner to outer membrane via adhesion zones (regions of contact between both membranes; Bayer, 1979). Most of the evidence to date supports the first hypothesis. In fact, periplasmic intermediates have been observed to accumulate when some outer membrane proteins or truncates are overexpressed (Freudl *et al.*, 1985 and 1986; Jackson *et al.*, 1986; Bosch *et al.*, 1986) while their ability to be translocated efficiently into inverted vesicles (Chen *et al.*, 1985; Yue, 1989) also suggests a periplasmic intermediate. On the other hand there is evidence that the assembly intermediates of some envelope proteins (such as TolQ and thioredoxin) are associated with envelope fractions enriched for possible adhesion or contact sites between inner and outer membranes (Bayer *et al.*, 1987; Bourdineaud *et al.*, 1989). In addition, the outer membrane porin, OmpF, has been observed from electron miscroscopy studies to emerge on the bacterial cell surface at distinct sites corresponding to these adhesion zones (Smit and Nikaido, 1978). However, the secretion of OmpF by spheroplasts (Metcalfe and Holland,1980) appears to contradict this proposed one step mechanism. Indeed, OmpF secreted by spheroplasts is monomeric in form and it has now been shown that they retain their ability to assemble into membranes and form trimers (Sen and Nikaido, 1990). This assembly process was apparently accelerated in the presence of LPS, which is absent from the inner leaflet of the outer membrane. Presumably therefore, *in vivo* the LPS required for assembly might only be accessible from the periplasm at special sites, corresponding perhaps to adhesion zones, in full agreement with the observations of Smit and Nikaido (1978). Thus, most of the evidence suggests that outer membrane proteins are first exported to the periplasm from where they assemble into the LPS containing surface membrane (see Section 1.2.4). The emergence of OmpF on the cell surface apparently at specific sites associated with adhesions zones remains a puzzle unless the final assembly step from the periplasm takes place at such regions.

1.5.2 Mechanisms for the secretion of proteins across the bi-membrane envelope of Gram negative organisms:

1.5.3 N-terminal signal peptide dependent secretion:

The majority of the Gram negative secretion systems investigated appear to utilise the general export pathway (SecA/Y) to translocate the protein, destined to be secreted, in a first step across the cytoplasmic membrane. This conclusion is based in most cases on the fact that such secreted proteins appear to carry a classical, transient N-terminal signal sequence. In addition, the essential role of the signal peptide of the secreted Exotoxin A from *Pseudomonas aeruginosa*, was demonstrated by the inability of the precursor to be translocated or processed in a *secA*⁻ mutant when expressed in *E. coli* (Douglas *et al.*, 1987). Similarly, a SecA/Y dependent step in the secretion of Pullulanase (PulA originating from *Klebsiella pneumoniae*) was demonstrated in *E. coli*, carrying all the information required for pullulanase secretion. In these studies, a PulA-LacZ fusion was found to block the export of other N-terminal targeted proteins, while PulA-PhoA fusions were at least partially exported to the periplasm signifying interaction with the general export pathway (d'Enfert and Pugsley, 1987).

There appear to be two classes of protein that utilise the SecA/Y export pathway in the secretion process, those that require specific ancillary protein components to facilitate secretion through the outer membrane and those that encode all the necessary information to secure their own specific release from the cellular envelope.

1.5.3a Secretion requiring ancillary proteins:

A number of secreted proteins appear to require additional specific elements to facilitate their complete secretion from their Gram negative host organism. This has been deduced from the isolation of mutants which can no longer secrete the polypeptide or the inability of such proteins to be secreted when expressed in a related organism. This latter method has been used to demonstrate the requirement for extra secretory proteins involved in the secretion of proteins such as Aerolysin from *Aeromonas hydrophilia*, *E. coli* Heat Labile and Cholerae toxins from *Vibrio cholerae*, Exotoxin A from *Psuedomonas aeruginosa*, and Pullulanase from *Klebsiella pneumoniae*.

i) Secretion of Aerolysin from *Aeromonas hydrophilia*

This secreted haemolytic toxin is synthesised as a precursor with a typical N-terminal signal sequence which is processed under normal conditions but accumulates in its precursor form upon the addition of the energy uncoupler CCCP (Howard and Buckley, 1985). In a pleiotropic export defective mutant the toxin was located in the periplasm suggesting a two step secretion mechanism. This was confirmed by pulse chase experiments where rapid processing occurred, indicative of translocation, followed by a lag before release of the toxin into the medium. However, when the Aerolysin gene was expressed in *E. coli* the toxin was not secreted, whereas the cellular associated toxin was processed and found to be located in the periplasmic shock fluid (Howard and Buckley, 1986). This result suggested that the signal sequence does normally function in exporting this protein from the cytoplasm but *E. coli* lacks the specific component(s) required to enable the translocation of the toxin from the periplasm to the external medium.

ii) Secretion of *E. coli* Heat Labile and Cholerae toxins from *Vibrio cholerae*

The cholera toxin (CT) from *Vibrio cholerae* and the heat labile toxin (HLT) of *E. coli* are structurally and functionally similar holotoxins composed of two separate subunits, A and B in a ratio of 1:5. Surprisingly, *V. cholera* can specifically secrete its own or HL toxin whereas *E. coli* cannot secrete either, and both toxins are exported only as far as the periplasm. Both of the subunits are synthesised with N-terminal signal sequences which are processed during translocation to the periplasm. Kinetic experiments have shown that the holoenzyme is assembled in the periplasm before it is finally secreted (Hirst and Holmgren, 1987a; 1987b). These results again demonstrate the requirement for additional, specific components to recognize the toxin and translocate it across the outer membrane. Indeed, it has recently been reported that the genes encoding these extra functions have been cloned (Marcus H and Holmes RK, 1990 ASM meeting Abstr. abstract No B-190)

iii) Secretion of Exotoxin A from *Pseudomonas aeruginosa*

The above two examples reflect that the secretion mechanism employed by many Gram negative organisms is a two step mechanism, export to the periplasm followed by recognition and translocation across the outer membrane by specific factors. However, in the case of Exotoxin A this does not appear to be the case. Lory *et al.*, (1983) have reported that this toxin is rapidly secreted in a co-translational manner, but if the outer

membrane is perturbed with 10% ethanol then secretion is prevented. Moreover, the toxin was located as a precursor enzymatically active form exposed on the outer membrane surface. The authors could not envisage a two step secretion process, but suggested that the protein was transferred from the inner to outer membrane via a Bayer-like junction with another specific function required for release.

When the gene encoding exotoxin A was expressed in *E. coli* it was not secreted (Douglas *et al.*, 1987). However, analysis by osmotic shock and Western blotting surprisingly, in this case identified the exotoxin in the periplasm, while *secA* mutants led to its accumulation in the cytoplasm. This result contrasts with the location of the toxin in *Pseudomonas*, where no periplasmic intermediate could be detected. Thus, caution must be taken when interpreting the existence of a periplasmic intermediate of a heterologous protein expressed in *E. coli* which might not necessarily reflect the normal mechanism.

iv) Secretion of Pullulanase from *Klebsiella pneumoniae*

Pullulanase (PulA) is a secreted lipoprotein which has starch debranching activity. PulA is synthesised with a typical N-terminal signal peptide which is processed by the lipoprotein signal peptidase (Pugsley et al., 1986). The addition of the fatty-acyl chains at the N-terminus apparently anchors the lipoprotein to the outer face of the outer membrane from which it is slowly released after the end of exponential growth (Pugsley et al., 1986). When the pulA gene was introduced and expressed in E. coli, PulA was not translocated to the outer membrane (Pugsley et al., 1990). However, co-expression with adjacent DNA fragments was found to permit PulA translocation to the outer membrane of E. coli, from which it was subsequently released (d' Enfert et al., 1987, Pugsley et al., 1990). Analysis of this extragenic DNA demonstrated that at least 7 proteins are involved in the secretion of PulA, with at least 4 more genes, as yet, uncharacterised (Pugsley and Reyss, 1990). No true periplasmic intermediate has been detected, either in the presence or absence of the specific secretory components, although PulA has apparently been localised during transport on the outer face of both the outer and inner membranes (Pugsley et al., 1990). A periplasmic intermediate, if it exists, might however be too transitory to isolate or impossible to detect in secretory mutants due to the properties of this lipoprotein which is expected to associate with the inner membrane if the export system is perturbed. Indeed, the preliminary localisation of a number of the ancillary proteins required for PulA secretion primarily to the inner

membrane (although some also fractionated with the outer membrane - d'Enfert et al., 1989; Pugsley and Reyss, 1990) may indicate an initial interaction with PulA following SecA/Y transport, which is then immediately translocated to the outer membrane with the help of other proteins before been slowly released by yet other specific functions.

1.5.3b Proteins encoding all the information required for their own secretion:

Both *Neisseria gonorrhoeae* and *Serratia marcescens* secrete proteases also synthesised with typical N-terminal signal sequences (Pohlner *et al.*, 1987; Yanigida *et al.*, 1986). The *Serratia* protease gene sequence has the capacity to encode for a 112KD polypeptide whereas the mature protein is only 40-50 KD in size. Partial amino acid sequencing of the secreted peptide by Yanigida's group showed that the protein was processed both at the N-terminus and C-terminus. Moreover, a 3' frameshift secretion defective mutant suggested an important role for the C-terminus in secretion. It was assumed that the N-terminal signal sequence directed the protein across the inner membrane but no periplasmic intermediate was isolated. Similar results were obtained by Pohlner and colleagues for the secretion of the 106KD Neisseria gonorrhoeae IgA protease from a potentially 169KD precursor. In this case the processed C-terminal peptide was found located in the outer membrane. On the basis of these results these authors proposed a model for secretion where the signal peptide targets the protein across the inner membrane and then the amphiphathic nature of the C-terminal "helper" domain allows its assembly into the outer membrane. The pore which is then formed provides the mechanism for the secretion of the protease located in the N-terminal domain. As a result of transport across the outer membrane, the protease then adopts an active conformation and is released from the membrane bound C-terminal domain by autoproteolysis. It is not known whether a periplasmic intermediate exists in this system or whether the helper domain directs the protease to the outer membrane via the putative adhesion zones.

Apparently two secreted chitinases (58 and 52KD respectively) from *Serratia marcescens*, encoded by *chiA* and *chiB* respectively, are secreted from *E. coli* when expressed in that organism (Jones *et al.*, 1986). The *chiA* gene was sequenced and an open reading frame encoding a 61KD protein was predicted. Sequence analysis also predicted the presence of a typical signal peptide, which was shown to be absent from the secreted

protein. The mechanism of secretion is not known but presumably also involves the SecA/Y general export pathway in the first step.

1.5.4 N-terminal signal peptide independent secretion - Colicins:

Of the secretion systems investigated so far only two examples have been identified where the secreted protein does not utilise an N-terminal signal sequence and thus would appear not to require the general SecA/Y export pathway. Instead, other specific proteins are required to translocate the polypeptide across both the inner and outer membranes in the case of the secretion of alpha-haemolysin from *E. coli* (see Section 1.6). In fact, haemolysin is the only known example of specific secretion of a polypeptide in *E. coli*, although some strains also release bacteriocins in a relatively non-specific manner (Baty *et al.*, 1987). These colicin molecules are expressed as a result of the SOS response following DNA damage and deletion/fusion analysis suggest that they do not encode discrete targeting signals (Baty *et al.*, 1987). Moreover, colicins are probably secreted directly across both membranes in a fully folded form (see de Graaf and Oudega, 1986; C. Lazdunski; personal communications). It is now known that their release requires the expression of another protein, the bacteriocin release protein (BRP), which is apparently synthesised with a typical N-terminal signal sequence which might be modified and recognized by the lipoprotein signal peptidase (reviewed by de Graaf and Oudega, 1986). BRP has been localised in both the inner and outer membrane (Oudega *et al.*, 1984; Cole *et al.*, 1985) and is postulated to promote the release of the bacteriocins by activating the outer membrane detergent resistant phospholipase A resulting in membrane perturbation (see Pugsley and Schwartz, 1984). The resulting increased permeability then allows the relatively non-specific release of the colicin along with some other soluble proteins, both cytoplasmic and periplasmic. The overexpression of the BRP protein is probably responsible for the observed cell lysis following induction but at low levels of induction the bacteriocins are somehow released together with a few other minor soluble proteins without resulting in cell death.

Now I will describe, in some detail, the mechanism of secretion, elucidated so far, for the specific secretion of the haemolysin toxin (HlyA) from some pathogenic strains of *E. coli*, as the work of this thesis is based on this system.

1.6 Secretion of Haemolysin from *Escherichia coli*:

1.6.1 Introduction:

It is almost 3 decades since it was observed that certain strains of *E. coli* release a toxin, haemolysin, into the extracellular medium (Smith, 1963). Indeed, this haemolysin molecule has been identified as an important factor in the virulence of pathogenic *E. coli* strains (Welch *et al.*, 1981) usually associated with urinary tract infections (Cavalieri *et al.*, 1984). The toxin is apparently cytotoxic for a number of cell types (see Mackman *et al.*, 1986) and is easily assayed quantitatively by the release of haemoglobin from red blood cells (Haemolytic assay).

The haemolysin determinant has been isolated from different pathogenic strains and is commonly chromosomally located on strains infecting Man whereas those infecting animals have the determinant on conjugative plasmids. All these determinants, however, show a high degree of conservation (Muller *et al.*, 1983). The entire determinant is carried on an approximately 7Kb DNA fragment (see Figure 1a), encoding four genes *hlyC,A,B* and *D* (Welch *et al.*, 1983; Mackman *et al.*, 1985a; Noegel *et al.*, 1979 and Hartlein *et al.*, 1983). The description and function of these gene products will be discussed below, together with the mechanisms of regulation and synthesis which are still not, by any means, fully understood. The *hlyA* and *hlyC* gene products are required for haemolytic activity whereas the other two genes are essential for the export of the toxin to the external medium as defined by transpositional mutagenesis (see below).

1.6.2 The haemolysin toxin and activating protein:

HlyC:

The *hlyC* gene product was first identified by Noegel *et al.*, (1979) using the mini-cell system as an 18KD protein, and is required only for haemolytic activity and not secretion. These authors, as it turned out, incorrectly proposed that the function of this protein was to process the intracellular haemolysin toxin (107KD HlyA) into a 60KD active molecule which was then secreted. Studies by Nicaud *et al.*, (1985a) demonstrated conclusively that the *hlyC* gene product (which was cytoplasmically located) was only required for the activation of the toxin in some as yet undefined manner, and when absent the inactive toxin was still secreted by the export proteins. Here I will concentrate on the properties of the LE2001 determinant, used in this laboratory, although all the haemolysin determinants studied so far are similar in organisation and function.

HlyA:

The size of the secreted toxin molecule (HlyA) was elusive for many years apparently due to its instability and ability to aggregate. Eventually, Mackman and Holland (1984a) were able to show conclusively that the toxin activity was associated with the secretion of a 107KD protein by the use of a robust laboratory strain which normally grows with very little lysis and by the addition of 10mM CaCl2 (which was found to stabilise the toxin). This molecular weight was later substantiated from the sequence analysis of the *hlyA* gene from the pSF4000 determinant (Felmlee *et al.*, 1985a) indicating the capacity to encode a 110KD protein. However, it has also been reported that two forms of HlyA can be detected by SDS-PAGE under partially denaturing conditions (Nicaud *et al.*, 1985a; Wagner *et al.*, 1988), apparently corresponding to the modified (active) or unmodified (inactive) forms (Wagner *et al.*, 1988). Sequence analysis of the *hlyA* gene (Felmlee *et al.*, 1985a) failed to identify a classical N-terminal signal sequence, as found transiently on most export proteins. This was peculiar as a periplasmic pool of haemolytic activity had been reported (Hartlein *et al.*, 1983). In addition, and surprisingly no processing of the toxin was detected during the secretion process (Felmlee *et al.*, 1985b). Indeed, further investigation by Gray *et al.*, (1986) into the reported periplasmic activity failed to detect significant levels of such activity in either the presence or absence of the haemolysin export proteins. These results indicated that haemolysin secretion did not involve a periplasmic intermediate. Studies with other related *hly* determinants have confirmed the absence of detectable periplasmic intermediates during the normal secretion process or when secretion is blocked by the deletion of the C-terminal targeting signal of HlyA (Felmlee and Welch, 1988; Koronakis *et al.*, 1989 and Oropeza-Wekerle *et al.*, 1989). Further evidence for the absence of a periplasmic intermediate was provided by Karen Baker (Baker, 1987; Gray *et al.*, 1989) who showed that the secretion of the haemolysin toxin was independent of the secA gene product. Kinetic experiments have also indicated that the translocation of the toxin to the medium is very fast (Felmlee and Welch, 1988) with a half life of 1 to 3 minutes. These data are consistent with the absence of a periplasmic intermediate, and indicate rather a mechanism that translocates the toxin directly across both membranes in a single step.

One of the unexpected and novel features of the haemolysin secretion system was the presence of a C-terminal, not N-terminal, targeting signal uncovered by the removal of the final 27 amino acids of HlyA and their

replacement by 9 unrelated residues (Gray *et al.*, 1986). The resulting molecule was not secreted by the HlyB,D export proteins, although intracellular haemolytic activity was detected, indicating the disruption of a specific recognition or targeting signal. The molecular analysis of this portion of the gene will be discussed in detail later as it is this region that I am primarily concerned with in this thesis (see Section 1.6.5 below). The presence of such a C-terminal signal region also implied that the secretion process must be, by definition, a post-translational process.

C-terminal deletion studies of *hlyA* by Felmlee and Welch (1988) suggested that at least the first ~80KD of the HlyA molecule constituted an independent domain retaining toxin activity, with the remaining segment encoding, and presumably involved in the correct presentation of, the novel targeting signal to the Hly export apparatus (Gray *et al.*, 1986; Felmlee and Welch, 1988). This was corroborated by the fact that the independent expression of the C-terminal 23KD of HlyA resulted in the efficient secretion to the medium of this peptide in a purely HlyB,D dependent manner (Nicaud *et al.*, 1986). Studies into the mode of action of the toxin (Bhakdi *et al.*, 1986; Menestrina, 1988) suggested that the HlyA molecule acts as a single monomer forming relatively non-specific pores (approximately 3nm in diameter) which result in rapid movement of cations and small uncharged molecules, presumably leading to cell death. Analysis of the sequence of the *hlyA* gene predicted the presence of three hydrophobic domains in the toxin domain (Felmlee *et al.*, 1985a), which might therefore be involved in the formation of a transmembrane pore in the target membranes. Such a role is supported by results from mutations to one of these domains that abolish haemolytic activity but not secretion (Ludwig *et al.*, 1987). At the interface of the apparently two separate functional domains (toxin and "signal" domains) a tandem nona-peptide repeat region (see Figure 1b) predicted to adopt random coils and turns has been identified (Felmlee *et al.*, 1985a; Gray *et al.*, 1986; Holland *et al.*, 1990a). This could conceivably act as a hinge region separating the two domains allowing their independent folding. However, more recent investigations into the role of this repeat region also suggests that its presence is infact required for haemolytic activity (Felmlee and Welch, 1988; Ludwig *et al.*, 1988). Indeed Ludwig *et al.*, (1988) and Felmlee and Welch (1988) obtained data suggesting that this region is involved with binding to Calcium which is required for haemolytic activity (Springer and Goebel; 1980). In contrast, the repeat structures are not essential for secretion since a C-terminal portion of HlyA, not carrying any repeats is

still secreted specifically in an HlyB,D dependent manner and such a peptide can support the secretion of heterologous proteins (see below).

1.6.3 The export proteins:
HlyB and HlyD:

Insertional mutagenesis (Wagner *et al.*, 1983; Mackman *et al.*, 1985b) demonstrated that an approximately 4Kb DNA fragment (encoding two cistrons, *hlyB* and *hlyD*) 3' to the *hlyA* gene was required for the translocation of the haemolysin toxin molecule to the extracellular media. The products of these two genes could not be identified *in vivo* so *in vitro* translation systems together with mini-cells and maxi-cell systems were used to identify the gene products. Hartlein *et al* (1983) used mini-cells to identify the export proteins HlyB and HlyD as 46KD and 62KD, respectively. Similarly, Felmlee *et al.*, (1985a) identified a single 54KD protein from *hlyD* and two products from *hlyB*, a 77KD and 46KD polypeptide, the 46KD protein was proposed to be an internal translation product predicted to be present from sequence data (Felmlee *et al.*, 1985a). In this laboratory *hlyB* from LE2001 was observed to produce both a 66 and 46KD protein in an *in vitro* translation system and only a 66KD protein in minicells (Mackman and Holland, 1985b). Pulse chase experiments could not provide any evidence for a precursor-product relationship and the significance of the two forms is still unknown. The LE2001 gene product of *hlyD* was again identified as a 54KD polypeptide. Breakage and fractionation of mini-cells by Osborn sucrose equilibrium gradients suggested an outer membrane location for HlyB,D (Hartlein *et al.*, 1983), whereas sarkosyl separations of the membranes of mini-cells showed a predominantly inner membrane location for HlyD, although small but significant amounts were also found to be associated with the outer membrane (Mackman *et al.*, 1985b). The absence of N-terminal sequences (from sequence data; Felmlee *et al.*, 1985a) for these proteins would also support an inner membrane location for HlyB and HlyD. Subsequent, more detailed localisation and protease studies in mini-cells overproducing HlyB and HlyD have also indicated that both HlyB and D are mainly inner membrane proteins (Wang *et al.*, 1991 In press J. Mol. Biol; Holland *et al.*, 1989).

Topological analysis of the export proteins by C-terminal fusions to beta-lactamase as a probe (only active when translocated to the periplasm) suggest that HlyB has 6 to 8 transmembrane loops in the N-terminal region with a large C-terminal cytoplasmic domain (Wang *et al.*, 1991 In press J. Mol. Biol). This C-terminal domain carries a highly conserved

ATP binding site which is found in several proteins involved in transport mechanisms in *E. coli* and *Salmonella typhimurium* (Higgins *et al.*, 1985; 1986). Similar topological and protease studies with HlyD suggested that it has one large transmembrane loop. On the basis of such data and the apparent absence of any periplasmic HlyA secretion intermediates it was proposed that HlyB may provide energy for the translocation of HlyA through a proteinaceous pore straddling the periplasm, constituted by the membrane domain of HlyB and the extramembrane domain of HlyD (Mackman *et al.*, 1986; Holland *et al.*, 1990a; 1990b).

In the last two to three years HlyB has been established as a member of a increasingly large family of proteins involved in the transport of such diverse molecules as polypeptides (from various bacterial species), cyclic polysaccharides (from species of bacteria), hydrophobic drugs (from Plasmodium falciparium and Man) and even a yeast mating factor (see Blight and Holland, 1990 for recent review).

1.6.4 Regulation of the *hly* determinant:

The *hly* determinant appears to be expressed from two major promoters, the first upstream of *hlyC* thought to initiate transcription of the *hlyC,A,B* and *D* genes and the second upstream of the *hlyD* gene. These conclusions were deduced from insertional mutagenesis, mutants of *hlyA* being polar for *hlyB* but not *hlyD* (Wagner *et al.*, 1983; Mackman *et al.*, 1985a) and supported by the observation that expression of *hlyD* is greater that *hlyB* in *in vitro* systems. In fact, the former group initially deduced the orientation of transcription of the *hlyD* gene as opposed to the other three genes, but sequence analysis later confirmed the same orientation of transcription and identified a putative upstream promoter for *hlyD* within *hlyB* (Felmlee *et al.*, 1985a). Regulation of the determinant appeared to be largely controlled by modulation of the promoter upstream of *hlyC* as fusions of beta-galactosidase to the *hlyC* gene showed that transcription from the upstream promoter reflected the secreted haemolytic activity, that is, there was high level transcription during exponential phase with a switch off during late exponential growth (Nicaud *et al.*, 1985b). Switch off of synthesis was observed to be accompanied by rapid loss of haemolytic activity in the medium and degradation of the toxin (Nicaud *et al.*, 1985b). However, the loss of detectable extracellular haemolytic activity is not necessarily due to rapid proteolysis of the toxin but rather to a gradual reversal of the activation-modification, mediated by HlyC (Oropeza-Wekerle *et al.*, 1989). Indeed, Wagner *et al.*, (1988) have shown that haemolysin can lose its activity

upon ultrasonication or treatment with phospholipase C which they suggest reverses in some way the HlyC induced activation.

The idea that the regulation of expression of the *hly* determinant was more complicated was indicated by the fact that the level of the export proteins appeared to be tightly regulated, especially HlyB, as demonstrated by the ability only to detect them in *in vitro* systems (Hartlein *et al.*, 1983; Mackman *et al.*, 1985b). Indeed, recent evidence by specifically probing levels of *hly* specific mRNA detected only two transcripts, an abundant *CA* and minor *CABD* transcript (Welch and Pellett, 1988). It is not known whether the smaller transcript represents a processed form of the larger although a strong rho-independent termination loop has been proposed to exist in the intergenic region between *hlyA* and *hlyB* from sequence analysis (Felmlee *et al.*, 1985a). Indeed, Hughes' group (Koronakis *et al.*, 1988) has shown that the presence of this intergenic region does indeed reduce the level of transcriptional readthrough by about 80%. The presence of a transcriptional enhancing sequence several kilobases upstream of *hlyC* (a cis-acting, non-transcribed element, *hlyR*, approximately 650 bp long) identified by Vogel *et al.*, (1988), which enhanced the level of secretion promoted by the cloned pHly152 determinant by at least 50 fold, acts to reduce the termination of transcription before *hlyB* to only 1%, demonstrating the strong regulational capacity of this region (Koronakis *et al.*, 1988).

Attempts by different laboratories to identify the 5' major promoter by sequence comparisons and mapping mRNA start points has produced different numbers of promoters and locations (Felmlee *et al.*, 1985a; Hess *et al.*, 1986; Welch and Pellett, 1988; Koronakis and Hughes, 1988). Although there is remarkable homology between the related haemolysin genes and intergenic regions used in these studies, there is very little homology 5' of the *hlyC* gene where the promoter resides. This inability to identify one common promoter may be due to the evolution and divergence of this region. This may also be correlated with the presumed mobility of the *hly* determinant within the genome suggested by the presence of IS-like sequences flanking several plasmid encoded *hly* determinants (Hess *et al.*, 1986)

Evidence from earlier studies was also presented that different media affected the level of haemolytic activity, whilst the addition of

haemoglobin to the medium also apparently increased haemolytic activity (Springer and Goebel, 1980). Subsequently, Nicaud *et al.*, (1985b) presented evidence using *hlyC-lacZ* fusions that these effects were not due to transcriptional regulation but rather to differential stability of the toxin, the half life at 37°C being 25 minutes in minimal media and 2.5 hours in rich media supplemented with 10mM $CaCl_2$. In contrast, there does appear to be evidence for some feedback regulation of the *hly* expression as the absence of the export proteins, led to reduced transcription from the *hlyC* promoter (Holland *et al.*, 1989; Oropeza-Wekerle *et al.*, 1990)

1.6.5 Molecular analysis of the C-terminus of the HlyA molecule:

As mentioned earlier the first indications that the HlyA molecule had an unusual translocation targeting signal came from the observation that substitution of the final 27 C-terminal amino acids abolished secretion but not haemolytic activity (Gray *et al.*, 1986). In addition, it had previously been demonstrated that the molecule possessed no recognisable N-terminal signal sequence and was secreted without any processing (Felmlee *et al.*, 1985a, 1985b). Indeed, subsequent experiments demonstrated the efficient secretion of either the C-terminal 23KD or 12KD peptides of HlyA in an HlyB,D dependent manner (Nicaud *et al.*, 1986; Mackman *et al.*, 1987) establishing the molecular location of the signal to within the last 110 residues (330bp of *hlyA*). Initial investigations of the capacity of the C-terminal 23KD signal region of HlyA to export heterologous polypeptides, by generating fusions, met with mixed results. Apparently, this portion of the HlyA molecule could not be employed to secrete fragments of two *E. coli* cytoplasmic proteins, beta-galactosidase and chloramphenicol trans-acetylase (unpublished results, this laboratory), whereas the *E. coli* outer membrane protein, OmpF, deleted for its N-terminal signal sequence was secreted with high efficiency when fused to either the 23KD or 12KD signal region (Mackman *et al.*, 1987).

A series of deletions were constructed in the *hlyA* portion of the *ompF-23KDhlyA* gene fusion in an attempt to locate the signal sequence more precisely as presented by Mackman *et al.*, (1987). This work revealed that deletion of 93 amino acids immediately upstream of the final 27 amino acids (AA) residues of the HlyA toxin still permitted the hybrid to be secreted (see Figure 2). This suggested that all the information required for recognition and secretion is within the last 27AA agreeing with the

previous observation that substitution of the final 27AA of HlyA abolished secretion. Surprisingly, extending this 93 residue deletion proximally by another 46AA abolished secretion of the hybrid protein. More surprisingly the additional amino acids removed by this deletion encompass a region not considered necessary for signal recognition, as it is not present or required, for example, for the secretion of an OmpF-120AA HlyA hybrid. Subsequent secondary structure predictions (Chou-Fasman) for these two deletions compared to that of the 23KD peptide itself, suggested that both fusions disrupted a predicted strong alpha-helical structure which was partially recreated in the first, but not in the second fusion (see Figure 2), indicating a role for this structure in the secretion process (Mackman *et al.*, 1987). An alternative hypothesis was that the necessary recognition information is indeed present within the last 27AA and that the spacing distance between this and the downstream OmpF domain was important for correct functioning of the targeting signal. The second fusion in this case was presumed to result in the occlusion of the signal sequence. The importance of the alpha-helix itself was later tested by constructing, with the aid of synthetic oligonucleotides, a fusion of 22 residues of HlyA (encoding only the strong alpha-helix structure) to the OmpF protein, again deleted for its signal sequence. Expression of this construct did not result in any detectable secretion (unpublished results this laboratory, Isabella Jazinski, Karen Baker). This result again, however, did not rule out a vital role for this region as the specific recognition sequence since in this minimal construct the sequence might have been rendered inaccessible to the export machinery by the OmpF domain.

1.6.6 Current secretion model:

The above studies suggested that haemolysin is secreted without a periplasmic intermediate and independently of the SecA/Y pathway. Indeed, haemolysin has not been detected in the periplasmic space of cells missing either or both HlyB and HlyD (export proteins) and spheroplasted cells do not secrete the toxin (Baker, 1987). These results suggest that HlyB,D can translocate the toxin from the cytoplasm directly to the external medium without first entering the periplasm. The localisation of both HlyB and D primarily to the inner membrane, with significant amounts associated with the outer membrane, together with the results of topological analysis, suggest that these proteins may interact together to span the periplasmic space. But do these inner membrane proteins interact with the outer membrane through the so called Bayer adhesion zones or do they span the envelope to generate their own haemolysin specific (transient) adhesion zones ?. Recent

evidence shows that haemolysin secretion is dependent on the presence of a minor outer membrane protein, TolC, a protein thought to be involved in membrane integrity and colicin uptake (Wandersman and Delepelaire, 1990). There is now also evidence for the role of adhesion sites in the entry of colicins (Baty *et al.*, 1990) which tentatively supports ideas of a role for adhesion sites in the secretion of HlyA involving TolC. Thus, it may be envisaged that HlyB,D interact together to form a specific transmembrane pore, with the proposed cytoplasmic domain of HlyB interacting with HlyA post-translationally and the membrane domain of HlyB, the extra membrane domain of HlyD plus the TolC protein then promoting the passage of HlyA directly to the medium. The proposed ATPase activity of HlyB might be required for the unfolding of the molecule before direct transfer across both membranes and release into the external medium, or ATP may be required for steps in the transport process itself.

1.7 Aims of study:

The initial aim was to probe the nature of the novel HlyA C-terminal targeting domain by testing its ability to promote the secretion of various heterologous-HlyA hybrid polypeptides via the Hly export channel. If this could be achieved then it was expected that it should be possible, by generating different hybrids, to investigate the location within the C-terminal fragment of the targeting signal and some of the principles/constraints involved in the secretion of passenger peptides.

A more rigorous investigation of the location and nature of the HlyA targeting signal was also envisaged using a variety of strategies for *in vitro* mutagenesis, either specifically or randomly, of the 3' region of *hlyA* encoding the targeting signal. The analysis, classification and identification of resulting mutants, defective or not for secretion, was anticipated to highlight residues both critical or non-essential for the secretion process. As a consequence of such an analysis of the C-terminal region of HlyA it was hoped that the structure/motif(s) specifically required for secretion would be elucidated.

Figure 1

Organisation of the *hly* LE2001 determinant and the *hlyA* gene.

a) Shows a representation of the *hly* determinant, with the origin of detected promoter activities indicated by the thick arrowed lines. The broken arrow line reflects the drop in transcriptional readthrough due to the presence of a strong rho-independent terminator situated between *hlyA* and *hlyB* (see Chapter 1; Section 1.6.4)

b) Depicts the features of the HlyA protein encoded by *hlyA*.
(see text)

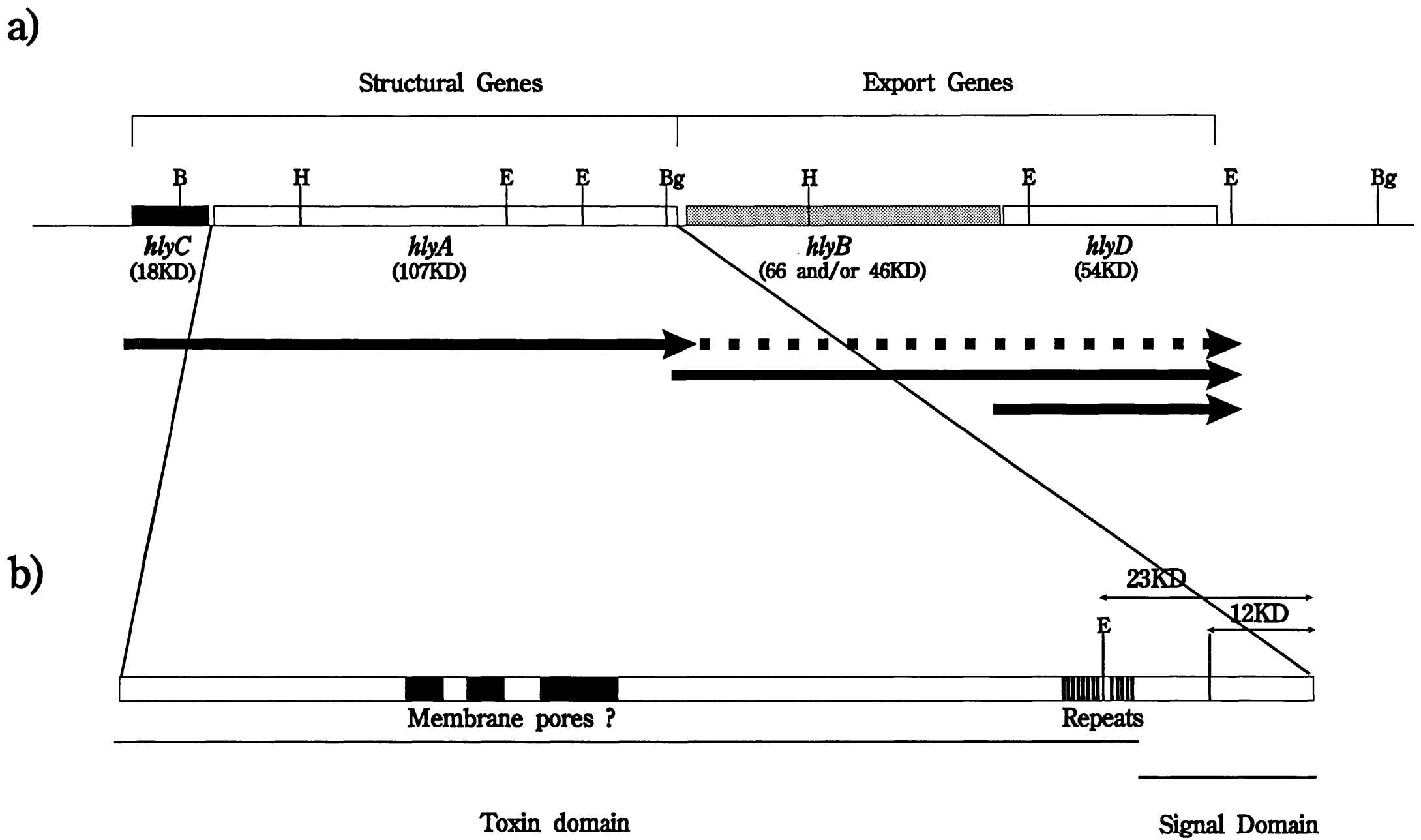

a)
Structural Genes
Export Genes
B
H
E
E
Bg
H
E
E
Bg
hlyC
(18KD)
hlyA
(107KD)
hlyB
(66 and/or 46KD)
hlyD
(54KD)
b)
23KD
12KD
E
Membrane pores ?
Repeats
Toxin domain
Signal Domain

Figure 2:

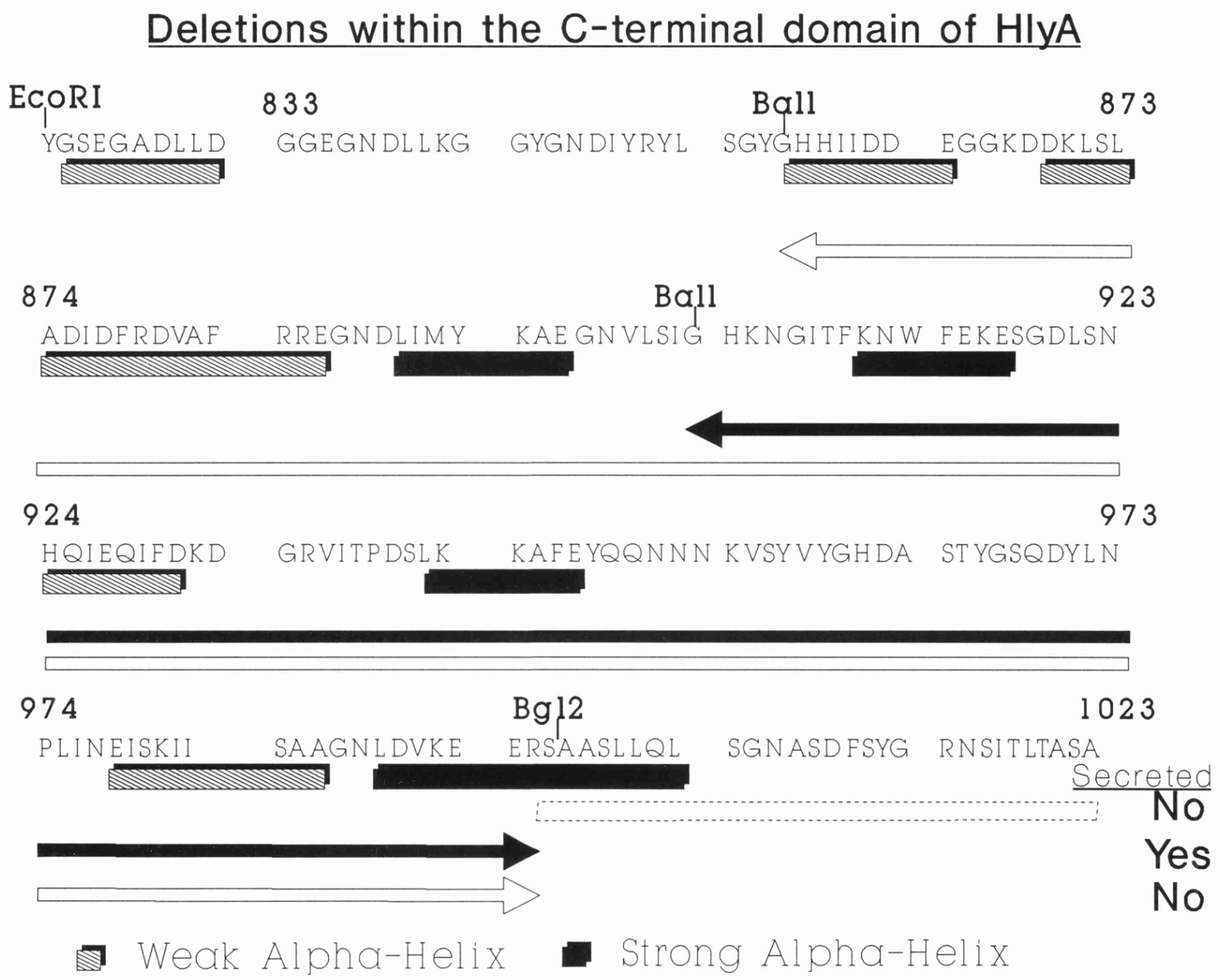
Deletions within the C-terminal domain of HlyA LE2001.

The primary amino acid sequence for the C-terminal 23KD domain of HlyA is given, along with predicted alpha-helical structures (using method of Chou and Fasman, 1978).

The dashed line represents a C-terminal deletion of HlyA (Gray *et al.*, 1986; see text).

THe arrowed lines show the extent of two internal C-terminal HlyA deletions within an OmpF-23KDHlyA hybrid protein (Mackman *et al.*, 1987).

The ability or inability of these truncated proteins to be recognised and secreted by HlyB,D is indicated in the bottom right hand corner.


Deletions within the C-terminal domain of HlyA
EcoRI
833
BglI
873
YGSEGADLLD GGEGNDLLKG GYGNDIYRYL SGYGHHIIDD EGGKDDKLSL
874
BglI
923
ADIDFRDVAF RREGNDLIMY KAEGNVLSIG HKNGITFKNW FEKESGDLSN
924
973
HQIEQIFDKD GRVITPDSLK KAFEYQQNNN KVSYVYGHDA STYGSQDYLN
974
Bgl2
1023
PLINEISKII SAAGNLDVKE ERSAASLLQL SGNASDFSYG RNSITLTASA
Secreted
No
Yes
No
Weak Alpha-Helix
Strong Alpha-Helix

CHAPTER 2

Materials and Methods:

2.1 Growth and Maintenance of strains and plasmids:

2.1.1 Strains:

The following *E.coli* strains were used:-

SE5000: K12- F^-, *rspL*, *ara*139, *thi*, Δ(*lacIPOZYA*), U169, *recA*57: K12-*desC*1, *flb*301, *relA*, *ptsF*25

NM522: K12- Δ(*lac-proAB*), *thi*, *sup*E, $r_k^- m_k^-$ [F'*proAB*,*lacI*^{q} ZDM15]

JM101: K12- *supE*, *thi*, Δ(*lac-proAB*), (F', *traD*36, *proA*$^+$*B*$^+$, *lacI*^{q}, ZDM15), $r_k^+ m_k^+$

DH1: K12- F^-, *recA*1, *endA*1, *gyrA*96, *thi*, $r_k^+ m_k^+$, *supE*44, *relA*1, λ^-

ΔH1: K12- F^-, *lac*$^-$, am Sm^R (λcI^{857} N7N53 ΔHI *bio*$^-$)

TG1: K12, Δ(*lac-pro*), *supE*, *thi*, *hsdD*5/F'*traD*36, *proA*$^+$*B*$^+$, *lacI*^{q}, *lacZ* ΔM15

BL21 (DE3): B- F^-, *ompT*,$m_b^- m_b^-$, λDE3

2.1.2 Plasmids:

pLG570 (pOU71: *hlyC,A,B,D*): Mackman and Holland, 1985a

pLG575 (pACYC184: *hlyB,D*): Mackman *et al.*, 1985a

pLG577 (pACYC184: *hlyC,A'*): Nicaud *et al.*, 1985a

pLG609 (pTTQ18: 3' hlyA: 23KD HlyA): Nicaud *et al.*, 1986

pLG612 (pUC: 3' hlyA: 23KD HlyA): R. Haigh this laboratory

pMG168 (Dual origin vector; *pcm*): Wright *et al.*, 1986

pMG196 (Dual origin vector): Wright *et al.*, 1986

pRK248 (cI^{857}): Bernard and Helinski, 1979

pEX-2 (*cro-lacZ*): Stanley and Luzio, 1984

pET-11d (T_7 RNA polymerase expression vector): Studier *et al.*, 1990

pPLEX (λ P_L expression vector): Sczakiel *et al.*, 1987

pACYC184: Chang and Cohen, 1979

pLG339: Stoker *et al.*, 1982

pBR322: Bolivar *et al.*, 1978

M13mp18/19: Yanish-Perron, 1985

pUC12: Messing, 1983

pUC18/19: Norrander *et al.*, 1983

Plasmid DNA was generally resuspended in distilled water at 1μg/μl and stored at -20^{o}C.

2.1.3 Growth media:

Nutrient Broth:

per litre

Oxoid No2 Nutrient Broth	25g

Luria Broth:
per litre of distilled water add

Tryptone	10g
Yeast extract	5g
NaCl	5g
pH 7.0	

M9-Glucose Minimal Media:
per litre of distilled water add

Minimal Salts	100ml
20% w/v Glucose	50ml
5% w/v Caso Amino acids	10ml (not added when selecting for the maintenance of F' *pro*+ plasmids).
Vitamin B1 (1%)	1ml
CaMg Salts	10ml

Minimal Salts: (boil to dissolve)
per litre of distilled water add

Na_2HPO_4	60g (anhydrous)
K_2HPO_4	30g (anhydrous)
NaCl	5g
NH_4Cl	10g

CaMg Salts:

$CaCl_2$	0.01M
$MgSO_4$	0.1M

Vitamin B1 (Thiamine) was prepared in distilled water, filter sterilised and stored at -20°C.

Solid nutrient, Luria or minimal agar was prepared as above but with the addition of agar to 1.5%.

Blood Agar Plate:
Identical to Nutrient agar plates but with the addition of 5% fresh sheep blood (Oxoid) to the agar just prior to setting and pouring, to avoid lysis of the erythrocytes.

Soft Nutrient Agar:

per litre of distilled water add

Oxoid No2 Nutrient Broth	25g
Davis Agar	5g

Media was sterilised by autoclaving at 15psi for 20 minutes.

2.1.4 Maintenance of strains and induction conditions for expression of plasmid encoded genes:

The *E. coli* strains were grown with aeration at either 30 or 37°C as appropriate in Nutrient broth, Luria or M9-glucose minimal media (see above). For plasmid containing strains chloramphenicol (25μg/ml), ampillicin (100μg/ml), tetracycline (12.5μg/ml), kanamycin (25μg/ml) were added as appropriate.

Strains carrying genes under the control of *lac* or *tac* promoters were grown usually to an A^{450} of 0.4 before inducing expression with IPTG (0.5mM final concentration). Strains carrying heat inducible promoters were grown at 30°C, usually to an A^{600} of 0.4 (equivalent to ~A^{450}=0.7) before heating to 44°C, by either immersing in a 44°C shaking waterbath for 5 minutes or shaking in a 65°C waterbath using a sterile thermometer to monitor the temperature change. After reaching 44°C the culture was returned to 37°C to continue growth.

2.2 Methods associated with DNA manipulation:

2.2.1 Preparation of Plasmid DNA:

DNA used for cloning procedures was usually prepared by the CsCl method while DNA required for rapid screening of transformants was prepared by the mini-preparation method.

2.2.2 Mini-Preparation of Plasmid DNA:

The method used was a modification of the Alkaline Lysis Method (Birnboim and Doly, 1979) as described in Maniatis *et al*, (1982) except that no Lysozyme was added to solution I.

2.2.3 Phenol-Chloroform Extraction:

Proteins were removed from DNA solutions by phenol/chloroform extraction as described in Maniatis *et al*, (1982). The phenol/chloroform extraction buffer was prepared as described in Maniatis *et al*, (1982).

2.2.4 Ethanol Precipitation:

DNA was routinely precipitated from solutions by the addition of $^1/_{10}$ volume of 3M ammonium acetate and 2.5 volumes of 100% ethanol. The DNA was pelleted after leaving the solution at -70°C for 20 minutes, by spinning at 13,000 rpm for 10 minutes. The pellet was washed in 70% ethanol, dried and resuspended in sterile distilled water.

2.2.5 CsCl Preparation of plasmid DNA:

Cells from a 100ml overnight culture (heat induced for copy number amplification if necessary) were concentrated by centrifugation (8K rpm, 20 minutes), and lysed in the same manner as the above mini-preparation method, except that the volumes were increased by 100 fold. The plasmid DNA was precipitated by ethanol precipitation and the pellet resuspended in a CsCl solution (1.1 grams/ml of distilled water final volume and 300ug/ml Ethidium Bromide) and centrifuged either at 100,000 rpm for 4 hours in the Beckman Benchtop TL-100 Ultracentrifuge using the TLA-100.2 rotor, or at 48,000 for 16 hours using the Beckman 70Ti rotor in the Beckman L5-65 Ultracentrifuge. The plasmid band was removed with a needle (0.6mm bore) and 2ml syringe, transferring to an eppendorf tube, and the ethidium bromide extracted using an equal volume of CsCl saturated iso-butan-2-ol (X 4). The plasmid was precipitated directly by diluting 0.4ml of DNA solution with 0.45 ml of distilled water to 0.54ml iso-propan-2-ol spinning 10 minutes at 13,000 rpm at room temperatures. The resulting pellet was washed in 70% ethanol, desiccated and resuspended in distilled water usually to a concentration of 1μg/μl.

2.2.6 Agarose Gel Electrophoresis:

Horizontal agarose slab gels (usually 0.7% w/v) were prepared using Seakem HGT agarose in ethidium bromide free TAE buffer (267mM Tris-Acetate pH8.0; 1mM EDTA) with the addition of Ethidium Bromide (1ug/ml). DNA samples were loaded in sample buffer (1X final concentration) and the gels run at 100 - 200 volts in TAE buffer. The DNA bands were visualised using ultra violet transilluminator and a photograph taken if appropriate.

10X Sample Buffer:

0.25M Tris-HCl, pH 6.8	10ml
Glycerol (100%)	4ml
EDTA	5mM
Orange C or Bromophenol blue as dye	
Distilled water to	40ml

2.2.7 Agarose gel DNA Markers:

The markers were obtained from BRL and consisted of the lambda phage digested with different restriction enzymes. The molecular sizes of the different markers is given in the appropriate Figures.

2.2.8 Restriction Digestion:

The restriction enzymes came from BRL and digestions were carried out using the restriction buffers provided and conditions described by the manufactures. 1µl of boiled RNase (10mg/ml) was added to digests involving mini-preparation plasmid DNA.

2.2.9 Isolation of DNA fragments from Agarose Slabs:

Isolation of DNA fragments, after separation on agarose gels was carried out either by using a modification of the Glass Extraction method described by Vogelstein and Gillespie (1979) or by the "Death Wish" method.

a) Glass-milk method:

Preparation of Glass-milk:

Packard scintillation vials were finely ground, autoclaved and resuspended in distilled water (100mg/ml).

Isolation of fragments:

The DNA fragments were separated on agarose gel and the DNA band(s) of interest cut out of the agarose slab and placed in eppendorf tubes. The agarose was dissolved in two volumes of saturated sodium iodide (1814.4g dissolved in 995ml warm water, allowed to stand until cold and crystals formed, adding Na_2SO_4 as stabiliser). 30µl of glassmilk was added and mixed for at least 10 minutes. The glassmilk was spun down (5 seconds eppendorf centrifuge) and washed in 600µl of an Ethanol-TNE solution. This procedure was repeated 4 times. The final glassmilk pellet was resuspended in 30µl of distilled water and the DNA released from the glass by incubation at 50°C for 3 minutes. The glassmilk was then centrifuged down (30 second spin in eppendorf centrifuge) and the

supernatant removed to a fresh tube. One-tenth of the final DNA solution was generally run on a 7% agarose gel to collect an estimation of the quantity of the DNA fragment isolated.

Ethanol-TNE solution:	TNE:
50% ethanol	0.01M Tris-HCl pH7.2
25% water	0.1M NaCl
25% TNE	1mM EDTA

b) Death wish procedure:

DNA samples were separated on agarose gels and the band to be isolated identified when visualised with ethidium bromide. The gel is cut just below the band of interest and an appropriately sized hemi-piece of dialysis membrane inserted into the slit. The DNA fragment is run onto the membrane and the membrane rapidly removed with the current still running, to a eppendorf tube using a pair of Millipore forceps. The edge of the dialysis tubing is trapped under the cap and the tube spun briefly to obtain the DNA containing liquid. The dialysis tubing is washed with 200µl of distilled water and centrifuged again. The DNA solution is then removed from any pieces of contaminating agarose, ethanol precipitated (see above) and resuspended in distilled water. One-tenth of the final DNA solution was generally run on a 7% agarose gel to obtained an estimation of the quantity of DNA fragment isolated.

Preparation of dialysis tubing:

Approximately 3 inch lengths of dialysis tubing were boiled for 15 minutes in a 10mM Tris-HCl pH 7.5, 1mM EDTA solution. The tubing was then washed with distilled water and stored at 4°C in 50% ethanol.

2.2.10 Ligations:

Ligations were usually carried out in a total volume of 20µl using the 5X BRL Ligase buffer provided with the enzyme T4 DNA Ligase (1 unit/ligation) incubating at 16°C overnight. Sticky end ligations were usually carried out with an insert to vector ratio of 1:1 while blunt end ligations involved a ratio of 2:1 or 4:1.

2.2.11 Blunting 5' overhangs using *E. coli* DNA Polymerase Klenow Fragment:

DNA fragments with 5' overhangs, to be filled in, were resuspended in 20µl distilled water and the following components added:

3µl	10X ligase buffer
3µl	T.M
1.2µl	0.1M Spermidine
2µl	dNTPs (2mM Stock)
	1 Unit Klenow.

The samples were incubated for 30 minutes at room temperature, stopping the reaction by the addition of 1µl 0.5M EDTA. The solution was then phenol-chloroform extracted to remove the enzyme, ethanol precipitated ready for subsequent reactions.

10X ligase buffer:		T.M:
0.5M	Tris.HCl pH7.4	0.1M Tris-HCl pH 8
0.1M	$MgCl_2$	0.1M $MgCl_2$
0.1M	Dithiothreitol	
10mM	Spermidine	
10mM	ATP	
1mg/mg	BSA	

2.2.12 Blunting 5' and 3' overhangs using T4 DNA Polymerase:

In the presence of 5' overhangs T4 DNA polymerase fills in the overhang to generate blunt ended fragments, whereas in the presence of 3' overhangs the enzyme "chews" back the 3' protruding overhang.

Reaction:

17µl	DNA solution
2µl	10X T_4 buffer
1µl	dNTPs (2mM Stock)
	2.5 Units T4 DNA polymerase

Reactions were incubated for 15 minutes at 37°C, stopping the reaction by placing at 70°C for 5 minutes. Again the solution was phenol-chloroform extracted to remove the enzyme, ethanol precipitated and resuspended in distilled water.

10X T_4 buffer:

0.33M	Tris- Acetate pH7.9
0.66M	Potassium Acetate
0.1M	Magnesium Acetate
5mM	Dithiothreitol
1mg/ml	BSA

2.2.13 Dephosphorylation of DNA fragment:

The DNA fragment for dephosphorylation was resuspended in 1X CIP buffer in a total volume of 20µl with distilled water. 0.01 units of calf intestinal dephosphorylase (CIP) enzyme was added per picomole of DNA. For dephosphorylation of fragments carrying 5' protruding termini this was incubated for 30min at 37°C, and this step repeated adding fresh enzyme. For fragments carrying blunt or 3' protruding termini incubation was for 15min at 37°C, adding more enzyme and incubating for a further 15min at 56°C. Reactions were stopped by removing the enzyme by phenol-chloroform extraction (see above), the DNA ethanol precipitated and resuspending it in distilled water for subsequent reactions.

10X CIP buffer:

0.5M	Tris- HCl pH9.0
0.01M	$MgCl_2$
1mM	$ZnCl_2$
5mM	Spermidine

2.2.14 Transformation Techniques:

Transformation is the method by which DNA is introduced into a cell and the two main methods used to introduce plasmid DNA into *E. coli* were as follows.

a) $CaCl_2$ Method

This method uses the property of $CaCl_2$ to render *E. coli* sufficiently permeable to permit the entry of DNA by heat shocking. An overnight culture of an *E. coli* strain was used to innoculate 25ml of fresh nutrient broth to an A^{600} ~0.1 and grown to an optical density equivalent to an A^{600} ~0.4. The cells were harvested by centrifugation (4°C, 3000rpm for 5 minutes; MSE Centaur 1 benchtop centrifuge) and the cells gently resuspended in 5ml of ice cold 0.1M $MgCl_2$. The cells were spun down again, resuspended in 5ml of ice cold 0.1M $CaCl_2$, and left on ice for 20 minutes. After this period the cells were spun down as before and resuspended in 1ml of 0.1M $CaCl_2$ and left on ice for at least 1 hour.

100µl of these "competent" cells were mixed with 1 to 100ng of plasmid DNA and left on ice for 20 minutes before placing at 42°C for 90 seconds (heat shock). 900µl of fresh broth was then added and the cells placed at 37°C for 1 hour to allow the expression of resistance genes. Usually 100 to 200µl of these cells are plated out on agar plates containing appropriate antibiotics. Competent cells prepared in this way could be store for several months at -20°C if resuspended in 25% glycerol and snap frozen in a dry ice/methanol bath.

b) Electroporation:

This method uses high voltage to force DNA into cells (Dower *et al.*, 1988). In order to subject the cells to a high voltage it is necessary that the cells be in a solution of low electrical resistance, that is in a salt free environment, so the cells have to be washed in water. This procedure also requires the cells to be concentrated.

The strain into which the plasmid DNA was to be introduced was grown overnight. The overnight culture was used to innoculate 100ml of fresh nutrient broth to an A^{600} ~0.1 and grown to an optical density of A^{600} ~0.9. The cells were harvested by spinning (4°C, 3000rpm for 15 minutes; MSE Centaur 1 benchtop centrifuge) and washed in an equal volume of cold sterile water. The cells were repelleted and washed in $^1/_2$ and $^1/_{50}$ of the original volume, respectively, with cold water. Finally the concentrated cells were resuspended in $^1/_{500}$ of the original volume with cold water. 40µl of the concentrated cells were mixed with 1 to 100ng of plasmid DNA and subjected to 1500 volts (capacitance dial set at 25µF) in a cuvette with a 2mm gap between the electrodes. 1ml of SOC medium was added immediately and the cells placed at 37°C with shaking for 1 hour to allow the expression of resistance genes, and 100-200µl plated out on agar plates containing appropriate antibiotics.

Note: The DNA has also to be present in a salt free environment, consequently ligation mixtures were ethanol precipitated, with the addition of 1µg of tRNA to ensure quantitative precipitation of DNA prior to electroporation.

SOC: per litre	
Bactotryptone	20g
Yeast extract	5g
1M NaCl	10ml
1M KCl	2.5ml

Add together to 970ml of distilled water, dissolve and autoclave. Add filter sterilised solutions of

$MgCl_2$	5ml
$MgSO_4$.	5ml
2M Glucose	10ml

2.2.15 Isolation of M13 transformants:

E. coli strains were transformed with the required M13 double stranded clone by either $CaCl_2$ or electroporation method (see 2.2.14). 100 to 200µl of the transformed cells were then plated out immediately by adding to 3ml of soft nutrient agar (kept at 42°C in a heating block), seeded with a 100µl of overnignt stationary phase F' carrying *E. coli* cells. The mixture was vortexed briefly and poured on top of a minimal media agar plate (selects for F'). After incubating overnight at 37°C individual plaques were transferred to 1ml Lambda Buffer with a tooth pick and stored at 4°C as phage stocks.

Lambda Buffer: (per liter)

1M Tris-HCl pH7.2	6ml
$MgSO_4.7H_20$	2.46g
Gelatin	0.05g

2.2.16 Isolation of Single Stranded DNA:

Single stranded DNA was obtained by using 100µl of phage stocks to infect 1.5ml of nutrient broth (seeded with 15µl of an overnight NM522 culture), growth being maintained with vigorous shaking in a 20ml test tube at 37°C for $5^1/_2$ hours. The cells were spun down (13K rpm, 10min) and the phage precipitated from 1ml of the supernatant by the addition of 200µl of 20%PEG (Polyethelene Glycol 6000 -Sigma)/ 2.5M NaCl, mixing and leaving at 15°C for 15 minutes. The phage was pelleted by spinning at 13K, 15min in an eppendorf centrifuge and the supernatant completely removed using a drawn out Pasteur pipette and suction. The pellet was resuspended in 100µl distilled water and the protein coat removed by phenol-chloroform extraction. The single stranded DNA, after the addition of 10µl of 3M sodium acetate, was then ethanol precipitated and resuspended in 20µl distilled water.

2.2.17 DNA Sequencing:

a) M13 single stranded sequencing:

Fragments to be sequenced were first cloned into the double stranded replicative form of M13 and used to generate single stranded DNA template for sequencing by the dideoxyribonucleotide method pioneered by Sanger *et al* (1977).

Sequencing Reactions:

i) Annealing of Primer to Single Stranded Template:

5µl of single stranded template was mixed with 0.8ng of an appropriate 17 to 21 mer oligonucleotide primer and 1µl of TM buffer (0.1M Tris-HCl pH8; 0.1M MgCl) in a total volume, with distilled water, of 10µl in capped eppendorf tubes. Annealing of primer to template was carried out by incubating at 65°C for 30 minutes, spinning the tubes briefly to collect the condensate. This procedure was then repeated.

ii) Labelling Reaction:

For each clone four tubes were taken and labelled either A, T, G or C and 2µl of the appropriate labelling mix (see below) aliquoted together with 2µl of the annealed template solution.

For each clone 0.43µl of Klenow polymerase (1 unit/µl) was added to 6.96µl of distilled water and 1µl of ^{35}S dATP (10µCi/µl; 400Ci/mmol). 2µl of this mixture was aliquoted into each reaction tube and after gentle mixing incubated for 20 minutes at 37°C. 2µl of sequence chase mix (0.25mM of each dNTP in TE) was added and incubation continued for another 20 minutes. The reaction was stopped by the addition of 4µl of formamide dye per tube (10ml formamide; 10mg xylene cyanol FF; 10mg bromophenol blue and 0.2ml 0.5M EDTA pH8). Reactions were run immediately or stored at -70°C for up to one week.

LABELLING MIXES: (store at -70°C)

	"T"	"C"	"G"	"A"
0.5mM dTTP	12.5	250	250	250
0.5mM dCTP	250	12.5	250	250
0.5mM dGTP	250	250	12.5	250
10mM ddTTP	6.2			
10mM ddCTP		4		
10mM ddGTP			8	
10mM ddATP				1.2
TE Buffer	500	500	500	500

Oligonucleotides: Various oligonucleotides specific for the *hlyA* gene were available and used as primers as well as primers specific for the M13 DNA (purchased from BRL).

iii) Acrylamide Sequencing Gels:

0.2mm gradient acrylamide gels were run using plates of dimensions 39 X 20 X 0.2cm glass or 42 X 33 X 0.2cm plates. 0.2mm combs with wells for 6 or 12 clones were used as appropriate. Gradient acrylamide gels were made from 0.5 and 2.5 times stocks.

0.5 X Stock		2.5 X Stock
17.1g	Acrylamide	2.28g
0.9g	Bis-acrylamide	0.12g
150g	Urea	20g
-	Sucrose	2g
-	Bromophenol blue	a pinch
15ml	10X TBE	10ml
	Make volume up to	
300ml	with distilled water	40ml

Acrylamide solutions were filtered through 2 pieces of Whatman 1 filter paper before being stored at 4°C in the dark

10X TBE Buffer:

109g	Tris base
55g	Boric acid
9.3g	Na_2EDTA

Make up to 1 litre. Check pH is 8.3

iv) Pouring Acrylamide gradient gels:

The following quantities were used to pour either the large or half volumes for the smaller acrylamide gels:

For the large plates 75ml of the 0.5X stock was mixed with 350µl of Ammonium persulphate (APS) and 24µl of TEMED while 20ml of 2.5X stock was mixed with 175µl of APS and 12µl of TEMED. Now 8ml of the first solution was mixed with 12ml of the 2.5X solution and poured to the bottom of the gel plates followed gently by the rest of the 0.5X solution until the plates were full and the combs inserted. The set gel was pre run (see below) for 20-60 minutes, with 0.5X TBE in the top tank and 2.5X TBE in the bottom tank. Extension gels were identical to the above except

that 14ml of 0.5X solution was mixed with 6ml 2.5X solution when pouring the gel, with 1.5X TBE instead of 2.5X buffer in the bottom buffer tank.

The sequencing reaction tubes were heated for 2 minutes at 80°C and 2-3µl of each sequencing reaction loaded. The gel was run between 1000-2000 volts maintaining the temperature of the plates at 40-50°C. After running the gel was fixed in 10%v/v methanol / 10%v/v acetic acid for 20 minutes before drying onto 3MM Whatman paper and exposing to a sheet of autoradiographic film at room temperature.

b) Double stranded sequencing:

i) Double stranded plasmid DNA was isolated by the CsCl method. 4-5µg of plasmid DNA was precipitated and resuspended in 50µl of a fresh solution of 0.2M NaOH, 0.2mM EDTA, incubating for 5 minutes at room temperature to denature the DNA. The solution was neutralised by the addition of 5µl 2M Ammonium acetate pH4.5 and precipitated with two volumes of absolute ethanol. The resulting pellet was washed in 70% ethanol and dried ready for sequencing.

ii) The prepared plasmid DNA was sequenced using the USB sequenase kit following the manufactures protocol. Sequencing gels made as described above were used to separate the products of the reactions.

Oligonucleotides: Various oligonucleotides specific for the *hlyA* gene were available and used as primers as well as other primers specific for regions within the pUC and pBR322 plasmids (purchased from BRL).

2.2.18 Site directed mutagenesis:

Oligonucleotides required to direct specific mutation events were synthesised by J. Keyte (Dept. of Biochemistry, Leicester university). The Amersham site directed mutagenesis kit was used to introduce the mutations into the *hlyA* gene as described by the manufactures.

2.2.19 Hydroxylamine mutagenesis reaction:

The method used was basically that described by Eichenlaub (1979). The DNA template was prepared by the CsCl method and specific DNA fragments isolated by the "death wish" method (see above).

Reaction:

DNA (1μg)	10μl
10X TES	10μl
0.1M Na Phosphate	40μl
1M Hydroxylamine pH6	40μl

Incubate 70°C 30minutes

Samples were then dialysed against 4 L 1X TES for 2 hour at 4°C, and the procedure repeated with fresh buffer. Finally, the samples were dialysed in 4 L 1X TES overnight. The plasmid DNA was ethanol precipitated and the hydroxylamine treated fragment was run in an agarose gel (7%) before glassmilk isolation to remove any trace of contaminating fragments.

2.2.20 Colony Hybridisation:

The procedure used was a modification of that developed by Grunstein and Hogness (1975). The colonies on an agar plate to be screened were first transferred to nitrocellulose filters (0.45μm Cellolosenitrat(e) purchased from Schleicher and Schuell). Then the cells were lysed and DNA denatured by incubating the filters in 0.5M NaOH and 0.1M NaOH/1.5M NaCl for 7 and 10 minutes respectively. The NaOH was neutralised with two 2 minute washes of 1M Tris-HCl pH 7.4, followed by immersion in 0.5M Tris-HCl pH7.5 / 1.5M NaCl for 4 minutes. The DNA was baked onto the filters at 80°C for 2 hours.

The probe DNA fragment was isolated by the glassmilk method (see above) purified and rerun down a second gel to remove minor contaminating fragments. The probe fragment was re-isolated and labelled with γ-dCTP by the method of oligo-labelling taken from Feinberg and Vogelstein (1984).

OLB	3μl
BSA	1.2μl (10μg/ml enzyme grade)
DNA	10μl boil 3min leave 37° 10'

before adding if double stranded

γ-dCTP	1.5μl (10μCi/μl; 3000Ci/mmol)
Klenow	0.3μl (1 unit/μl)

Incubate for 30 minutes at room temperature and stop the reaction with 75μl stop buffer (20mM NaCl, 20mM Tris-HCl pH7.5, 2mM EDTA; 0.25%w/v SDS, 1μM dCTP)

OLB: made from the following components

Solution A: (store at -20°C)

625µl	2M Tris-HCl pH8
25µl	5M $MgCl_2$
350µl	Water
18µl	2-mercaptoethanol
5µl	each of dATP,dTTP,dGTP (each dissolved in 3mM Tris-HCl pH8, 0.2mM EDTA pH7.0 at a concentration of 0.1M)

Solution B: (store at 4°C)

2M	HEPES pH6.6 (with NaOH)

Solution C: (store at -20°C)

Hexadeoxyribonucleases (Pharmacia) evenly suspended in 3mM Tris-HCL, 0.2mM EDTA pH 7.0 at 900 units/ml.

Mix the solutions A,B and C in the ratio of 2:5:3 to make OLB. Store at -20°C

Hybridisation conditions:

The filters were washed in 3X SSC / 0.1% SDS and left overnight in fresh 3X SSC / 0.1% SDS solution. A 1 hour prehybridisation preceded the hybridisation with the labelled probe - boiled for 3 minutes - overnight at 65°C in Marvel milk solution (1.5X SSPE; 0.5% Marvel milk; 1% SDS and 6% PEG 8000). 3X SSC / 0.1% SDS solution was used to wash the filters 4 times at 65°C for 5 minutes and the final high stringency washes - two 30 minute washes - were in 0.1X SSC / 0.01% SDS at 65°C. Autoradiography was preceded by the drying of the filters at room temperature.

SSC X 20: 3.0M NaCl; 0.3M Trisodium citrate

1.5X SSPE: 0.27M NaCl; 15mM Na Phosphate pH7.7 and 1.5mM EDTA

2.2.21 Radioactivity:

All radioactive isotopes were purchases from Amersham Ltd.

2.2.22 Computer Programmes:

The Winsconsin package of DNA and protein manipulation programmes available on the Leicester Vax system were widely used to aid these studies. Also a suite of 8 secondary structure prediction programmes (E.

Eliopoulos, Dept. of Biophysics, University of Leeds) available on the Leeds BioVax system was used extensively in the analysis of mutations on the predicted structure of the HlyA molecule.

The Harvard Graphic and Coreldraw packages were used to generate the diagramatic Figures and plasmid maps in this thesis using an IBM Personal System 2 Model 50.

2.3 Methods for Protein Work:

2.3.1 Determination of culture optical densities:

Samples of overnight cultures were generally dilute 1 in a hundred in 1ml of fresh broth before taking an optical density reading using a Gilford micro-sampling spectrophotometer. The machine was first zero'd with fresh broth and the samples diluted to keep the reading less than 1 to obtain a linear response.

2.3.2 Separation of Cellular and Supernatant Protein Samples:

At appropriate times samples, indicated in the text, samples were taken and centrifuged at 15,000 rpm for 10 minutes in the SS34 rotor to pellet the cells. The cellular pellet was resuspended directly in SDS-Sample buffer. The supernatant was transferred to a fresh tube (avoiding contamination with whole cells) and the proteins precipitated with TCA (Trichloroacetic Acid added to a final concentration of 10%). This mixture was incubated on ice for 1 hour and the precipitated proteins harvested by centrifugation in an HB4 rotor (10,000 rpm, 15 minutes). The tube was inverted and the pellet allowed to drain. The pellet was solubilised in SDS-sample buffer (containing 10% saturated tris to neutralise the TCA). The cellular pellets were generally resuspended in 40µl of sample buffer / A_{450} Unit of cellular sample. Concentrated supernatant samples were resuspended in 4µl of sample buffer / A_{450} equivalent unit of cellular sample. Samples were analysed on SDS-polyacrylamide gels following incubation in boiling water for 3 minutes.

Sample Buffer (2X):

0.125M	Tris-HCl pH6.8
20%v/v	Glycerol
4% w/v	Sodium Dodecyl Sulphate
0.05% w/v	Bromophenol Blue
10% v/v	Beta-Mercaptoethanol

2.3.3 SDS-Polyacrylamide Gel Electrophoresis:

The procedure used was that of Laemmli (1970) using either the Biorad Protean II system or small mini-gel kits. The mini-gel glass plates measured 80mm X 100mm by 0.2mm wide and were poured as the large kits but using $^1/_{10}$ the volumes (see Table). Samples were boiled in the sample buffer for 3 minutes prior to loading and running at 30mA until the dye front reached the end of the gel.

Separating Gel:

	Percentage of gel			
	9%	11%	15%	18%
Buffer A				
(0.75M Tris-HCl pH8.8 0.2% w/v SDS)	13.5	13.5	13.5	13.5
Acrylamide (44:0.8 w/v acrylamide: bisacrylamide)	5.6	6.8	9.2	11.1
H_2O	7.3	6.0	3.6	1.7
Ammonium Persulphate (10mg/ml)	All 0.95			
TEMED (N,N,N',N-tetramethyl ethylenediamine)	All 75µl			

	7%
Stacking Gel:	
Buffer B	
(0.25M Tris-HCl pH6.8 0.2% w/v SDS)	10
Acrylamide (44:0.8 acrylamide: bisacrylamide)	3.3
H_2O	6.7
Ammonium Persulphate (10mg/ml)	0.5
TEMED (N,N,N',N-tetramethyl-ethylenediamine)	40µl

Volumes are in millilitres unless stated otherwise.

2.3.4 Western Blotting:

Protein samples were separated on SDS-Polyacrylamide gels and transferred to nitrocellulose (0.45μm Cellulosenitrat(e) purchased from Schleicher and Schuell) using the method of Towbin et al (1979). However, the transfer buffer did not contain methanol. The Protein bands were first of all visualised by staining with Ponceus S solution (purchased from BDH) and the positions of molecular weight markers indicated, the stain was then removed by washing in water.

The nitrocellulose filter was blocked overnight in Phosphate buffered saline solution pH7.2 or Tris buffered saline solution (see overleaf) usually containing 5% Marvel Milk Powder but could be substituted for by either 0.5% Tween-20 or 1% BSA. Incubation with the primary antibody (isolated from rabbits) in saline buffer (see below) was for 1 hour (with very slow shaking motion) followed by 3-5 washes in saline buffer in a 15 minute period. The secondary antibody (goat anti-rabbit biotin conjugate: BRL) was added at a $^{1}/_{1000}$ dilution. This was incubated and washed as above before the tertiary antibody (streptavidin-horseradish peroxidase conjugate) was added at a $^{1}/_{1000}$ dilution. This was again incubated for 1 hour and washed as above.

The Western blot was developed by the addition of 30mg of 4-chloro-1-naphthol dissolved in 10ml of methanol and mixed with 50ml of 50mM Tris-HCl pH 7.5, followed by 30 to 100μl of hydrogen peroxide. The reaction was stopped by removing the solution and washing with water. Overnight incubation in water was found to increase the intensity of the colour reaction.

Antibodies:

Rabbit anti alpha-haemolysin antibodies were provided by N. Mackman and used in $^{1}/_{1000}$ dilution while antibodies raised against the C-terminal 23KD HlyA peptide were provided by R. Haigh and used in $^{1}/_{100}$ dilution.

Rabbit prochymosin antibodies were a gift provided by G. Yarranton (Celltech, Slough) and used in $^{1}/_{1000}$ dilution.

10X PBS:

per litre of distilled water

NaCl	80g
KCl	2g
Na_2HPO_4	14.4g
KH_2PO_4	2.4g
pH 7.2	

10X TBS:

Tris-HCl pH 7.5	30.285g
NaCl	43.83g

2.3.5 Inclusion body Preparation:

Separation of soluble and insoluble fraction of total cell extracts was a method modified from that of Schoemaker *et al.*, 1985.

1ml cell culture samples were centrifuged in an Eppendorf centrifuge (13K rpm, 1 minute) and the supernatant discarded. The cells were resuspended in 450µl lysis buffer (100mM Tris-HCl pH 6.5, 5mM EDTA, lysozyme 260µg/ml, 1mM PMSF) and incubated on ice for 5 minutes. Sodium deoxycholate was added to 0.1% (w/v) and incubated for a further 5 minutes on ice. DNase (25µg/ml) and $MgCl_2$ (10mM) were added and again incubated for a further 5 minutes. The lysed cells were centrifuged for 10 minutes at 4°C, the supernatant contained the soluble proteins. The pellet was washed twice in 100mM Tris-HCl pH 6.5, 5mM EDTA and consists of the insoluble proteins. Samples of these fractions solubilised with equal volumes of 2X sample buffer were analysed on acrylamide gels following incubation in boiling water for 3 minutes.

2.3.6 Haemolytic Assay:

The secreted haemolytic activity was measured by the release of haemoglobin from sheep red blood cells (Oxoid).

The red blood cells were centrifuged in eppendorf tubes (13K, 30 seconds) and washed 3 times in blood buffer (155mM NaCl, 5mM KCl, 20mM $CaCl_2$, 2mM $MgCl_2$, 10mM Tris-HCl pH 7.4) before resuspending to a final 5% concentration.

1ml culture samples were taken at appropriate intervals and the cells spun out (13K, 2 minutes). 100µl of the supernatant was taken and added to 900µl of the 5% blood assay solution. This was incubated at 37°C for 10 minutes and the unlysed red blood cells and cellular debris centrifuged down (13K, 1 minute). 700µl of this supernatant was placed in a 1ml acryl-cuvette (Sarstedt No.67.740) and the absorbency measured at A^{543}. The assayed sample volume was reduced when appropriate

(making up the volume with fresh nutrient broth) to maintain the reading less than 1 as this gives a linear response (Gray, 1987). The spectrophotometer was adjusted to zero for each sample by incubating 100µl of nutrient broth with 900µl of the blood assay mixture to determine the background release of haemoglobin. Samples were always carried out in duplicate and the average result given.

2.3.7 Quantification Methods:

a) Quantification via ^{3}H Leucine Labelling experiments:

i) ^{3}H Leucine Labelling reactions:

Cultures containing appropriate plasmids were grown in M-9 minimal medium (minus caso amino acids- Difco Laboratories) supplemented for all required amino acids (250µg/ml final) except for leucine. Cultures were grown and plasmid gene expression induced as described in text. Growth was continued for 30 minutes before the addition of 50µCi of ^{3}H-leucine and 1µg/ml of cold leucine (to ensure gradual uptake of the isotope) to 10ml of culture. Protein synthesis and growth were arrested after 30 minutes by the addition of chloramphenicol (250µg/ml) and cold leucine (100µg/ml). Cell and supernatant samples were taken and prepared as described above for analysis by SDS-PAGE.

ii) Estimation of incorporation of counts by scintillation counting:

Generally a one hundredth of the cellular sample was placed on two individual filters (Millipore 4.5µM filter) and one sample washed 4 times with 20ml of cold TCA (10% w/v). The other sample was left untreated. Both filters were dried ready for scintillation counting. In addition a one hundredth of the culture supernatant sample was applied to two filters and treated as above.

Due to the nature of the precipitated supernatant protein sample it was first solubilised in sample buffer (containing 10% saturated tris to neutralise the TCA used for concentration). Now a similar sample volume was removed and placed in 10% TCA before transferring to filters. Again only one sample was washed with cold TCA as before and the filter dried. All samples were carried out in duplicate.

The dried filters were placed in plastic inserts, appropriate for packard scintillation vials, and 5ml of scintillant fluid added (0.5% PPO, 0.03% POPOP in toluene), placed in scintillation vials and counted using the

Packard Tri-Carb scintillation counter. Each sample was counted for 1 minute in duplicate on the ^{3}H detecting channel.

Estimation of eluted counts:

Cell and supernatant samples were separated by SDS-PAGE and the gels stained with Coomassie blue and dried down. The bands of interest were identified and excised from the gel and placed into scintillation counting inserts (see above). The bands were rehydrated with 50µl distilled water and left for 10 minutes at room temperature. 5ml of NCS scintillation fluid was added, vortexed and incubated overnight at 37°C. The samples were revortex before counting for 1 minute using the Packard scintillation counter.

NCS scintillation fluid:

NCS solubiliser	83.5g
PPO	4g
Dimethyl POPOP	60.5mg
Toluene	1000ml

b) Quantification by Laser Scanning:

Cell and supernatant samples were separated by SDS-PAGE and the gels stained with Coomassie blue. The gels were either dried down between two sheets of cellulose (purchased from Hoefer Scientific Instruments Cat No. SE1142) or transferred directly to the LKB ultrascan laser densitometer for scanning. Each track was scanned and the output for each detected bands expressed as a percentage of total absorbency for that lane or as a relative absorbance value. Thus a supernatant band could be related to an intracellular band (whose value as a percentage of the total cellular profile is given in the output) from the corresponding cellular track, and thus its level as a percentage of total cellular protein can be estimated, by comparing relative absorbancy values.

2.3.8 Protein Markers:

Both low and high molecular weight protein markers were purchased from BRL and the sizes are indicated in the appropriate Figure.

CHAPTER 3

Secretion of heterologous proteins using the haemolysin secretion system

3.1 Introduction:

The haemolysin secretion system is unique in its ability to specifically transfer the haemolysin toxin (a 107KD polypeptide) from the cytoplasm of *E. coli* across its two enclosing membranes and into the external medium. It had already been shown that the C-terminal 23KD domain of this toxin is sufficient to promote its own specific secretion via the haemolysin export channel (Nicaud *et al.*, 1986). Here, we attempt to address the problem of whether the haemolysin C-terminal 23KD domain encodes all the information required for the secretion of the entire toxin molecule or whether further N-terminal signals are also required. To do this we ask whether this 23KD domain can promote the secretion of heterologous proteins. These experiments will also test the specificity of the secretion system and ultimately provide information on the mechanism of secretion.

The C-terminal 23KD signal domain of HlyA had already been used, unsuccessfully, in preliminary attempts by others in this laboratory to secrete portions of the *E. coli* cytoplasmic proteins - chloramphenicol acetyltransferase and beta-galactosidase (unpublished results; this laboratory). Both of these proteins are not normally translocated across membranes, and although subsequent studies indicated that secretion could in fact be achieved, at the onset of this study it was felt more appropriate to investigate secretion of a foreign protein which is normally translocated across membrane bilayers *in vivo*. The protein chosen for this study was Calf Prochymosin (Pcm), the major milk clotting enzyme of the fourth stomach of unweaned calves. Chymosin is synthesised in a prepro form, the secreted zymogen (prochymosin) being irreversibly converted by a complex autocatalytic process to active chymosin under acidic conditions (Pedersen *et al.*, 1979).

It should be acknowledged, that at the same time as this work was initiated, Mackman *et al.*, (1987) demonstrated that the C-terminal 23KD domain of HlyA when fused to the majority (92%) of the *E. coli* outer membrane porin (OmpF), deleted for its N-terminal signal sequence, successfully promoted the secretion of this chimeric polypeptide to the medium, in an HlyB,D dependent manner. This illustrated the capacity of the haemolysin system to secrete heterologous proteins specifically through its export "channel" when fused to the HlyA secretion signal domain. Nevertheless, the ability to secrete heterologous, non *E. coli* proteins had still to be addressed.

Due to the apparent failure to secrete portions of the *E. coli* cytoplasmic proteins beta-galactosidase and chloramphenicol transacetylase (N. Mackman personal communication), it was considered appropriate to construct Pcm-HlyA fusions progressively in case particular regions were inhibitory to the haemolysin secretion pathway.

3.1.1 Prochymosin and the Dual Origin Vectors:

Initial difficulties in isolating *pcm*-23KD*hlyA* fusions using the pTTQ series of vectors (Stark, 1987) were encountered, apparently resulting from plasmid deletion during construction and isolation of transformants. The pTTQ expression vector is a high copy number plasmid carrying a multiple cloning site downstream of a *tac* promoter, expression being regulated by the *lacI*q gene present on the same plasmid. However, this vector contains a small duplication of the *lac* promoter region providing opportunities for intramolecular recombination (N. Mackman, personal communication). The observed deletions were presumably due to deleterious effects of low levels of the chimeric protein arising from basal level expression from the *tac* promoter. Consequently, an alternative vector was sought in a bid to reduce these undesired effects. The dual origin vector (DOV) was one such plasmid system and was kindly provided by Celltech Ltd (Slough). The dual origin vectors have been described elsewhere (Yarranton *et al.*, 1984) but the principle features of the vector are as follows. Firstly, the vector carries two origins of replication, a low copy number pSC101 origin together with its partition function (*par*) and a 'hybrid' *Col*E1 based origin, whose function is controlled by the lambda P_R promoter. The vector also carries the lambda temperature sensitive repressor gene (cI^{857}), a selectable bacterial resistance gene and a constitutive tryptophan promoter (*trp*) from which the expression of the cloned gene is directed. This arrangement allows the vector to be maintained at 3-4 copies per cell at low temperature, but upon heat shock the cI repressor protein is rendered inactive allowing extensive transcription from the lambda P_R promoter through the adjacent ColEI *ori* region. This consequently leads to frequent initiation of new rounds of replication resulting in copy number amplification, to at least 200-300 copies per cell. This amplification leads in turn to maximal expression of the protein under the control of the *trp* promoter, which at high copy number is presumably completely derepressed. The maintenance of the plasmid at low copy number facilitates the cloning of genes, whose products if expressed might be deleterious to the cell. The

results of high level expression are then monitored after heat induction to amplify copy number. The prochymosin gene (*pcm*) was previously cloned for optimal expression into the dual origin vector (Wright *et al.*, 1986) under the control of the *trp* promoter to create pMG168 (Figure 1).

3.2 Construction of an N-terminal Prochymosin-23KDHlyA Chimera:

The complete sequence of the prochymosin gene is known (Harris *et al.*, 1982) and from this the restriction enzyme sites can be inferred. The sites most relevant to this study are shown in Figure 1. The *hlyA* 1.6Kb EcoRI/HindIII fragment (encoding the C-terminal 23KD of HlyA and an N-terminal fragment of HlyB), was previously cloned in front of the *tac* promoter of the vector pTTQ18 (Stark, 1987) generating pLG609 (see Figure 2a; Nicaud *et al.*, 1986). A series of synthetic oligonucleotide linkers were inserted into the unique EcoRI site of pLG609, to create a unique SmaI/XmaIII site, generating a series of vectors encoding the 23KD portion of HlyA and allowing the isolation of the encoding fragment in all three reading frames (Figure 2b), designated pLG609-1/ -2/ -3. Another similar vector, pLG609-4, was also generated by the insertion, this time, of a synthetic oligonucleotide linker carrying a BamHI site (see Figure 2b).

The prochymosin gene carries a unique SmaI/XmaIII site (XmaIII has the same recognition sequence as SmaI but generates a sticky rather than blunt end) 456bp from the initiator ATG triplet, which was used for the construction of the first fusion to *hlyA*. This was a convenient site, also, in that it permits the fusion of only the N-terminal (42%) portion of Pcm, a sizeable segment of foreign polypeptide (15KD), to the 23KD HlyA signal domain, to test the versatility of the haemolysin secretion system to translocate foreign polypeptides. The *pcm*-Xma-23KD*hlyA* fusion (pLG800-1) was generated as described in Appendix 1 and the resulting construct depicted in Figure 3. The 3' portion of *hlyA*, encoding the C-terminal 23KD of the toxin, was isolated on an XmaIII/ HindIII fragment from the pLG609-1, -2, -3 series of vectors (Figure 2b) and used to generate fusions to the XmaIII site of prochymosin in all three reading frames. This was done to aid the construction of future *pcm-hlyA* fusions (see below). As the host vector, pMG168 (DOV-Pcm), carries two XmaIII sites and several HindIII sites, the required vector fragment had to be isolated in two sections, resulting in a final three way ligation in order to construct the required fusion (see Appendix 1). The resulting plasmids,

carrying the 3' end of *hlyA* fused in all three frames to the XmaIII site within prochymosin, were named pLG800-1/ -2/ -3, with the inframe fusion predicted to be encoded on pLG800-1.

3.2.1 Identification of the in phase Pcm-Xma-23KDHlyA hybrid polypeptide:

Expression of the hybrid polypeptide was analysed by transforming the plasmid pLG800-1 (*pcm*-Xma-23KD*hlyA*) into *E. coli* DH1. A culture was inoculated and grown in nutrient broth to an A^{600} of approximately 0.4 before high level expression of the hybrid was induced by rapidly heating the cultures to 44°C, via a 65°C waterbath, and then continuing growth at 37°C, as described in Materials and Methods. Cell samples were taken before and at 1 and 3.5 hours post heat induction of expression, resuspended in sample buffer and analysed by SDS-PAGE. It should be mentioned here that the predicted 23KD HlyA peptide migrates at about 27KD on SDS-acrylamide gels and thus the expected size of the hybrids are correspondingly higher. The gels were analysed for an over expressed band, of the predicted 38KD in size, by Coomassie blue staining (Figure 4) and by Western blot analysis (data not shown, but see Figure 7). Figure 4 lanes 1-3 reveal the induction of an approximately 38KD band while Western blot analysis probing with both antibodies against HlyA and Pcm confirmed its identity (data not shown). In contrast, the product of pLG800-2 (a predicted out of phase Pcm-Xma-23KDHlyA fusion) was only detected, as expected with antibodies directed against prochymosin (data not shown). It should be noted here that expression levels from these vectors is very high (Yarranton *et al.*, 1984), and can represent up to 30% of cellular protein usually resulting in the formation of insoluble inclusion bodies (Schoemaker *et al.*, 1985). The expression of pLG800-1 led to the accumulation of high levels of the hybrid protein, the majority of which is in an insoluble form as determined by its recovery in an inclusion body preparation (see Materials and Methods; Figure 4, Lanes 4 and 5).

3.2.2 Secretion of the Pcm-Xma-23KDHlyA chimera by the haemolysin export mechanism:

The dual origin vector pLG800-1 (*pcm*-Xma-23KD*hlyA*) was transformed into *E. coli* SE5000, with or without (+/-) pLG575 (*hlyB,D*), and maintained at 30°C to avoid induction of expression. Exponentially growing cultures were heat induced at A^{600} of 0.4 as described above (see Materials and Methods). Again cell samples were taken before, 90 and 180 min after induction. Supernatant samples were also prepared at the same time, by centrifugation of the culture sample at 15K rpm for

10min to remove all cells. The secreted proteins were then precipitated by the addition of TCA (10% final concentration) and placed on ice for 1hour. The precipitated proteins were pelleted by spinning at 10K rpm for 15min and the pellet resuspended in sample buffer containing 10% saturated Tris to neutralise the acidic pH. An SDS-acrylamide gel was loaded with 0.5 OD units (A^{600}) of cells and 10 times that of an equivalent OD of supernatant samples, the proteins were then separated and stained with Coomassie blue stain. Figure 5 (Lanes 1-6) clearly shows the visible accumulation of an approximately 38KD polypeptide in the cellular samples of strains with or without pLG575 (*hlyB,D*), but only after heat induction. However, comparison of the supernatant samples lanes (Lanes 7-9 versus 10-12), demonstrates the specific export of a 38KD band only in the samples derived from the strain expressing the export proteins HlyB,D (Figure 5, Lanes 10-12). The identity of this polypeptide was confirmed by Western analysis using antibodies raised against haemolysin (see Figure 7, Lanes 1-4). The secreted polypeptide can be seen to exist in two forms. This is thought to be an effect of the TCA (used to concentrate the supernatant samples) which may be promoting some form of acid cleavage at the N-terminus of prochymosin, at a position close to the normal point of autocatalytic cleavage, acidic conditions being a requirement for catalysis of prochymosin to chymosin (Emtage *et al.*, 1983).

3.3 Can the haemolysin C-terminal secretion signal domain promote the secretion of larger portions of the prochymosin polypeptide ?:

Since the previous experiment demonstrated that the haemolysin export system could be utilised to secrete at least the N-terminal portion of prochymosin, further experiments were designed to resolve whether larger segments of the prochymosin molecule could be secreted in a similar manner. This increase in size, raises both the potential of the prochymosin moiety to take up any natural secondary or tertiary structure and might also result in the introduction of translocation 'blocking' sequences.

3.3.1 Generation of fusions of 73% and 99% of Pcm to the 23KD C-terminal domain of HlyA:

Figure 3 shows the resulting constructs (pLG801 and pLG802) described in this section. The fusion of the N-terminal 73% of Pcm to the 23KD portion of HlyA was achieved by using the unique ApaI restriction site in *pcm* (see Figure 1), removing the 3' overhang with T_4 DNA polymerase

(see Materials and Methods), followed by ligation to the correct frame of the 3' of *hlyA*, isolated now from pLG800-2 (see section 3.2 and Appendix 2 for details). The successful creation of an in phase hybrid polypeptide was confirmed by Western blot analysis, probing with anti-HlyA antibodies, of the SDS-PAGE cellular protein profile of a strain carrying pLG801 (*pcm*-Apa-23KD*hlyA*), induced for expression, as before. This resulted in the detection of a single band of the expected molecular weight, 50KD, (data not shown, but see Figure 7, Lanes 5-8).

The construction of the third and final fusion of the 3' end of *hlyA* to the *pcm* gene lacking only the final 11bp (including the stop codon triplet), involved the use of the third BalI site within the *pcm* gene (see Appendix 3). This construction was more difficult, since this required the isolation of three separate DNA fragments and a final three way ligation. The details of the construction are shown in Appendix 3. In view of anticipated difficulties in isolating the correct clone in this case, the resulting transformants of the ligation were screened first by colony hybridization. For this purpose, a small fragment (280bp ApaI/BalI), encoding the 3' region of *pcm* was used as the probe (see Materials and Methods). This analysis confirmed that the vast majority of the transformants, surprisingly, carried the insert (data not shown). The nature of some of these transformants was verified by heat inducing the expression of eight of the transformants, followed by analysis of total cell lysates by SDS-PAGE and Western blot analysis using antibodies directed against HlyA. This identified that 7 out of 8 transformants expressed the expected polypeptide of ~60KD in size (data not shown; but see Figure 7, Lane 9). This new construct was named pLG802 (*pcm*-Bal-23KD*hlyA*) - see Figure 3 and Appendix 3.

3.3.2 Can the haemolysin export machinery secrete 73 or 99% of the Pcm molecule when fused to the C-terminal 23KD signal domain of HlyA ?:

Both of the plasmids, pLG801 (*pcm*-Apa-23KD*hlyA*) and pLG802 (*pcm*-Bal-23KD*hlyA*) were transformed into *E. coli* SE5000 +/- pLG575 (*hlyB,D*). The resulting cultures were induced, at an A^{600} ~0.4, for high level expression of the hybrid proteins and total cell and supernatant protein samples obtained and analysed as described before (see Materials and Methods). The resulting Coomassie blue stained SDS-acrylamide gels are shown in Figure 6a/b with Figure 6a (Lanes 1-6) and Figure 6b (Lanes 1-8) showing the cellular accumulation, after induction, of a 50KD and a 60KD hybrid band, respectively, in either the cellular samples

carrying or lacking pLG575 (*hlyB,D*). Figure 6a (Lanes 7-9 versus 10-12) and 6b (Lanes 9-12 versus 13-16) demonstrates the specific secretion, respectively, of both the 50KD and 60KD polypeptides into the culture supernatant only from the strains that express the export proteins HlyB and HlyD. It should be noted that the non-specific background smear of peptides visible in the concentrated supernatant profiles originate from the nutrient broth (see Chapter 4: Figure 4). The identity of both of these secreted Pcm-23KDHlyA polypeptides was confirmed by Western analysis using anti-haemolysin antibodies (see Figure 7, Lanes 5-8 and 9, respectively). Again it should be emphasised that two secreted forms of each Pcm-HlyA hybrid peptide can be resolved, under appropriate gel conditions into two bands differing in molecular weight by 2-4KD (The two "50KD" Pcm-HlyA secreted hybrid bands are not resolved on this gel), apparently due to non-specific cleavage under acid conditions used to concentrate the supernatant samples (see Section 3.2.2).

3.4 Analysis of the possible folding of the Pcm domain of the secreted Pcm-Bal-23KDHlyA chimeric polypeptide:

It is of some interest to us whether foreign polypeptides when fused to the HlyA secretion domain can behave independently and take up their proper folded tertiary structure once secreted. Some preliminary attempts to assess this were therefore carried out. One way to obtain an insight into correct folding is to carry out the published chymosin milk clotting assay (Emtage *et al.*, 1983), assuming that Pcm deleted for the final 3 amino acids (as is present in the chimeric peptide Pcm-Bal-23KDHlyA) retains activity. However, normal activation of Pcm unfortunately requires a complex pH and concentration dependent sequential autocatalytic cleavage of the pro sequence under acid conditions to produce active chymosin (McCaman *et al.*, 1985). Achieving the correct conditions for complete acid cleavage of the entire pro- sequence was beyond the scope of this work so another indicator was sought.

The prochymosin molecule carries 6 cysteine residues which form 3 disulphide bridges required for protein integrity and activity (Foltmann *et al.*, 1979), while the C-terminus of haemolysin carries none. These bonds are formed under oxidising conditions and might be expected to form once the polypeptide is secreted from *E. coli*, that is, if the molecule can fold independently of the attached C-terminal 23KD HlyA extension and does not require accessory factors to aids its proper folding (such as BIP-like proteins implicated in the correct assembly of some proteins; see Chapter 1: Section 1.4.5). One way to obtain an impression whether these

disulphide bridges are formed is to monitor the mobility of the chimeric polypeptide on SDS-polyacrylamide gels in the presence or absence of a reducing agent, such as 2-mercaptoethanol.

Thus, a portion of the 60KD, Pcm-Bal-HlyA23KD polypeptide secreted in a previous experiment was resuspended in either reducing or non-reducing sample buffer and separated by SDS-PAGE. The resulting gel is shown in Figure 8. It is evident that there is a difference in mobility of the hybrid secreted protein, due solely to the presence or absence of the reducing agent. In the presence of reducing compound the secreted band (as detected by Coomassie staining) migrates slower than its counterpart in the absence of the reducing agent (Figure 8; Lanes 3,4 versus 1,2). This was interpreted to indicate that the addition of the reducing agent is breaking intramolecular disulphide bridges resulting in the molecule losing some of its compact structure, therefore running more slowly in SDS-polyacrylamide gels than its tightly folded oxidised counterpart. This result suggests the formation of disulphide bridges in the secreted form of the hybrid but it does not address the question of the authentic linking of the cysteine residues. It should be reemphasised that the 23KD portion of HlyA does not possess any Cys residues and thus could not participate in disulphide bonding.

3.5 Secretion of beta-galactosidase via the Hly export pathway:

3.5.1 Background:

Beta-galactosidase is a large (c.117KD) cytoplasmic protein which many groups have attempted to translocate across the *E. coli* membrane by attaching classical N-terminal export signal sequences with the general conclusion, until recently (see below), that this molecule is not exportable. Indeed, the inability of such hybrid polypeptides to be translocated, leading to the jamming of the export channels, was widely utilised to identify the boundaries and nature of the N-terminal signal sequence by selecting and analysing mutants that were no longer recognised by the translocation apparatus (see Chapter 1; Section 1.2.2). Lee *et al.*, (1989) investigated the nature of the blocking effect of the LacZ molecule and found that multiple regions of the LacZ molecule hindered export, concluding that the premature folding of LacZ or of domains of it was probably responsible for the block in export. Thus, I decided to test whether the attempted export of the LacZ moiety, when directed to the haemolysin export pathway would also result in the specific blockage of

this secretion channel and allow the selection of targeting mutants which are no longer targeted to the export pathway.

3.5.2 Generation of a Cro-LacZ-HlyA, 200KD hybrid protein:

The details of the construction of this inframe *cro-lacZ-hlyA* fusion is depicted in Figure 9. The vector employed, pEX (Stanley and Luzio, 1984), was designed primarily to generate fusions to the C-terminus of the 117KD Cro-LacZ polypeptide for hybrid overexpression and subsequent utilisation for the production of antibodies against both domains. Thus, a series of vectors were available to generate inframe fusions to the 3' end, via a multiple replacing the final 53 bases of *lacZ*. Insertion into this cloning site unfortunately abolishes LacZ activity and confers insolubility on the hybrid under conditions of high level expression (Stanley and Luzio, 1984). The *cro-lacZ* gene is under the control of the Lambda P_R promoter but as the plasmid does not encode the cI^{857} repressor protein to regulate expression, it has to be provided separately.

Due to the apparent inability in previous studies to translocate the LacZ molecule, presumably due to rapid folding and tetramerisation, I decided to generate a fusion to the largest *hlyA* fragment possible to try to "cushion" the signal domain from the large Cro-LacZ 117KD domain. Thus, I ligated the 2.4Kb ScaI/SalI *hlyA* fragment (encoding the final c.80KD of HlyA) from pACYC-CA,Cla (see Chapter 7; Section 7.5.1) into the SmaI/SalI sites of pEX-2 to generate pLG811 encoding an in phase c.200KD Cro-LacZ-HlyA hybrid protein, as described in Figure 9.

3.5.3 Expression of the Cro-LacZ-HlyA hybrid protein in the absence of export protein:

pLG811 was transformed into the *E. coli* strain ΔH1, which carries the cI^{857} gene on the chromosome. This controls expression from the lambda P_R promoter and thus expression of the Cro-LacZ-HlyA hybrid. A culture was initiated at an A^{450} of 0.1, and growth maintained at 30°C until an A^{450} of 0.7 was reached. At this stage a cellular sample was removed and the culture heat induced (as before and see Materials and Methods). This inactivated the CI^{857} repressor, leading to high level expression from the lambda P_R promoter. After a further 1.5 hours of growth a cellular sample was taken for both an inclusion body preparation (see Materials and Methods) and for analysis of total cell protein by SDS-PAGE. The results of the gel analysis are shown in Figure 10a together with the results of Western blot analysis, probing with anti-23KDHlyA antibodies

(Figure 10b). The Coomassie stained gel shows the intracellular accumulation of a protein of approximately 200KD (Lane 1,2) while the inclusion preparation shows that most of the hybrid polypeptide is indeed present in an insoluble form (Lane 3) as reported by Stanley and Luzio (1984). Western blot analysis confirmed the identity of this hybrid polypeptide and also demonstrated the presence of some hybrid protein even before induction of expression (Figure 10b; Lane 1), representing basal level expression.

3.5.4 Secretion of the Cro-LacZ-HlyA hybrid protein in an HlyB,D dependent manner:

The *E. coli* strain chosen to test for secretion of the 200KD Cro-LacZ-HlyA hybrid was the *E. coli* strain, SE5000, used routinely in these secretion experiments as there is little evidence of lysis during normal growth. As this strain does not encode the lambda temperature sensitive repressor protein (CI^{857}) required to control expression of the hybrid protein, a low copy number, tetracycline resistant, plasmid pRK248 (Bernard and Helinski, 1979) carrying the cI^{857} gene was transformed into the cells. This strain was then transformed with pLG811 (*cro-lacZ-hlyA*) and pLG575 (*hlyB,D*) and maintained at 30°C. Overnight cultures of SE5000 pRK248/pLG811 +/- pLG575 were used to inoculate fresh Luria broth to an A^{450} of ~0.1. At an A^{450} of ~0.7, cell and supernatant samples were taken for later analysis by SDS-PAGE and the culture heat induced for hybrid expression as usual (see Materials and Methods). More cell and supernatant samples were taken 1 and 3 hours post-induction, and analysed on a 9% SDS-acrylamide gel as before (see Figure 11a).

Lanes 1-3 and 4-6 show the cellular accumulation of a protein band of approximately 200KD, after induction of high level expression, with more apparently accumulating in the cells (Lane 2-3) minus pLG575 (*hlyB,D*) than plus (Lane 5-6) pLG575 (*hlyB,D*). Lanes 7-9 demonstrate that this hybrid is not released to the supernatant in the absence of HlyB,D, whereas Lane 10 reveals the secretion of the 200KD polypeptide, even before induction of high level expression, in an HlyB,D dependent manner. The identity of this secreted polypeptide was confirmed by Western blot analysis with antibodies directed against the 23KD C-terminal peptide of HlyA (data not shown). Lane 10 also shows the presence of a low background of cellular bands, presumably due to some cell lysis (or whole cell contamination). However, comparison of the samples before induction, Lanes 4 (Cells) versus 10 (Supernatant), clearly show the enrichment of the level of the released 200KD protein indicating

the specific secretion of the hybrid protein. Unfortunately, heat induction of high level expression of the hybrid protein inhibited growth (Figure 11b), and led to more extensive cellular lysis as evident in Lanes 11 and 12 of Figure 11a, obscuring the possible secretion of the hybrid protein.

The growth curves from the above experiment (Figure 11b) also show, as is often the case, that the presence of the Hly export proteins seems to inhibit the growth of cultures to some degree (unpublished observations; this laboratory). However, upon induction of high level expression of the Cro-LacZ-HlyA hybrid the cells expressing HlyB,D quite dramatically ceased to grow and lysed. In contrast, over-expression of the hybrid in the absence of export proteins had a much reduced inhibitory effect on growth. This suggests that targeting of the large hybrid protein to the export channel under conditions of high level expression leads to abortive translocation complexes which interfere with cellular growth, and lead to eventual lysis.

3.5.5 Secretion of the Cro-LacZ-HlyA hybrid protein in an HlyB,D dependent manner, under conditions of basal level expression:

An identical experiment was carried out to that described above except that the cultures were kept at 30°C, permitting only basal level expression of the hybrid from the λ P_R promoter, and using nutrient broth instead of Luria broth. Cellular and supernatant samples were taken 1, 2.5, 4 and 4.5 hours after initiation of growth at an A^{450} of 0.1 with an inclusion body preparation carried out on the final sample (see Materials and Methods). The proteins from the samples were separated on a 9% SDS-acrylamide gel with only the supernatant profiles shown in Figure 12.

The cellular protein profiles did not show, as expected, visible levels of the 200KD Cro-LacZ-HlyA hybrid as confirmed by the inability to isolate insoluble inclusion bodies (data not shown). Figure 12, however, demonstrated the apparently constant low level but specific secretion of the 200KD Cro-LacZ-HlyA hybrid in an HlyB,D dependent manner (Lanes 1-4 versus 5-8). The background smear in the supernatant lanes represents nutrient broth peptides which are precipitated by TCA even in the absence of cells (see Chapter 4: Figure 4). The amount of these peptides decrease as less supernatant sample is loaded, to compensate for the increasing culture density, in order to give equivalent A^{450} loading. The growth curves for these cultures are also displayed in Figure

11b. Again, as stated above, the presence of the export proteins reduces the growth rate of that strain. However, no major toxic effects are observed this time due to the combined presence of the Cro-LacZ-HlyA hybrid and the Hly export proteins. This is presumably due to the ability of the export machinery to successfully secrete the low basal levels of the hybrid without jamming the export channel.

3.6 Discussion:

The results described in this chapter demonstrate the first known example of the specific secretion, from the *E. coli* cytoplasm to the external medium, of the majority of both a mammalian polypeptide (Prochymosin) and the *E. coli* cytoplasmic protein (Beta-galactosidase), achieved by fusing them to the C-terminal HlyA signal domain, and harnessing the haemolysin export polypeptides/channel. The ability to secrete a fragment of Pcm when fused only to the C-terminal 23KD HlyA domain implies that not only does this domain encode all the information required to secrete itself (as shown by Nicaud *et al.*, 1986), but more importantly does not require any HlyA sequences further upstream to allow the secretion of N-terminally attached domains, whether HlyA or heterologous.

An important observation from this work was the ability of the 23KD portion of HlyA to secrete not only the N-terminal (42%) portion of Pcm but also 72 and 99% of the molecule. This contrasts, surprisingly, with the unsuccessful attempts of Little *et al.*, (1989) to export Pcm to the periplasm by fusing it to the N-terminal signal sequence of alkaline phosphatase (an *E. coli* periplasmic protein). These authors identified a translocation "blocking sequence" within the first 29 amino acids of prochymosin, inhibiting both the processing of the signal sequence and translocation of the molecule. However, processing and presumed translocation of at least part of the molecule was permitted when these residues were deleted. This inhibitory region contains a high percentage of charged hydrophilic residues (8 out of the first 20 amino acids of the molecule are glutamic acid, lysine or arginine - with an overall positive charge). The authors speculated that this may disrupt the normal association of the amino terminus of the polypeptide with elements of the secretory mechanism. This proposition is supported by the observations that increasing the positive charge at the N-terminus of alkaline phosphatase (PhoA) reduces its ability to localise in the *E. coli* periplasm (Li *et al.*, 1988). Moreover, the insertion of a positively charged 51 residue

segment of beta-galactosidase inhibited the translocation of alkaline phosphatase when inserted between the signal peptide and mature portions of the export protein (Lee *et al.*, 1989). Indeed, the inhibitory effect of this latter fragment was lessened by reducing its overall positive nature by site directed mutagenesis (Lee *et al.*, 1989).

Further attempts by Little *et al.*, (1989) to export the remaining portion of the Pcm molecule (deleted for this "blocking" sequence) fused only to the alkaline phosphatase signal sequence, failed to show processing and no translocation was detected. The authors suggest that this might be due to the ability of the overexpressed Pcm molecules to form insoluble aggregates, blocking their entry into the secretory pathway. Alternatively, it could be due to the presence of other translocation "blocking" sequences.

Thus, the ability of the haemolysin export system to secrete not only the N-terminus of Pcm but essentially the whole molecule (361 out of 364 residues) when fused to the C-terminus of HlyA must be highlighting the existence of some intrinsic differences in the translocation mechanisms/ channels of the *hly* and *secA/Y* pathways. This difference could be interpreted to indicate the presence in the *hly* post-translational translocation system, of an unfoldase activity (proposed to exist for the post-translational import of proteins targeted to the mitochondria: Rothmann and Kornberg, 1986; Ostermann *et al.*, 1989). This could for example, be inherent in HlyB (an ATP binding protein; Higgins *et al.*, 1986), to allow this system to unfold strong structures which otherwise might block translocation via the *secA* pathway. Alternatively, it might also reflect a difference in the size or specificity of the translocation pore/channel, affecting the ability of the two systems to accommodate different size/ conformations of polypeptides.

The question of the efficiency of secretion of these chimeric polypeptides will be addressed in more detail in Chapter 5. It is appropriate here, however, to point out that the efficiency of secretion is extremely difficult to access in this system, for example by comparing the level of intracellular to extracellular protein. This is due to the fact that high level expression of Pcm leads to the formation of insoluble inclusion bodies. This has been suggested to reflect the improper folding during synthesis of the Pcm molecules in the bacterial cytoplasm, triggering the formation of inclusion bodies (McCaman, 1985). Therefore, the levels of soluble Pcm-23KDHlyA hybrid molecule available to the secretion machinery is

probably limited. Indeed, the secreted protein could be looked upon as that rescued from the inclusion body pathway. We cannot guess what the kinetics might be between the competition for the nascent polypeptide for incorporation into the inclusion body or the haemolysin secretion pathways. However, it is not unreasonable to assume that the secretion process may in any case reach saturation level under the conditions used, the limiting step probably being the number of export "channels".

It has been suggested by McCaman and co-workers (1985) that the correct folding of Pcm, when overexpressed in *E. coli* is not possible due to the environment of the bacterial cytoplasm, leading to the formation of inclusion bodies. They suggest that correct folding of Pcm is normally dependent upon co-translational secretion into the extracellular space following its synthesis on the bovine endoplasmic reticulum. Therefore, we might envisage that the ability to secrete Pcm from *E. coli* might allow the Pcm molecule to fold correctly, unless other components are required to enable proper folding. Whether this secreted near full length Pcm-HlyA hybrid product can be activated, under the correct conditions of pH and concentration, when fused to 23KD of the HlyA polypeptide (but deleted for its final 3 residues) remains to be tested. As stated earlier (Section 3.4) the Pcm activity assay for this secreted hybrid was beyond the scope of this study. However, the presence/ absence of the reducing agent, 2-mercaptoethanol, was used to detect any differences in the mobility of this chimeric protein on SDS-polyacrylamide gels. Under non-reducing conditions a clearly faster moving form was detected consistent with some intramolecular cross linking, causing the molecule to adopt a more compact configuration. This is an indication that the Pcm molecule can fold to some degree, despite the C-terminal presence of the 23KD of HlyA.

The significance of the observed partial cleavage of all three secreted chimeric Pcm-23KDHlyA polypeptides (38, 50 and 60KD) to a second form apparently reduced in size by around 4 kilodaltons (evident using SDS-PAGE conditions that can resolve the small difference in sizes) under acidic conditions (TCA concentration of the supernatant samples), is unknown. Acidic conditions are a requirement for the normal autocatalytic cleavage of the pro sequence of prochymosin, but as the whole molecule is not present in all the chimeric molecules, we cannot consider this an indication of autocatalysis and ability to be activated. Perhaps, this result reflects a non-autocatalytic cleavage induced under acidic conditions which nevertheless occurs at or close to the normal position of cleavage within the Pcm molecule. Now, that it has been

shown that the Pcm molecule has the capacity to be secreted directly to the medium via the haemolysin export pathway, it is worthwhile to generate, in the future, a fusion to the entire Pcm molecule incorporating a unique protease site to allow the specific removal of the HlyA signal domain after secretion. The secreted Pcm molecule could then be easily isolated from the supernatant and tested for activity.

In relation to the use of the HlyA secretion signal domain to secrete heterologous proteins that might have wild-type activities, it is interesting to note an example by Steipe and Pluckthun (personal communication; Holland *et al.*, 1990c). These workers successfully secreted the V_L and V_H variable domains of the murine immunoglobulin A (which as a heterodimer normally binds phosphorylcholine) when fused to the 23KD HlyA domain. Moreover, the IgA molecules were then successfully released from the HlyA moiety by proteolytic cleavage at a specific site engineered at the junction of the two proteins. The two peptides (V_H and V_L) were then shown to be capable of forming specific heterodimers permitting the expected binding to phosphorylcholine. This provided additional evidence therefore that molecules secreted by the HlyA system can adopt at least some folding characteristics of the native molecules.

Finally the secretion of a virtually entire *E. coli* cytoplasmic protein, beta-galactosidase across both the inner and other membranes of *E. coli* was a surprising and unique achievement. Many previous attempts to export this protein only across the *E. coli* cytoplasmic membrane by fusing it to appropriate N-terminal signal sequences, with or without intervening portions of the mature export protein, have invariably failed. This repeated observation has led to the general conclusion that this protein encodes export incompatible features or folds too rapidly thereby preventing its transfer. Indeed, Lee *et al.*, (1989) carried out a study on the exportability of different portions of the LacZ molecule to determine whether there are specific inhibitory sequences in the molecule. Their conclusion was that some portions of the molecule did not inhibit translocation, but that the molecule as a whole consisted of a number of weakly inhibiting sequences which together strongly inhibited transfer. Other evidence by researchers such as Randall and Hardy (1986) and Eilers and Schatz (1986) suggest that it is the premature folding of a molecule into a stable conformation that can lead to its inability to be

translocated. Thus, Lee *et al.*, (1989) suggest that in the case of beta-galactosidase either the whole or particular sub-domains of the molecule have evolved which allow it to rapidly fold and preclude its translocation.

On the other hand, Freudl *et al.*, (1988) obtained data suggesting that the entire beta-galactosidase (deleted for only 6 N-terminal residues) molecule can be exported across the cytoplasmic membrane when fused to the N-terminal 219 residues (including its signal sequence) of the outer membrane protein, OmpA. Export was obtained when the hybrid was expressed at low levels but the hybrid was toxic and the cells lysed when over-expressed. About 20% of the hybrid protein was found associated with the outer membrane as deduced from the properties of the OmpA fragment, although all the hybrid protein was apparently digested by trypsin when the outer membrane was permeabilised, indicating a non cytoplasmic location. These experiments can, however, be critised on a number of grounds, including whether the apparent export was still *secA,Y* dependent. In addition the authors did not show in the trypsin accessibility experiment that cytoplasmic proteins were not also accessible to digestion. Moreover, this paper does not conclusively prove that complete translocation of the entire beta-galactosidase molecule across the inner membrane had occurred.

In contrast to their other findings Freudl *et al.*, (1988) also report that LacZ hybrids fused to the N-terminal 85 or 44 residues of OmpA were not secreted which strongly suggests that the export channel does not in fact have the capacity to translocate the LacZ moiety. The authors suggested that the folding of the exported mature 219 amino acid portion of OmpA (see above) "contributes to the vectorial forces which drive translocation" across the cytoplasmic membrane. Thus, this may be regarded as a special case and it remains a possibility that the large hybrid was only partially translocated with the mature exported portion of OmpA processed and inserted in the outer membrane, leaving the LacZ moiety spanning the inner membrane.

Recently, Philips and Silhavy (1990) have published data claiming that the LacZ moiety of a hybrid protein can be targeted across the *E. coli* inner membrane in the presence of over-expressed levels of the *E. coli* heat shock proteins, DnaK and GroEL. These proteins are members of the heat shock family of molecular chaperones and are thought to participate in bacterial export by maintaining polypeptides in a translocation competent conformation (Bockkarera *et al.*, 1988; Lecker *et al.*, 1989 and

Kusukara *et al.*, 1989). In the presence of high levels of either of these heat shock proteins the processing of the signal sequence of a LamB-LacZ hybrid was observed to increase while the jamming of the export pore to normal export protein decreased. However, the location of the processed hybrid was not determined although it was membrane associated. Whilst these results are consistent with successful avoidance of folding and therefore transport of the LacZ moiety they do not rule out the possibility that high levels of DnaK and GroEL might function by pulling a jammed LamB-LacZ hybrid, spanning the inner membrane with the signal sequence having already been processed, back into the cytoplasm. Such reversible translocation of certain mutant proteins has been observed previously (T. Silhavy, personal communication; unpublished data this laboratory).

In contrast to the experiments in bacteria, Keng *et al.*, (1986) demonstrated that beta-galactosidase can be translocated across both the inner and other membranes of the yeast mitochondria, targeted by as few as 9 N-terminal residues of an imported mitochondrial matrix protein. Thus, it would appear that translocation mechanisms involved in the transfer of proteins across double membranes (such as that of mitochondrial import and haemolysin secretion) are intrinsically more capable of translocating a wider range of proteins including cytoplasmic proteins compared with the SecA,Y translocation pathway. Mitochondrial import and haemolysin secretion are truely post-translational (see Hartl and Neupert, 1990 and Chapter 1; Section 1.4.2 and below). In contrast, the SecA,Y pathway although able to support post-translational export of some proteins (that is, those that do not rapidly adopt highly folded conformations), are in fact mostly co-translational *in vivo* (see Chapter 1; Section 1.3.1), presumably minimising the opportunity for the export protein to fold. This may indicate that although the haemolysin and mitochondrial translocators recognise quite different targeting signals, they both have the capacity to contend with or unfold stable structures, which must inevitably form before the polypeptides post-translationally engage the export pathway. However, in the case of mitochondria it is thought that unfolding is a prerequisite for import, catalysed probably by heat-shock like proteins (see Chapter 1: Section 1.4.5) but there is also a possible role for other cytoplasmic proteins, such as PBF (presequencing binding factor: Murakami and Mori, 1990) in binding precursor proteins and maintaining them in a import-competent conformation.

Such accessory factors appeared to be ruled out in the case of the Hly export machinery where the targeting signal is at the extreme C-terminus and no additional secretion signals are present in the remainder of the molecule. Thus, the ability of the *hly* system to translocate *E. coli* cytoplasmic and foreign proteins (greater than 90% of the Pcm, LacZ and chloramphenicol transacetylase molecules: Kenny *et al.*, to be submitted; Holland *et al.*, 1990c), supports the idea that the majority of such molecules must have already folded even before the HlyA signal domain is synthesised. Thus it appears reasonable to propose that the Hly translocation machinery at least differs from that of the SecA,Y pathway, in translocating highly "folded" molecules, either in this form or concomitant with active unfolding. Interestingly, there is some evidence that the import of mitochondrial proteins requires an unfoldase activity (see Chapter 1; Section 1.4.5). HlyB has a putative ATP binding site (Chapter 1; Section 1.6.3) which might be involved in the provision of energy through ATP hydrolysis to unfold structures found incompatible for transport via the SecA,Y export pathway. However, energy generation by HlyB might be required for alternative steps in the transport process.

Concerning the ability of the HlyA signal to direct the secretion of heterologous proteins it is important to note that in the great majority of constructs so far produced the passenger protein lacks the full sequence. Thus, the Cro-LacZ-HlyA hybrid lacks the C-terminal 17 residues of LacZ resulting in the loss of LacZ activity and possibly its capacity to tetramerise. This might be the case especially in the presence of the 80KD HlyA domain. This could reduce the capacity of the LacZ domain to adopt a stable highly folded quaternary structure rendering it more compatible for secretion by the Hly export machinery. However, this does not diminish the fact that the Hly export machinery is apparently able to translocate molecules/structures which inhibit the translocation of mammalian Pcm (this Chapter), and *E. coli* beta-galactosidase (see above) across the *E. coli* cytoplasmic membrane when fused to N-terminal signal sequences (Little *et al.*, 1989; Lee *et al*, 1989). It should be pointed out that Gentz *et al.*, (1988) only attempted to and failed to translocate the entire CAT molecule via the SecA,Y pathway. However, the results indicated that the CAT protein was rapidly degraded by the mere interaction of the signal sequence and the export pathway, suggesting that transport would be prevented irrespective of the size of the CAT fragment. In studies in this laboratory (Kenny *et al.*, to be submitted; R.

Haigh this laboratory) the Hly pathway secreted at least 91% of the CAT molecule to the medium but the secretion of an intact CAT protein was inhibited with the molecule accumulating within the cells. This failure to be exported in this case may be due to the rapid folding and in particular trimerisation of the intact CAT molecule. Thus, the Hly pathway although apparently able to translocate many structures inhibitory to the SecA,Y pathway, conceivably involving some unfoldase activity may not be able to cope with complex quaternary or multimeric structures.

Figure 1:
Prochymosin gene under the control of the trp promoter in the dual origin vector pMG168 (Wright et al., 1986).
(see text for details)

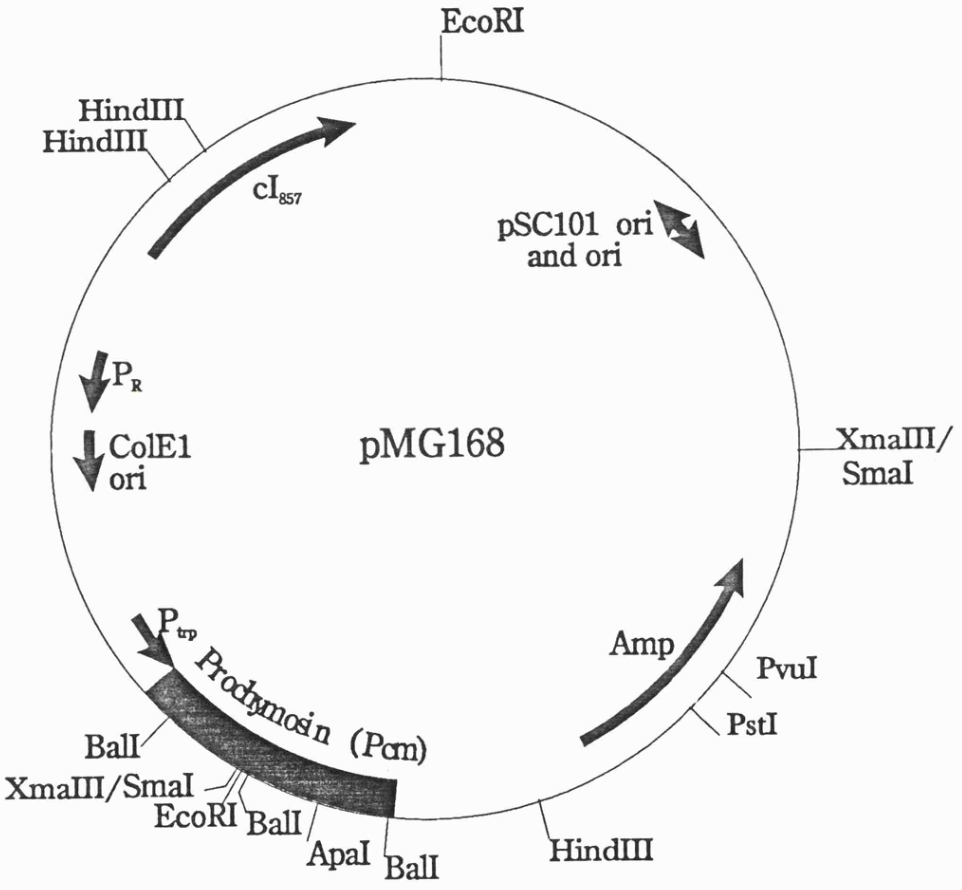
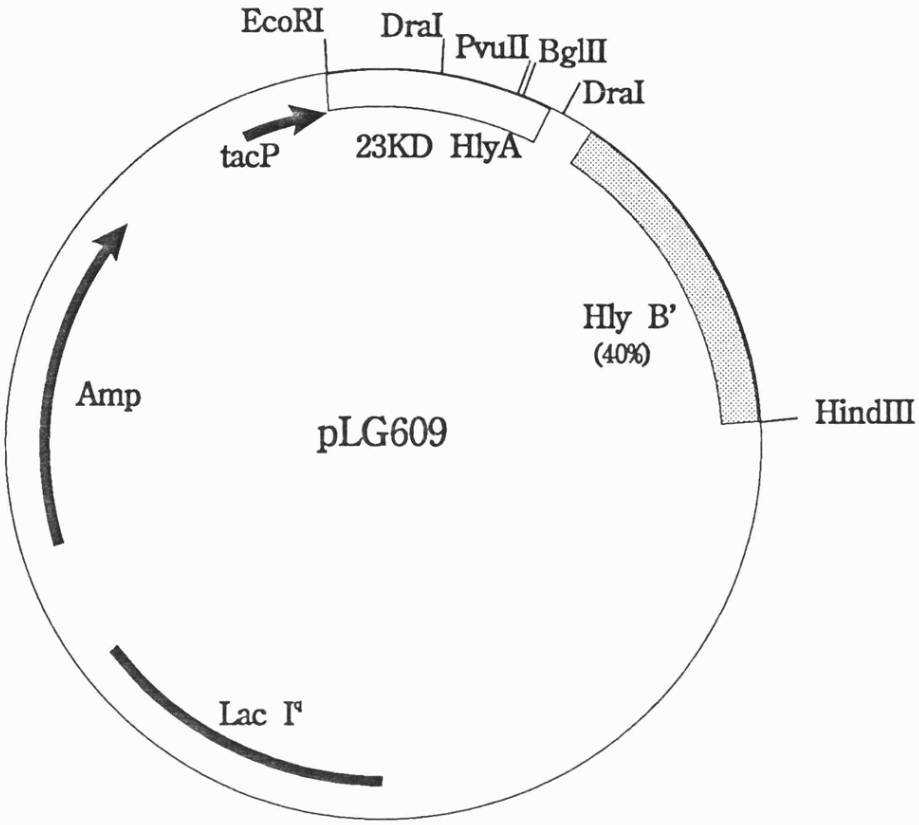

EcoRI
HindIII
HindIII
cI857
pSC101 ori
and ori
PR
ColE1
ori
pMG168
XmaIII/
SmaI
Ptrp
Prodymosin (Pcm)
Amp
PvuI
PstI
BalI
XmaIII/SmaI
EcoRI
BalI
ApaI
BalI
HindIII

Figure 2:

a) Map of pLG609 carrying the 3' end of *hlyA* (encoding the C-terminal 23KD) and the 5' end of *hlyB*, under the control of the inducible *tac* promoter (Nicaud et al., 1986)

b) Alternative versions of pLG609 generated by the insertion of oligonucleotide linkers into the EcoRI site of pLG609 to permit the isolation of the *hlyA* fragment in all three reading frames (pLG609-1/-2/-3), and using a BamHI restriction site (pLG609-4).

a)



b)

	EcoRI
pLG609	G| AAT T CT CTT
	C TTA A |GA GAA
	SmaI(XmaI)
pLG609-1	G AAT TTT CCC C |GGG GAA AAT TCT CTT
	C TTA AAA GGG G |CCC CTT TTA AGA GAA
	SmaI(XmaI)
pLG609-2	G AAT TTT CCC | GG GAA AAT TCT CTT
	C TTA AAA GGG | CC CTT TTA AGA GAA
	SmaI(XmaI)
pLG609-3	G AAT TTC CC | G GGA AAT TCT CTT
	C TTA AAG GG | C CCT TTA AGA GAA
	BamHI
pLG609-4	G ATT TTG G|GA TC C CCA AAT TCT CTT
	C TAA AAC C CT AG| G GGT TTA AGA GAA

Figure 3:

Diagrammatical representation of the *pcm*-23KD*hlyA* gene fusions generated in this chapter.

Filled in boxes represent regions of the *pcm* gene.

Dotted lines represent *pcm* region deleted in the fusions.

Open boxes represent the *hlyA* region encoding the C-terminal 23KD segment.

Pcm-23KDHlyA hybrid fusion-site junctions:

	Pcm		C-terminal HlyA
pLG800-1	ACC CAG GAG CCC		*GGG GAA* AAT TCT
	Thr Gln Glu Pro		*Gly Glu* Asn Ser
pLG801	AAG CTG GTC G		*GG GAA* AAT TCT
	Lys Leu Val(Gly)	Gly	(G*ly*)*Glu* Asn Ser
pLG802	GTG GGG CTG G		*GG GAA* AAT TCT
	Thr Gln Glu(Ala)	Gly	*(Gly)Glu* Asn Ser

Underlining = *hlyA* derived

Italics = linker derived (see Figure 2b)

Brackets = natural residue at this position prior to fusion

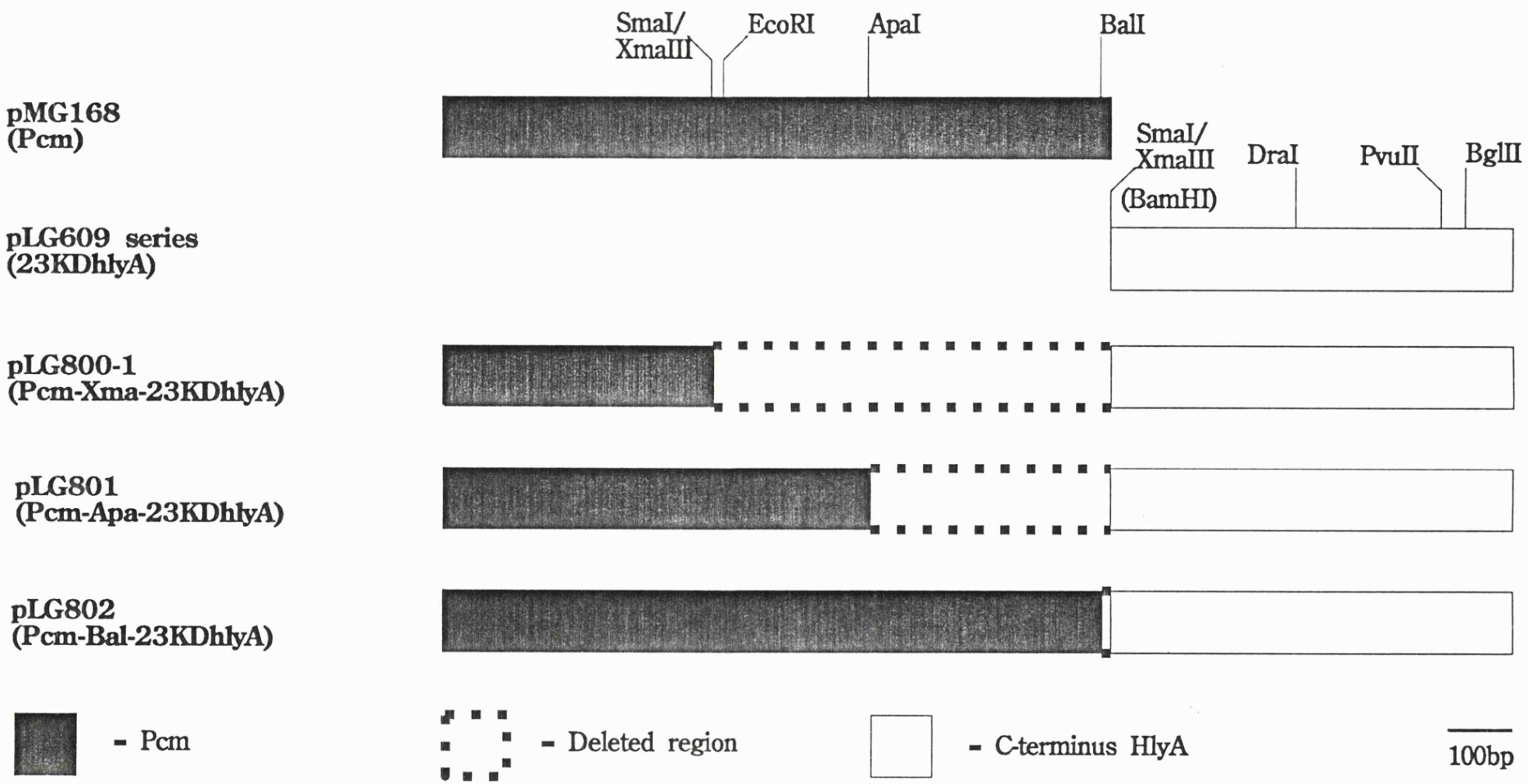
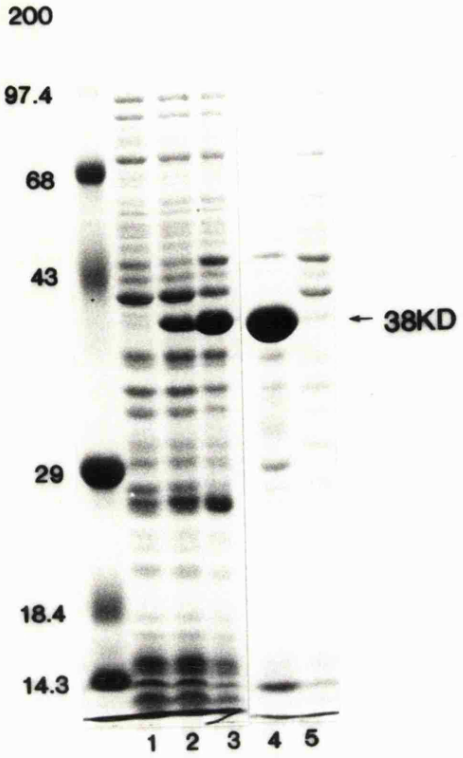

SmaI/
XmaIII
EcoRI
ApaI
BalI
pMG168
(Pcm)
SmaI/
XmaIII
DraI
PvuII
BglII
(BamHI)
pLG609 series
(23KDhlyA)
pLG800-1
(Pcm-Xma-23KDhlyA)
pLG801
(Pcm-Apa-23KDhlyA)
pLG802
(Pcm-Bal-23KDhlyA)
- Pcm
- Deleted region
- C-terminus HlyA
100bp

Figure 4:
Expression and insolubility of the 38KD Pcm-Xma-23KDHlyA hybrid:
The *E. coli* DH1 strain containing pLG800-1 (Pcm-Xma-23KDHlyA) was grown in nutrient broth (with appropriate antibiotics) to an A^{600} of 0.46 before heat inducing high level expression. Samples were taken before, 1 and 3.5 hours post-induction. At t = 3.5 hours a sample was also taken for separation into soluble and insoluble fractions using an inclusion body preparation (see Materials and Methods). Proteins were analysed by SDS-PAGE (15%) and visualised by Coomassie blue. 0.5 units of cells (where 1 unit is defined as 1ml at A^{600} = 1) and 1 equivalent units of the insoluble/ soluble cellular fractions were loaded. Cellular samples: lane 1-3, time = 0, 1 and 3.5 hours, respectively. Lane 4 is the insoluble protein (inclusion body preparation) with lane 5 the corresponding soluble fraction.
Arrow marks the position of the 38KD hybrid protein with molecular weight markers also shown.

Figure 5:
Secretion of the 38KD Pcm-Xma-23KDHlyA hybrid protein from *E. coli* in an HlyB,D dependent manner.
E. coli SE5000 containing pLG800-1 (Pcm-Xma-23KDHlyA) +/- pLG575 (*hlyB,D*) was grown in nutrient broth (with appropriate antibiotics) to an A^{600} of 0.4 before heat inducing high level expression. Samples were taken before, 1.5 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{600} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were analysed by SDS-PAGE (11%) and visualised by Coomassie blue.
Cellular samples: lane 1-3, time = 0, 1.5 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) and lane 4-6, time = 0, 1.5 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).
Supernatant samples: lane 7-9, time = 0, 1.5 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 10-12, time = 0, 1.5 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).
Note: The highly expressed band in Lanes 4-6 corresponds to the CAT protein expressed from the plasmid pLG575 (*hlyB,D*). Also the non-specific background peptides visible in the concentrated supernatant profiles originate from the nutrient broth (see Chapter 4; Figure 4)
Arrows mark the position of the 38KD hybrid protein and a lower molecular weight form, with molecular weight markers also shown.


200
97.4
68
43
29
18.4
14.3
38KD
1 2 3 4 5

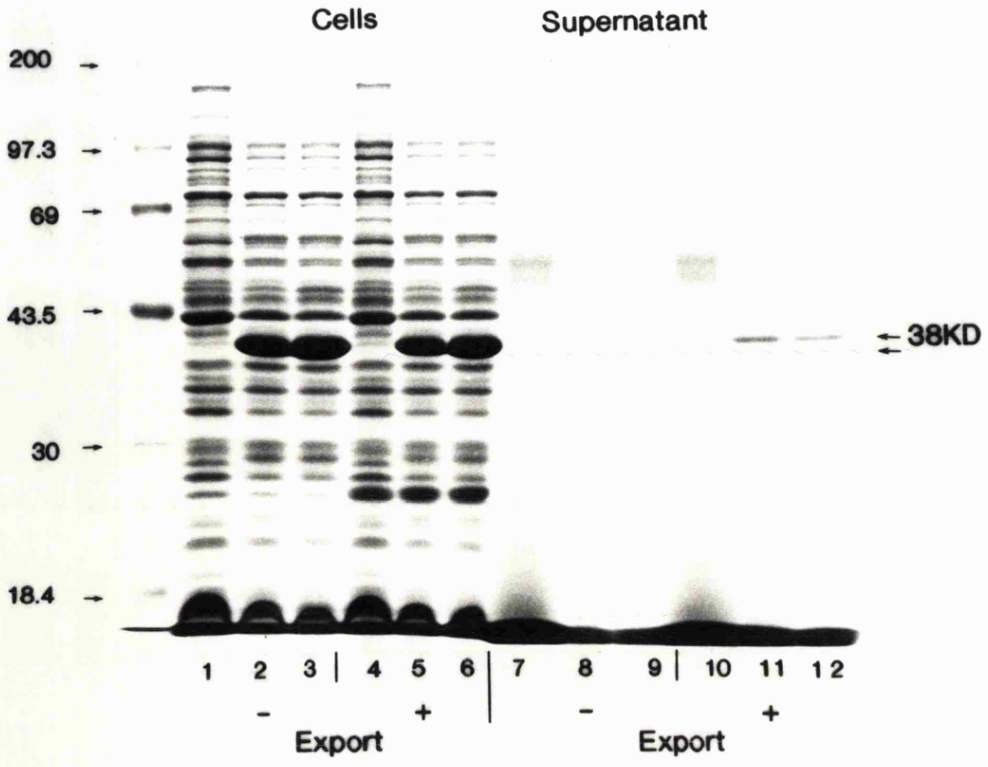

Cells
Supernatant
200
97.3
69
43.5
30
18.4
38KD
1 2 3 4 5 6 7 8 9 10 11 12
- + - +
Export
Export

Figure 6:

Secretion of the 50KD and 60KD Pcm-23KDHlyA hybrid proteins from *E. coli* in an HlyB,D dependent manner.

a) *E. coli* SE5000 containing pLG801 (Pcm-Apa-23KDHlyA) +/- pLG575 (*hlyB,D*) was grown in nutrient broth (with appropriate antibiotics) to an A^{600} of 0.4 before heat inducing high level expression. Samples were taken before, 1.5 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{600} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were analysed by SDS-PAGE (11%) and visualised by Coomassie blue.

Cellular samples: lane 1-3, time = 0, 1.5 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 4-6, time = 0, 1.5 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 7-9, time = 0, 1.5 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 10-12, time = 0, 1.5 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Note that the migration of the secreted 50KD peptides is slower in this gel, for unknown reasons, compared to cellular hybrid band. Not apparent in other gels (see Figure 7). Arrow marks the position of the 50KD hybrid protein with molecular weight markers also shown.

b) *E. coli* SE5000 containing pLG802 (Pcm-Bal-23KDHlyA) +/- pLG575 (*hlyB,D*) was grown in nutrient broth (with appropriate antibiotics) to an A^{600} of 0.4 before heat inducing high level expression. Samples were taken before, 1, 2 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells and 5 equivalent units of supernatant sample were loaded. Proteins were analysed by SDS-PAGE (11%) and visualised by Coomassie blue.

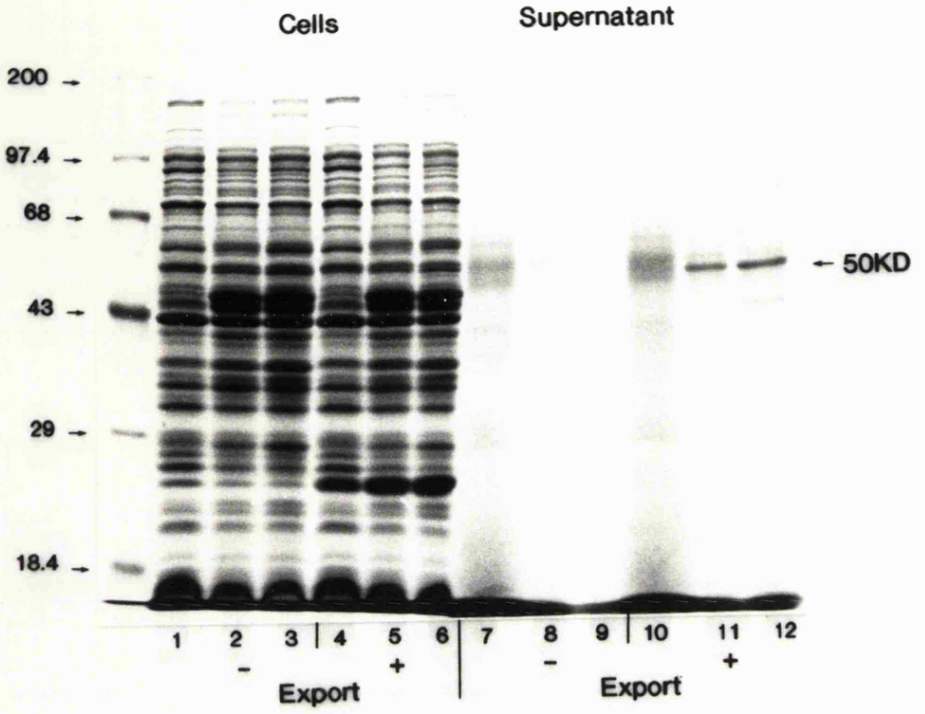
Cellular samples: lane 1-4, time = 0, 1, 2 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 5-8, time = 0, 1, 2 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 9-12, time = 0, 1, 2 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 13-16, time = 0, 1, 2 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Arrow marks the position of the 60KD hybrid protein with molecular weight markers also shown.

Again <u>note</u> that the non-specific background smear of peptides visible in the concentrated supernatant profiles originate from the nutrient broth (see Chapter 4; Figure 4).

a)


Cells
Supernatant
200
97.4
68
43
29
18.4
50KD
1 2 3 4 5 6 7 8 9 10 11 12
-
+
-
+
Export
Export

b)

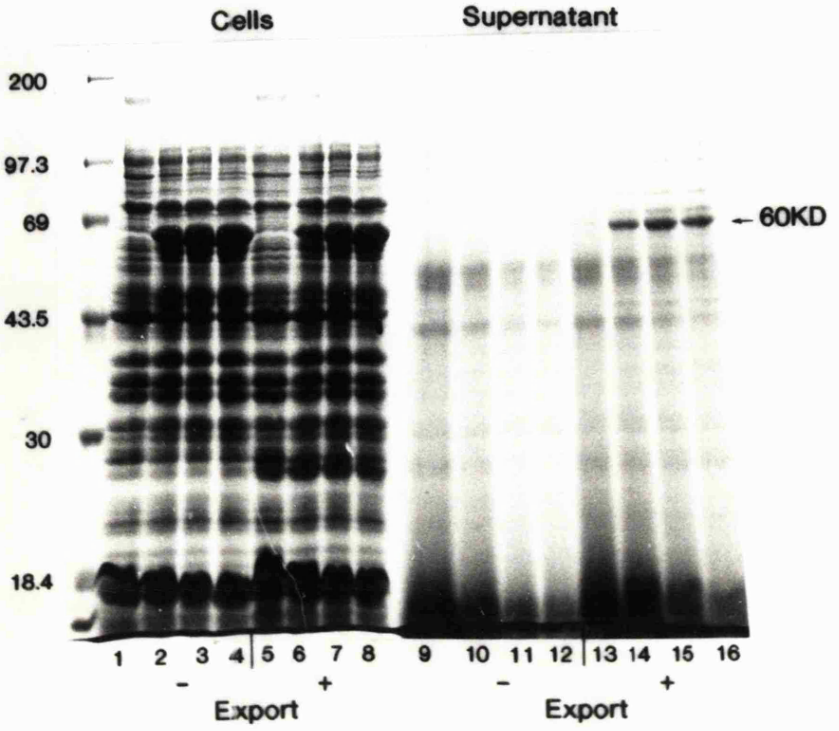

Cells
Supernatant
200
97.3
69
43.5
30
18.4
60KD
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
-
+
-
+
Export
Export

Figure 7:

Western blot analysis of Pcm-23KDHlyA Hybrids.

Additional aliquots of cell and supernatant samples previously analysed by SDS-PAGE in Figure 5 and 6a/b (Secretion of the 38, 50 and 60KD Pcm-23KDHlyA hybrids from *E. coli*) were run on an 11% SDS-acrylamide gel and probed by Western analysis using anti-23KDHlyA antibodies. 0.1 units of cells (where 1 unit is defined as 1ml at A^{600} = 1) and 0.5 equivalent units of supernatant sample were loaded (see footnote)

Lane 1: Cellular samples: pLG800-1 **minus** pLG575, t = 3 hours

Lane 2: Cellular samples: pLG800-1 **plus** pLG575, t = 3 hours

Lane 3: Supernatant samples: pLG800-1 **minus** pLG575, t = 3 hours

Lane 4: Supernatant samples: pLG800-1 **plus** pLG575, t = 3 hours

Lane 5: Cellular samples: pLG801 **minus** pLG575, t = 3 hours

Lane 6: Cellular samples: pLG801 **plus** pLG575, t = 3 hours

Lane 7: Supernatant samples: pLG801 **minus** pLG575, t = 3 hours

Lane 8: Supernatant samples: pLG801 **plus** pLG575, t = 3 hours

Lane 9: Supernatant samples: pLG802 **plus** pLG575, t = 3 hours

Lane 10: Positive control of secreted 23KDHlyA peptide

The Pcm-23KDHlyA secreted hybrids and their lower molecular weight forms (if resolved on this percentage gel) are indicated with arrows together with the positions of molecular weight markers.

pLG800-1 (38KD *pcm*-Xma-23KD*hlyA*)

pLG801 (50KD *pcm*-Apa-23KD*hlyA*)

pLG802 (60KD *pcm*-Bal-23KD*hlyA*)

pLG575 (*hlyB,D*)

Note: Unfortunately, the cellular samples for Lanes 5 and 6 were inadvertently diluted during the loading procedure, so volumes actually loaded are unknown.

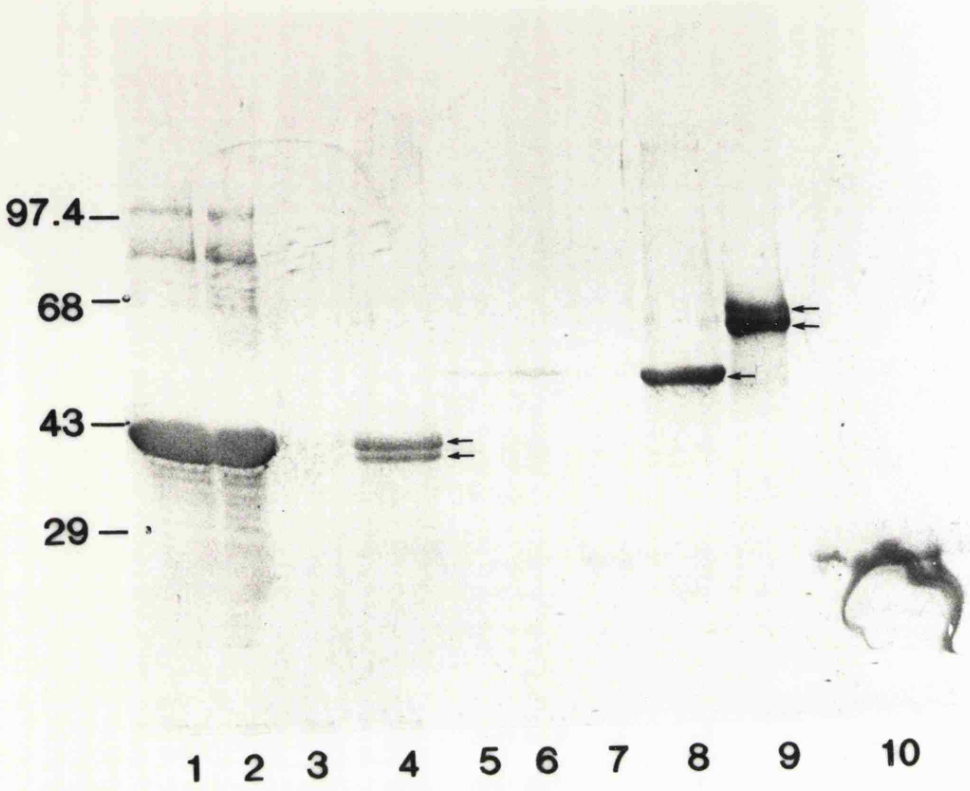

97.4
68
43
29
1 2 3 4 5 6 7 8 9 10

Figure 8:

Effect of the reducing agent, 2-mercaptoethanol, on the migration of the 60KD Pcm-Bal-23KDHlyA secreted protein.

SE5000 containing pLG802 (*pcm*-Bal-23KD*hlyA*) and pLG575 (*hlyB,D*) was grown to an A^{600} of 0.4 before heat inducing high level expression of the hybrid. 3 hours after induction a supernatant sample was taken and the proteins precipitated with TCA (10% w/v Final). Half the sample was resuspended in sample buffer either with or without 2-mercaptoethanol (0.716M final concentration). Samples were run slowly overnight (current = 10mA) on a 15% SDS-acrylamide gel and stained with Coomassie blue.

Lane 1: **minus** 2-mercaptoethanol 50ul loading

Lane 2: **minus** 2-mercaptoethanol 25ul loading

Lane 3: **plus** 2-mercaptoethanol 50ul loading

Lane 4: **plus** 2-mercaptoethanol 25ul loading

Arrow marks the position of the 60KD Pcm-23KDHlyA secreted hybrid with the molecular weight markers also shown.

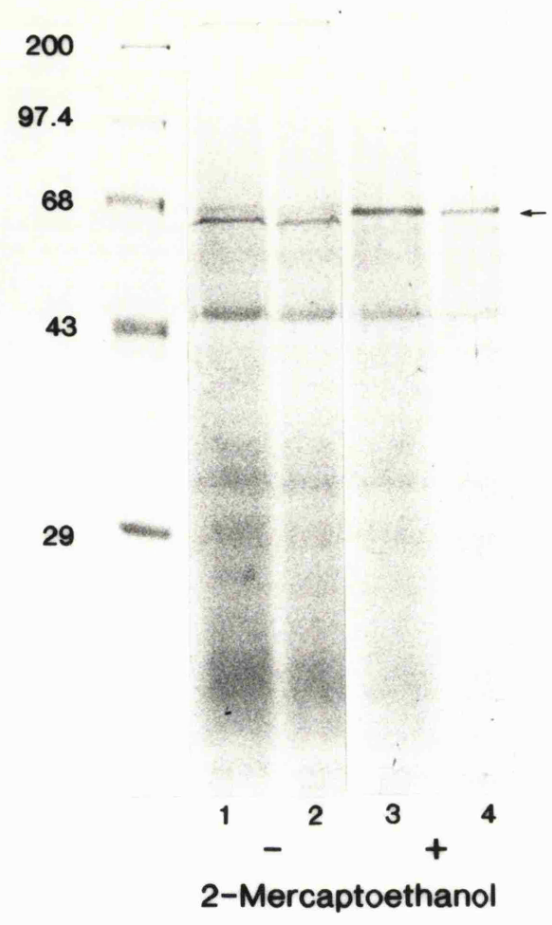

200
97.4
68
43
29
1
2
3
4
–
+
2-Mercaptoethanol

Figure 9:
Construction of pLG811 (a *cro-lacZ-hlyA* gene fusion) encoding a 200KD hybrid protein.
(see text for details)
pEX-2: see Stanley and Luzio, 1984
pACYC-CA,ClaI: see Chapter 7; Section 7.5.1

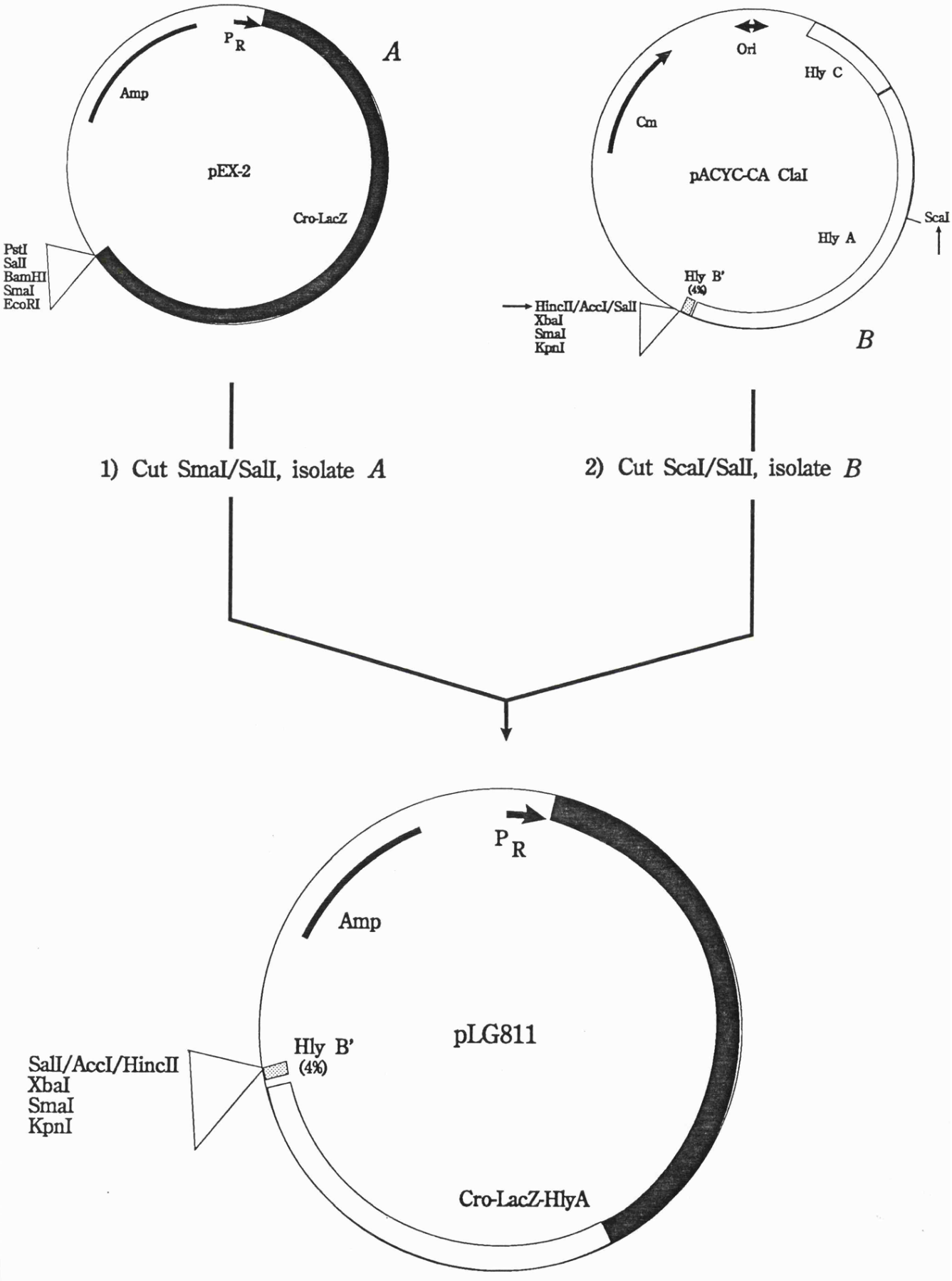

P R
A
Amp
pEX-2
Cro-LacZ
PstI
SalI
BamHI
SmaI
EcoRI
Ori
Hly C
Cm
pACYC-CA ClaI
ScaI
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
B
1) Cut SmaI/SalI, isolate A
2) Cut ScaI/SalI, isolate B
P R
Amp
pLG811
Hly B'
(4%)
SalI/AccI/HincII
XbaI
SmaI
KpnI
Cro-LacZ-HlyA

Figure 10:

Expression and insolubility of the 200KD Cro-LacZ-HlyA hybrid protein

The *E. coli* strain ΔH1 containing pLG811 (*cro-lacZ-hlyA*) was grown at 30°C in nutrient broth (with appropriate antibiotics) to an A^{450} of 0.7 before heat inducing high level expression. Samples were taken before and 1.5 hours post-induction. At t = 1.5 hours a sample was also taken and the insoluble protein fraction isolated using an inclusion body preparation (see Materials and Methods). Proteins were analysed on a mini- 15% SDS-acrylamide gel and visualised either by Coomassie blue staining (Figure 10a) or Western blot analysis (Figure 10b) using anti-23KDHlyA antibodies. As a positive control for the Western blot analysis a cellular sample from a previous experiment containing an induced 44KD hybrid protein composed of 99% of Pcm fused to the final 39 residues of HlyA was used (see Chapter 4: Section 4.4). 0.1 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 0.5 equivalent units of the insoluble cellular fraction were loaded.

a) and b) are Coomassie stained and Western blots respectively.

Lane 1: Cellular samples: pLG811 (*cro-lacZ-hlyA*): t = 0 hours

Lane 2: Cellular samples: pLG811 (*cro-lacZ-hlyA*): t = 1.5 hours

Lane 3: Insoluble fraction: pLG811 (*cro-lacZ-hlyA*): t = 1.5 hours

Lane 4: Cellular samples: pLG807 (*pcm*-Bal-39residue *hlyA*): t = 3 hours

Arrows mark the positions of the 200KD Cro-LacZ HlyA and 44KD Pcm-HlyA hybrid proteins with molecular weight markers also shown.

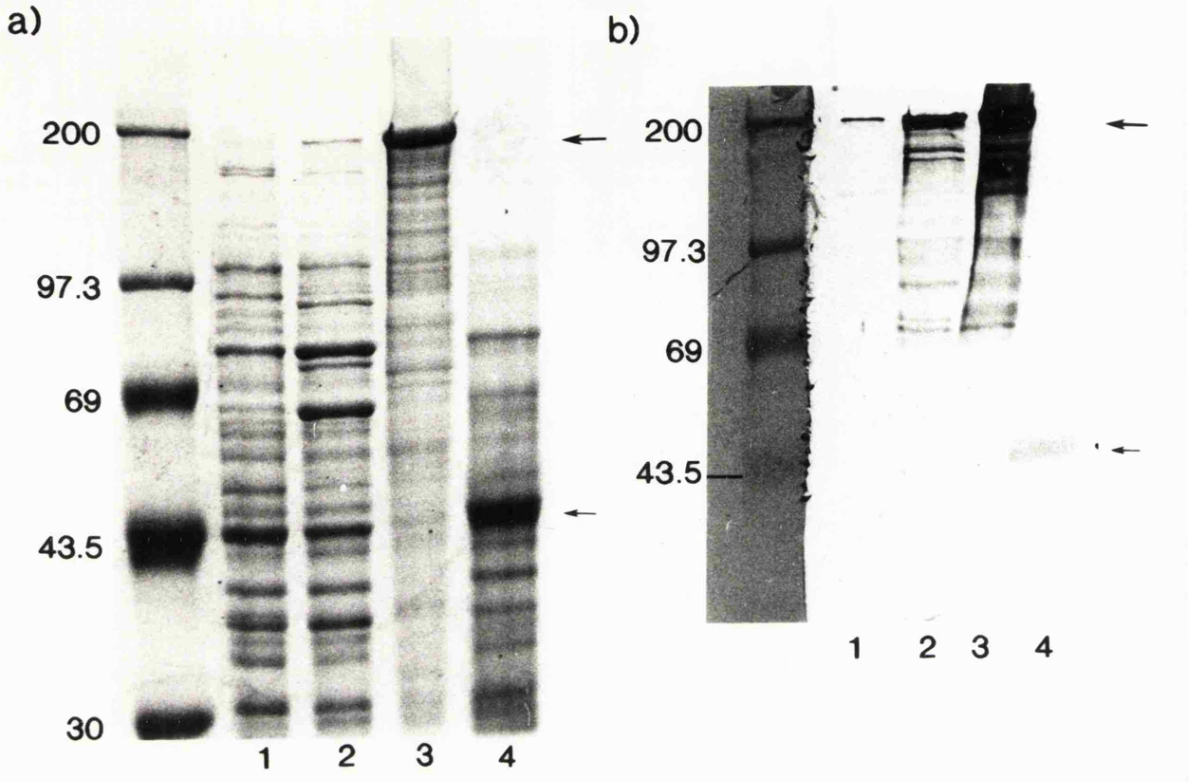

a)
200
97.3
69
43.5
30
1
2
3
4
b)
200
97.3
69
43.5
1
2
3
4

Figure 11:

Secretion of the 200KD Cro-LacZ-HlyA hybrid protein from *E. coli* in an HlyB,D dependent manner.

a) *E. coli* SE5000 containing pLG811 (*cro-lacZ-hlyA*), pRK248 (cI^{857}) +/- pLG575 (*hlyB,D*) was grown in Luria broth (with appropriate antibiotics) at 30°C to an A^{450} of ~0.7 before heat inducing high level hybrid expression. Samples were taken before, 1 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 10 equivalent units of supernatant sample were loaded. Proteins were analysed by SDS-PAGE (9%) and visualised by Coomassie blue staining.

Cellular samples: lane 1-3, time = 0, 1 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) and lane 4-6, time = 0, 1 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 7-9, time = 0, 1 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 10-12, time = 0, 1 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

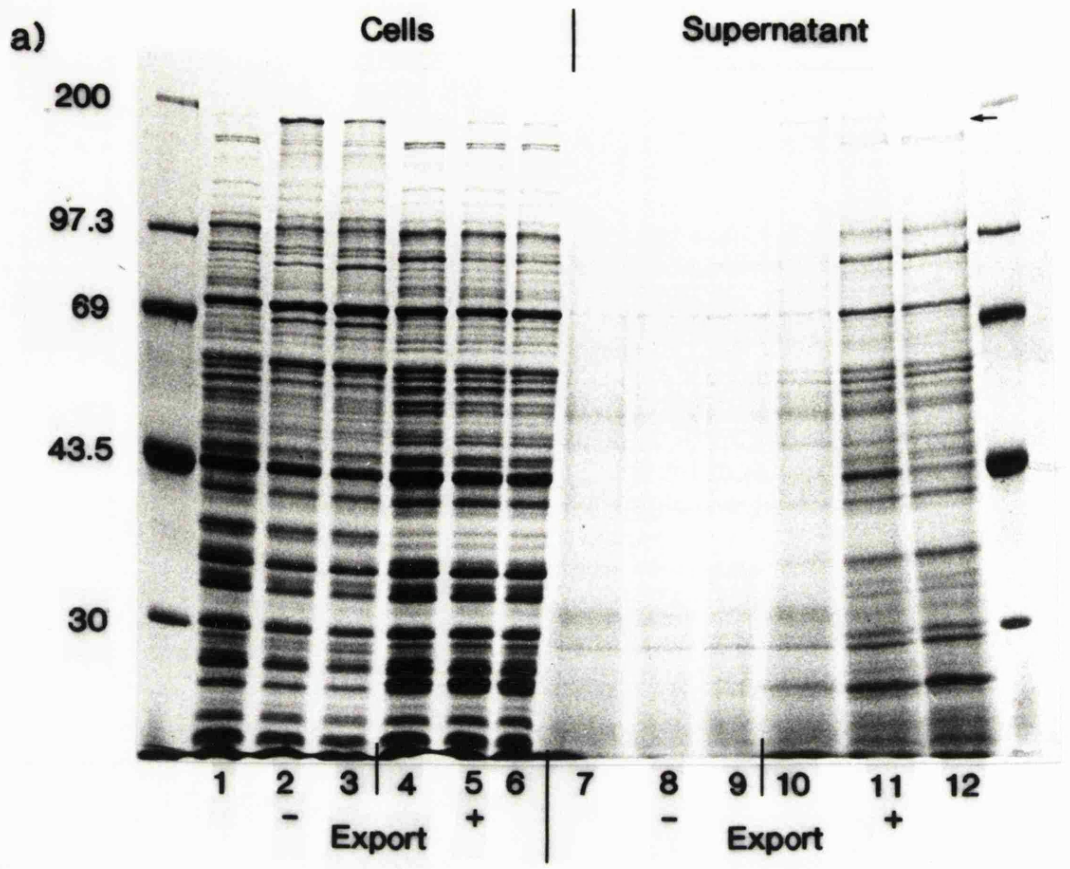
Arrow marks the position of the 200KD hybrid protein with molecular weight markers also shown.

b) Growth curves of cultures expressing high and low levels of the 200KD Cro-LacZ-HlyA hybrid in the presence/absence of HlyB,D:

At time intervals during the experiments outlined in 11a (above; induction of high level expression of the Cro-LacZ-HlyA hybrid) and Figure 12 (see below; basal level expression of the Cro-LacZ-HlyA hybrid) samples were taken and the optical density monitored as the experiment progressed.

Open squares represent cultures grown in the absence of HlyB,D (pLG575), with induced high level of expression of the hybrid Cro-LacZ-HlyA (solid lines) or basal level expression only (broken lines).

Stars represent cultures grown in the absence of HlyB,D (pLG575), with induced high level of expression of the hybrid Cro-LacZ-HlyA (solid lines) or basal level expression only (broken lines).



Growth curves of SE5000 pRK248/pLG811 +/- pLG575 under basal level or high level hybrid induction conditions

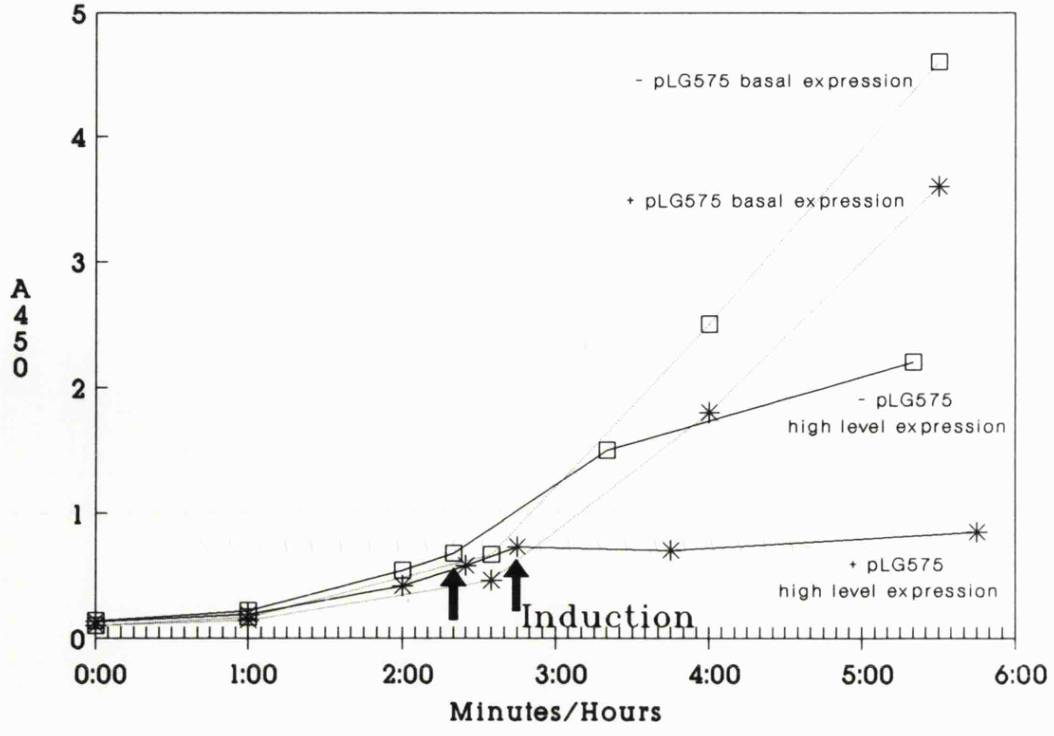


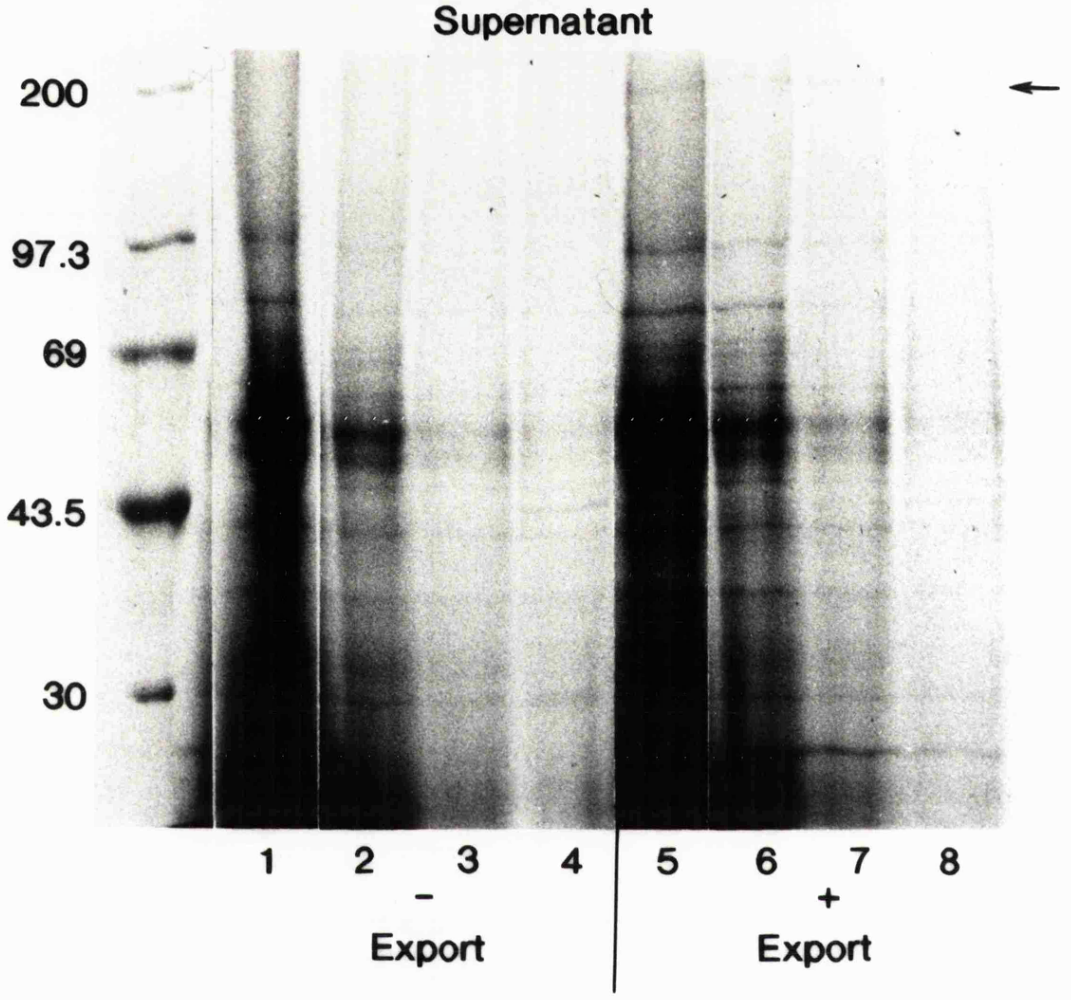
Figure 12:

Secretion of the 200KD Cro-LacZ-HlyA hybrid protein from *E. coli* in an HlyB,D dependent manner under conditions of basal level expression.

E. coli SE5000 cultures containing pLG811 (*cro-lacZ-hlyA*), pRK248 (cI^{857}) +/-pLG575 (*hlyB,D*) were initiated at an A^{450} of 0.1 and maintained at 30°C in nutrient broth (with appropriate antibiotics) throughout the course of the experiment. Samples were taken 1, 2.5, 4 and 4.5 hours after initiation of growth. Cell and culture supernatant samples were prepared as described in Materials and Methods. 1 equivalent unit (where 1 unit is defined as 1ml at A^{450} = 1) of supernatant sample was loaded in each case. Proteins were analysed by a mini- SDS-PAGE (9%) and visualised by Coomassie blue staining.

Supernatant samples: lane 1-4, time = 1, 2.5, 4 and 4.5 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) and lane 5-8, time = 1, 2.5, 4 and 4.5 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Arrow marks the position of the 200KD hybrid protein with molecular weight markers also shown. Background supernatant smear represent concentrated nutrient broth peptides, as witnessed by the decrease in intensity as less concentrated samples are loaded with time, in order to load equivalent (A^{450} cell equivalents) samples.


Supernatant
200
97.3
69
43.5
30
1
2
3
4
5
6
7
8
–
Export
+
Export

CHAPTER

4

Investigations into the nature of the HlyA targeting signal

4.1 Introduction:

Now that it had been established that virtually the whole of the Pcm molecule could be specifically secreted from *E. coli*, when fused to the HlyA 23KD C-terminal secretion signal domain, I endeavoured to use this fusion system (Pcm-HlyA) to investigate some of the properties of the HlyA signal domain. Two of the questions to be addressed are whether smaller portions of the HlyA C-terminal signal domain retain sufficient information to continue to support the secretion of the Pcm polypeptide, and secondly whether the HlyA signal domain must be C-terminally located in order to be recognised and promote secretion of the passenger domain.

4.2 Secretion of the Pcm-Sma-12KDHlyA hybrid via the haemolysin export machinery:

4.2.1 Background:

The C-terminal 12KD portion of HlyA can promote the secretion of itself, as well as the majority of the *E. coli* outer membrane protein OmpF (deleted for its NH_2-terminal signal sequence - see Chapter 1; Section 1.6.5.; Mackman *et al.*, 1987). However, can this 12KD domain still be recognised and promote specific secretion when fused to foreign heterologous polypeptides, in this test case, a portion of the mammalian Pcm molecule ?. The 12KD of HlyA is encoded by the final 342bp of *hlyA* and the fragment can be isolated on a 353bp DraI/DraI (blunt ended) fragment from pLG609 (see Chapter 3; Figure 2a). As the DraI fragment is only available in one reading frame the most convenient restriction site within *pcm* to fuse it to, whilst maintaining the open reading frame, was the unique SmaI (XmaIII) restriction site within *pcm* (see Chapter 3; Figure 1). This fusion site has the added advantage that it was previously used to demonstrate the ability of the HlyA 23KD signal domain to secrete the N-terminal (42%) of the Pcm molecule (fused at the same SmaI site; see Chapter 3; Section 3.2.2). Consequently, if the 12KD peptide is also able to promote the secretion of this same domain, the levels of secretion can be directly compared to try to determine the effect, if any, of reducing the size of the HlyA signal domain on the efficiency of secretion.

4.2.2 Generation of a *pcm*-Sma-12KD*hlyA* gene fusion:

Due to the lack of suitable unique restriction sites within pMG168 (DOV-*pcm*) and the need to clone a 353bp *hlyA* blunt ended fragment, the required construct had to be generated through a series of intermediate plasmids (see Appendix 4a/b for details). The final, dual origin based plasmid was named pLG806 (*pcm*-Sma-12KD*hlyA*) and the fusion is

depicted in Figure 1. In the initial construction strategy the final vector carrying the *pcm*-Sma-12KD*hlyA* fusion was to be a tetracycline resistant derivative of the dual origin vector, pLG806Tc (see Appendix 4b). However, problems were encountered in maintaining strains carrying this plasmid at 30°C, possibly due to interference by *tet* transcription on the pSC101 *ori* and *par* functions. Thus, the *tet* gene and pSC101 functions were replaced by the ampicillin resistance gene and pSC101 functions from the dual origin based vector pMG196, generating pLG806 (see Appendix 4b). This plasmid was transformed into SE5000 +/- pLG575 (*hlyB,D*), the resulting strain was grown at 30°C and heat induced for high level expression of the hybrid at an A^{600} =6.6 as described before and in Materials and Methods. Cellular and supernatant samples were taken prior to and 3 hours after heat induction and the protein bands resolved by SDS-PAGE (Figure 2). The induction of a protein of approximately 30KD in size (the predicted size of this inphase chimeric polypeptide) is evident in both cellular samples (Figure 2; Lanes 1-4). Again it is worthwhile remembering that the 23KD HlyA peptide migrates approximately 3-4KD more slowly than predicted from its molecular size. Similarly, the chimeric proteins containing the 23KD domain also run more slowly (Chapter 3; Section 3.2.1). The mobility of the expected hybrid in this case might be expected to migrate somewhat more slowly than predicted from its size. Analysis of the supernatant protein profiles demonstrated the specific secretion of polypeptide in fact about 32KD, after induction of expression, only from the strain expressing HlyB,HlyD (Figure 2, Lane 7-8 versus 5-6). It should also be noted that the secreted chimeric Pcm-12KDHlyA polypeptide is again present in two forms, presumably due to the acid conditions used to concentrate the sample, resulting in some cleavage of the Pcm moiety (see Chapter 3; Section 3.2.2).

4.3 Attempted secretion of the Pcm-Sma fragment fused to the terminal 37 amino acids of HlyA via the haemolysin export machinery:

4.3.1 Background:

The observation that the last 12KD of HlyA can promote the secretion not only of itself but of a heterologous protein fragment, prompted us to investigate whether an even smaller C-terminal portion could duplicate this function and hopefully further define the boundaries within which the targeting signal resides. This experiment had previously been designed with regard to secreting a portion of OmpF (deleted for its signal sequence) fused to the last 37 residues (114bp) of HlyA. This was to be

generated using several oligonucleotides, reconstructing the final 114bp of *hlyA*, kindly supplied by Celltech Ltd. These were assembled and ligated to the *ompF* fragment, deleted as usual for its signal sequence (Karen Baker unpublished results this laboratory) but unfortunately the size of the resulting hybrid was indistinguishable from that of the wild-type OmpF. Hence, the secretion of this hybrid could not be assessed, as small amounts of natural OmpF are apparently released to the media in budded vesicles (Mug-Opstelten and Witholt, 1978; Mackman *et al.*, 1987). In addition Western blot analysis could not be used to distinguish between the two species as the haemolysin antibodies, then available, did not recognise any epitopes within, at least, the C-terminal 12KD segment. Therefore, the Pcm molecule seemed an ideal candidate for this experiment as specific antibodies are available for the Pcm moiety.

4.3.2 Construction of a *pcm*-Sma-114bp*hlyA* fusion:

The generation of this construct was a long and complicated procedure, involving the generation of three intermediate plasmids preceding the construction of the desired end product. This lengthy procedure was due, again, to the absence of convenient or unique restriction sites within the *pcm* expression vector (pMG168; Chapter 3: Figure 1). The details of the construction of this plasmid, pLG805 (*pcm*-Sma-114bp*hlyA*) are shown in Appendix 5a/b with the resulting fusion depicted in Figure 1.

4.3.3 Expression of Pcm-Sma-37residue HlyA chimeric protein in the presence of HlyB,D export proteins:

pLG805 (*pcm*-Sma-114bp *hlyA*) was transformed into SE5000 $^{+/-}$ pLG575 (*hlyB,D*), the strains grown at 30°C and high level hybrid expression heat induced at an A^{600} =0.4 as described previously and in Materials and Methods. Cell and supernatant samples were isolated before, 1.5 and 3 hours post-induction, prepared as usual and resolved by SDS-PAGE (see Materials and Methods). The resulting Coomassie blue stained gel is shown in Figure 3 with lanes 1-3 and 4-6 revealing the induction of a band of the size expected for the hybrid protein (~23KD) in both cellular samples ($^{+/-}$ pLG575; *hlyB,D*), after induction. However, scrutinising the supernatant protein profiles did not reveal any secretion of this polypeptide into the supernatant.

Subsequent findings however rendered it difficult to draw any firm conclusions from this experiment. Thus, comparison of DNA sequencing gels of the 3' end of the LE2001 *hlyA* gene in this study with that of the previously published sequence (Gray *et al.*, 1986), highlighted 2 errors

which had been carried through into the synthesized oligonucleotides used in the construction of the *pcm*-Sma-114bp *hlyA* hybrid. The first was a silent change whereas the second altered a conserved (in closely related haemolysin molecules) phenylalanine (35 residues from the end of the molecule) to a leucine. This substitution alone, from later mutagenesis studies (see Chapter 8; Section 8.2.4: Table 2 Mutant N^{o}. 19) was found to reduce the level of secretion of the 23KD portion of HlyA by 70 to 80%. This alteration may therefore reduce the level of the secreted chimeric protein to an amount too low to be easily detected. This possibility is not unreasonable as we have observed that reducing the signal from 23KD to 12KD results in an approximately 70% decrease in the level of chimeric Pcm-HlyA polypeptide secreted to the supernatant (see Chapter 5; Section 5.5.3). Hence, reducing the size of this signal domain from 12 to 4KD might therefore be expected to further reduce secretion levels. This reduction on the secretion levels, which if compounded by this debilitating mutation, could indeed make the secreted levels too minor to be readily detectable. Therefore, to try to resolve the initial question and by-pass the above mentioned difficulties, I decided to generate a hybrid protein comprising the majority of the Pcm molecule fused to the last 39 (natural) residues of HlyA.

4.4 Construction of a *pcm*-Bal-120bp*hlyA* gene fusion:

The *hlyA* gene carries 2 PvuII restriction sites, one present c.120bp from the 3' end (encoding the final 39 residues). This blunt site was used to make an inframe fusion to the third BalI site within the *pcm* gene, 11bp from the end of this gene, the details of which are shown in Appendix 6. The resulting fusion (predicted to produce a 40KD peptide) is illustrated in Figure 1, with the recombinant plasmid named pLG807 (*pcm*-Bal-120bp*hlyA*).

4.4.1 Fate of the Pcm-Bal-4KDHlyA chimeric polypeptide in the presence of the haemolysin export machinery:

Initially, pLG807 (*pcm*-Bal-4KD*hlyA*) was transformed into the *E. coli* strain NM522 +/- pLG575 (*hlyB,D*), the resulting cultures grown and heat induced as before (see Materials and Methods), for high level hybrid expression upon reaching an A^{450} ~ 0.6. Cell and supernatant samples were isolated before and 3 hours post-induction and separated on an 11% SDS-PAGE (see Figure 4). Induction of high levels of expression results in the accumulation of an approximately 44KD band, which is evident in both cellular samples (Figure 4, Lanes 1-4). Examination of the supernatant samples indicated the secretion of low but significant levels

of this chimeric polypeptide to the supernatant only from the culture also expressing the haemolysin export proteins (Figure 4, compare Lanes 5-6 to 7-8). Although there are signs of low levels of cell lysis (often seen from this particular host strain) this appears to be constant in all supernatant lanes with the 44KD band only present in the supernatant samples obtained from the cultures expressing HlyB and HlyD. The background smear present in the supernatant lanes is due to the TCA precipitation of the nutrient broth peptides. This is demonstrated in Figure 4 (Lane 9), which shows the resulting SDS-PAGE stained profile from nutrient broth alone after TCA precipitation. However, attempts to duplicate the secretion of this 44KD hybrid, with the standard more robust laboratory strain (SE5000) as the host generally used for these secretion experiments failed. Moreover, repeating this experiment using the same strain, NM522, did not result in detectable levels of secreted chimeric. This may reflect weak recognition by the export proteins of the signal motif resulting in inefficient levels of secretion or alternatively rapid sequestration into the inclusion body pathway removing soluble hybrid available for secretion, with the levels of secretion being variable from experiment to experiment. The underlining picture from this series of experiments suggests that the last 39 residues of the HlyA molecule are not capable of promoting efficient secretion when fused to virtually the entire prochymosin molecule. Is this due to the absence of part / all of the recognition motif or due to the absence of other upstream sequences required for the correct presentation of the signal motif to the export machinery in a manner independent of the upstream domain ? One way to answer this question is to determine whether the final 39 residues of HlyA can be secreted in an HlyB,D dependent manner when expressed alone.

4.5 Cloning of the final 120bp of *hlyA* into an expression vector:

4.5.1 Background:

In order to express the 39 C-terminal residues of HlyA it was necessary to clone the encoding DNA fragment in frame with an initiation methionine codon (Met) downstream of a ribosome binding site and an appropriate inducible promoter. It was also desirable to minimise the number of codons fused to the 5' end of the DNA fragment during the construction, so as to minimise any alteration to the size or the properties of the resulting small "4KD" HlyA peptide. The vector chosen was pPLEX (Sczakiel *et al.*, 1987; see Figure 5) which carries a multiple cloning site downstream of the Lambda P_L promoter and a ribosome binding site.

More importantly, the initiation codon (Met) is overlapped by an NcoI restriction enzyme site allowing the subsequent transfer, if required, of the *hly* gene fragment together with the newly acquired initiation Met codon. As there were doubts of the ability to monitor the expression and fate of such a small peptide (4KD), it was decided to also clone and express the C-terminal 15KD encoding fragment in an identical context, as a positive control, as this peptide is known to retain all the information required for its recognition and secretion by the haemolysin export proteins (HlyB,D).

4.5.2 Generation of plasmids expressing the C-terminal 4 and 15KD of HlyA:

The plasmid (pPLEX-4KDHlyA) permitting the expression of the final 39 residues of HlyA was constructed as shown in Figure 5. The DNA fragment encoding this *hlyA* region was isolated from pLG609-1 (see Chapter 3: Figure 2a) on an approximately 1.1 Kb PvuII (blunt)/HindIII fragment (which also carries 40% of the 5' of *hlyB*). In order to generate an inframe fusion as close as possible to the initiation Met codon in pPLEX, this vector was digested with AccI, the 3' overhang filled in by T_4 DNA polymerase (see Materials and Methods), and then digested with HindIII. The *hlyA* PvuII/HindIII fragment was then ligated in, the resulting plasmid being named pPLEX-4KDHlyA (Figure 5). This plasmid encodes the last 39 residues of HlyA preceded only by a valine (Val) and the initiation Met codons derived from pPLEX. As it is not unusual for fill in reactions to produce aberrant fusion sites, double stranded plasmid DNA was isolated and purified by CsCl gradients and the expected nature of the fusion site confirmed by double stranded sequencing using the USB sequenase kit (data not shown).

The plasmid allowing the expression of the last 148 residues of HlyA (predicted molecular weight ~15KD) again only preceded by a Val and initiation Met codons is depicted in Figure 5. This construction was simplified due to the previous introduction of an unique ClaI site c.447bp from the 3' end of *hlyA* (encoding the last 148 amino acids) by site directed mutagenesis (see Chapter 7; Section 7.2.3.). This restriction site allowed this 447bp *hlyA* fragment to be cloned directly into the compatible AccI site within the multiple cloning site of pPLEX generating an inframe fusion to the Met codon. The 3' of *hlyA* was isolated on a 600bp ClaI/SmaI(blunt) fragment from pACYC-CA,ClaI (see Chapter 7; Section 7.5.1) and ligated into the AccI/HpaI(blunt) sites in pPLEX, generating pPLEX-15KDHlyA (see Figure 5).

4.5.3 Investigation of the ability of the *hly* export machinery to recognise and secrete the 4KD and 15KD C-terminal peptides of HlyA:

In order to regulate the expression of the 4 and 15KD peptides from the pPLEX lambda P_L promoter, the plasmids had to be introduced into a strain expressing the CI^{ts} repressor gene, *E. coli* ΔH1 being chosen for this purpose. Overnight cultures of these strains +/- pLG575 (*hlyB,D*) were grown at 30°C and subcultured into 75ml of fresh Nutrient broth to an A^{450} ~0.2. The cultures were grown to an A^{450} ~0.7, at which point cell and supernatant samples were taken prior to heat inducing expression, by incubating at 42°C for 10 minutes. Growth was continued at 37°C for 3 hours and further cell and supernatant samples taken, and analysed on a 20% acrylamide-SDS gel. The protein profiles were visualised either by Coomassie blue staining (data not shown) or Western blot analysis using anti-23KDHlyA antibodies (See Figure 6a/b). The stained samples did not reveal the intracellular accumulation of either a 4 or 15KD band after induction (data not shown). Moreover, no such protein bands were detected by Western blot analysis of the cellular samples using anti-23KDHlyA antibodies (Figure 6a/b, Lanes 1-4) although the antibodies did cross-react with a band of about 25KD in size (see below). Analysis of the Coomassie blue stained supernatant protein profiles revealed a subset of cellular bands in all lanes, which appear to signify some degree of cellular lysis (data not shown). This could be due either to the particular host strain being used or the presence of this recombinant plasmid.

When the same supernatant samples were probed with anti-23KDHlyA antibodies by Western blot analysis it was found that the 15KD HlyA peptide was specifically secreted in an HlyB,D dependent manner (Figure 6a; Lanes 5-8). This protein could not have been released by cellular lysis as it is not stable within the cells as indicated by the inability to detect this band by Western analysis in the cell samples (Figure 6a; Lanes 1-4). Indeed, the 15KD HlyA peptide appeared even to be unstable when secreted, indicated by the level of secretion and two forms detected on stained gels and Western blots (Figure 6b, Lane 11). The, Western blot analysis also failed to detect either a cellular or a secreted 4KD HlyA related peptide, using antibodies raised against the 23KDHlyA molecule. Nevertheless, the antibodies did recognise the control hybrid protein (composed of 99% of the Pcm molecule fused to the last 4KD of HlyA; See Section 4.4: pLG807), showing the capacity of the anti-23KDHlyA

antibodies to recognise epitopes in this C-terminal region at least under certain conditions (Figure 6b; Lane 10; also see Chapter 3; Figure 10b). Surprisingly, in these experiments the antibodies also appeared to cross react with an unrelated band of approximately 25KD which might have been co-purified with the 23KD HlyA peptide used to generate these antibodies.

These results demonstrated that the 15KD fragment as found previously for the 23KD and 12.5KD C-terminal fragments could also be secreted to the medium in an HlyB,D dependent manner. In contrast no evidence for secretion of the 4KD peptide was found. This suggested that whereas the export machinery could recognise and export the C-terminal 15KD of HlyA to the medium and thus protect it from rapid degradation, the final 4KD peptide did not appear to be secreted. This initially suggested that this small peptide either did not encode all the information required for recognition and secretion or was too unstable to be detected and or secreted. Unfortunately, the non-specific release of cellular proteins to the medium could contribute to the inability to detect minor levels of secreted 4KD peptide (directly masking identification of a small 4KD band or through the release of proteases). Consequently, it was decided to repeat these experiments after transferring the DNA encoding these peptides to another expression vector and under the control of a powerful promoter other than a heat inducible one, which often results in increased cell lysis and protein degradation.

4.6 Expression of the 4 and 15KD C-terminal peptides of HlyA from an IPTG inducible T7 RNA polymerase promoter:

4.6.1 Background:

In an attempt to obtain maximal expression of the two HlyA C-terminal peptides, I decided to clone the encoding fragments into a vector carrying a T7 RNA polymerase promoter (under the control of the *lac* repressor gene product) which directs high level expression in the presence of T7 RNA polymerase. The vector chosen was pET11d (Studier et al., 1990; see Figure 7) which carries a T7 promoter, a ribosome binding site, the initiating Met residue plus 10 subsequent codons of the highly expressed T7 gene S10. This Met codon also corresponds to the cleavage site of the restriction enzyme NcoI which allows the T7 gene S10 fragment to be substituted for the *hlyA* fragment cloned previously into pPLEX (see above). Just upstream of the T7 promoter is the *lacI* repressor gene which in turn only allows expression by the T7 RNA polymerase when IPTG is

added. Such, a system provides for tight regulation of expression to reduce any possible deleterious effects of the polypeptide under study.

Specific, high level expression from the T7 promoter as indicated above requires T7 RNA polymerase which is encoded in strain BL21 downstream of a *lac* promoter in the prophage DE3 inserted into the host chromosome. Consequently, addition of IPTG results both in the expression of the T7 RNA polymerase gene and de-repression of the hybrid T7-*lac* promoter allowing powerful and specific expression of the cloned *hlyA* fragments.

4.6.2 Subcloning DNA fragments encoding the final 4 and 15KD fragments of HlyA into the pET (T7) expression system:

As mentioned previously, the DNA fragments encoding the final 4 and 15KD of HlyA plus the additional valine and initiation methionine codons could be easily transferred from the pPLEX vectors (see Section 4.5.2) due to the presence of an NcoI site overlapping the Met codon, and cloned into the pET11d vector using a similar NcoI site (Figure 7). However, the only other available cloning site in pET11d was a BamHI site, so it was decided to fill in this site, using T4 DNA polymerase (see Materials and Methods), thus allowing insertion of the *hlyA* fragments on NcoI/blunt ended fragments. Thus, the 275bp NcoI/HpaI (blunt end) fragment from pPLEX-4KDHlyA was ligated into pET11d (NcoI/BamHI fill in) to generate pET11d-4KDHlyA.

Unfortunately, construction of pET11d-15KDHlyA was more difficult to generate as no convenient blunt site was present at the 3' of the *hlyA* gene within pPLEX, so the downstream KpnI (see Figure 5) site was blunted, again using T4 DNA polymerase (see Materials and Methods), before ligating into pET11d (see Figure 7).

4.6.3 Studies of the secretion of the 4 and 15KD C-terminal HlyA peptides with the T7 expression system:

pET11d-4KDHlyA and pET11d-15KDHlyA were transformed into BL21 (DE3 - encoding the inducible T7 RNA polymerase gene) +/- pLG575 (*hlyB,D*). Overnight cultures were used to inoculate 100ml of nutrient broth to an A^{450} of 0.2 and grown at 37°C to A^{450} of approximately 0.7, before inducing expression of T7 RNA polymerase by the addition of IPTG (0.5mM final concentration). Cell and supernatant samples were taken prior to and 3 hours post induction and prepared as usual (see Materials

and Methods). The samples were separated on a 10-18% gradient gel containing 0-10% sucrose (as described by Hashimoto *et al.*, 1983; see Materials and Methods). This gel system was chosen as it is supposedly capable of resolving proteins in the broad range of 1.5KD to 100KD, especially showing high resolving power in the 1.5 to 25KD molecular weight range. The resulting Coomassie blue stained gel is shown in Figure 8. Lanes 9-12 show the cellular samples from the strains carrying pET11d-15KDHlyA with lanes 9,10 (minus export functions) showing the induction of a minor ~15KD cellular band. This band was not detected in lysates from the corresponding strain containing export functions (Lanes 11,12), but instead appeared to be secreted to the supernatant (see Lane 16).

Analysis of lanes 1-8 did not reveal the induction or secretion of a 4KD peptide, although it is not apparent whether this is due to the inability to stain this peptide or indeed resolve it under these gel conditions. Unfortunately, TCA precipitation in rich broth, to concentrate supernatant samples, leads to a high background of small peptides present in the broth (see Section 4.4.1), perhaps obscuring the detection of low level secretion. To avoid this problem, it was decided to repeat this experiment in M9-glucose minimal medium which lacks such peptides. It is worthwhile noting here the unexpected, apparently specific release/ secretion, of two novel protein bands (approximately 46 and 30KD) to the medium after IPTG induction from this *E. coli* B strain. These may well be related to the presence of the DE3 prophage in the strain but this was not further investigated.

4.6.4 Studies of the secretion of the 4 and 15KD C-terminal HlyA peptides from the T7 expression system in M-9 glucose Minimal Medium:

The same strains as above were streaked out on M-9 glucose minimal media plates and used to inoculate overnight cultures. The following day 100ml of fresh M9 glucose Minimal Medium was inoculated to an A^{450} = 0.2 and growth continued with shaking at 37°C. At an A^{450} = 1 cell and supernatant samples were taken for analysis and the remaining cultures induced for T7 RNA polymerase expression by the addition of IPTG (0.4mM final concentration). Growth was only continued for a further hour before a second set of cellular and supernatant sample were taken for analysis on a 20% acrylamide-SDS gel (see Figure 9a). Lanes 1-4 and 5-8 show the Coomassie blue stained cellular profiles of the strains expressing the 4 and 15KD HlyA peptides, respectively, but there is no

sign of either of the peptides accumulating within the cells even under the very specific high level induction system used. In contrast, comparison of Lanes 13-14 with 15-16 demonstrates the specific secretion of the 15KD peptide in an HlyB,D dependent manner (its identity confirmed by Western analysis; Figure 9b; Lane 16, again noting a smaller, presumed, breakdown product) whereas the 4KD peptide was not detectable in either the cell or supernatant samples (Figure 9a; Lanes 1-4 and 9-12, respectively). The Western blot analysis was carried out on the samples resolved in a 25% SDS-PAGE gel system, as described by Fling and Gregson (1986) except a gradient gel was not used since we were only interested in the low molecular weight protein bands. The amido black stained Low and High molecular markers used demonstrated the ability to resolve only 2 of the 3 smallest markers (6.2, 3.4 and 2.3KD). The smallest two marker bands are the insulin A and B chains and the above authors report that the 2.3KD band does not stain with at least Coomassie blue or silver stain. They suggested that either extremely analogous migration, leaching from the gel or lack of banding under the conditions used as possible reasons for lack of detection of certain low molecular weight peptides. The authors also report the apparent faster migration of other small molecular weight protein markers under these gel conditions.

The results described in this section indicate that the 4KD peptide is highly unstable within the cells and also lacks sufficient information for efficient recognition and hence secretion via HlyB,D. However, the results do not rule out the possibility that the protein is degraded before secretion can take place. In addition it remains a possibility that the 4KD HlyA peptide may indeed be secreted but fails to be resolved, stained or fixed by the gel systems used (as reported for other low molecular weight proteins; Fling and Gregson, 1986) and possibly may not be transferred to or remain on the nitrocellulose during Western blot analysis.

4.7 Analysis of position on the function of the HlyA targeting signal:

4.7.1 Background:

The novel HlyA signal domain is C-terminally located in nature but is this an absolute requirement for recognition and secretion by the export pathway ?. To obtain an insight into this question I designed a strategy to alter the stop codon of *hlyA* in order to generate an EcoRI restriction site, now permitting its novel fusion upstream of a portion of the *pcm* gene.

4.7.2 Generation of the 23KDHlyA-Pcm hybrid protein:

In another series of experiments (see Chapter 6; Section 6.2.3) the *hlyA* 1.6Kb SmaI/HindIII fragment (encoding the C-terminal 23KD of HlyA and 42% of the N-terminal portion of HlyB) had been subcloned from pLG609-1 (see Section 3.2.1: Figure 2b) into the bacteriophage M13mp18 double stranded vector (see Appendix 7a). This M13mp18-23KDHlyA clone was then used to generate single stranded DNA for use in a site directed mutagenesis reaction (see Materials and Methods). A synthetic 21-mer oligonucleotide was designed incorporating 2 alterations which would generate an EcoRI site overlapping the *hlyA* stop codon triplet. The Amersham site directed mutagenesis kit was used successfully to generate the required site as confirmed by restriction and sequence analysis (data not shown). The altered 1.6Kb SmaI/HindIII fragment was then subcloned into the corresponding sites of pUC12 for expression studies from the inducible *lac* promoter (see Appendix 7a), the resulting plasmid being named pLG618 (pUC-23KDHlyAEcoRI).

This experiment was designed to generate an inframe fusion within the *pcm* gene using the internal EcoRI site, resulting in the deletion of the 5' (43%) of the *pcm* gene. Thus, the 23KD polypeptide of HlyA would be fused at the N-terminus to the final c.20KD of Pcm. The required gene fusion is depicted in Figure 1, the construction of which, generating pLG803 (pUC23KDHlyA-'Pcm), is shown in Appendix 7b. This was achieved by ligating the 3' portion of *pcm* (isolated from pMG168 on a 0.9Kb EcoRI/HindIII fragment; see Appendix 7b) into the corresponding sites at the 3' end of the *hlyA* fragment in pLG618, generating pLG803 (see Appendix 7b). The plasmid was analysed with a variety of restriction enzymes to verify the correct orientation of insertion of the vector derived EcoRI/EcoRI fragment.

4.7.3 Investigation of the ability of the HlyA signal domain to promote secretion when placed upstream of a passenger peptide:

As an additional test for the effect of position on the function of the HlyA signal domain, advantage was taken of the plasmid pLG618 described in the previous section (pUC-23KDHlyAEcoRI; see Appendix 7a). The mutagenesis at the *hlyA* stop codon generating the EcoRI site in this construct alters this codon and results in subsequent read through creating a novel 29 residue C-terminal extension. This mutant HlyA-C-terminal peptide consequently allows us to also study the effect of a small extension on secretion. Thus, pLG618 (pUC-23KDHlyAEcoRI) and

pLG803 (pUC-23KDHlyA-'Pcm) plasmids were individually transformed into *E. coli* JM101 (carries *lacI*q gene to repress expression from the *lac* promoter) +/- pLG575 (*hlyB,D*). Overnight cultures of these strains were used to inoculate fresh nutrient broth to an A^{450} of 0.1 and growth initiated by shaking at 37°C. At an A^{450} of 0.5, cellular samples were removed for future analysis and the expression of the cloned genes induced with IPTG (0.5mM final concentration). Cell and supernatant samples were also taken after a further 2 hours of growth and prepared as usual (see Materials and Methods). The samples were separated on a 15% SDS-polyacrylamide gel and visualised by either Coomassie blue staining (data not shown) or Western blot analysis probing with antibodies directed against HlyA (see Figure 10). The stained gel showed the induction of the 23KDHlyA-Pcm chimeric protein (expected size ~47KD) in the cellular samples but not that of the pLG618 (pUC23KDHlyA-3KD extension) gene product (expected size ~32KD; data not shown). Neither of these gene products appeared to be secreted to the external medium but detection is obscured by minor background lysis proteins (data not shown). However, Western blot analysis using anti-HlyA antibodies detected the expected gene products inside the cells, after IPTG induction (Figure 10; Lanes 1-4 and 5-8, respectively) but not in the supernatant samples (Lanes 9-12). Lane 13 is a positive control represented by the chimeric Pcm-Bal-23HlyA fusion protein (see Chapter 3; Section 3.3.1). Therefore, it can be concluded that the addition of these particular peptides to the C-terminus of the HlyA secretion domain strongly inhibits recognition and/or secretion by the transport machinery.

4.8 Discussion:

An important observation from the above results is the positional effect of the HlyA targeting domain upon secretion. First of all, the 23KD signal domain cannot support the secretion of a C-terminally located 20KD domain of Pcm. From previous studies (see Chapter 1; Section 1.6.5.) it has been inferred that at least the last 27 residues constitute an important part of the signal motif (Gray *et al.*, 1986; Mackman *et al.*, 1987). Therefore, it is possible that the downstream Pcm peptide in this hybrid might assume some tertiary structure masking/impeding recognition of the signal motif. However, the C-terminal addition of only 29 residues (~3KD) to the 23KD HlyA signal peptide also resulted in a secretion minus phenotype. This suggests that it is indeed important that the signal domain is C-terminally located, the addition of extra amino acids presumably hindering the adoption of the correct tertiary structure,

its presentation to and/or access of the targeting signal to the membrane translocator protein. It should also be mentioned that the feature of the new C-terminal 29 residues extension

(EFINLNNSNLTGLCHIRLLFFWSHNGFLS)

is that of a strongly hydrophobic tail with a high proportion of large amino acids. This contrasts with the natural C-terminal tail which is weakly hydrophobic and contains a high incidence of hydroxylated and small residues. Hence, it cannot be ruled out that it is the specific nature of this extension which inhibits translocation through the haemolysin export channel rather than non-specific interference with the structural organisation of the target domain. Nevertheless, the export machinery could not secrete the hybrid consisting of the final 23KD of HlyA fused at the N-terminus of a 20KD C-terminal segment of Pcm, even though it has been demonstrated that the export pathway has the capacity to secrete this entire molecule (missing only its final 3 residues). Thus, it would appear that it is the position of these HlyA C-terminal extensions and not their nature that is the primary inhibition to the secretion process.

Studies concerning the boundaries of the HlyA signal motif by reduction of the HlyA C-terminal portion fused to Pcm demonstrated that the signal can be reduced from 23KD to the final 12KD and still promote efficient secretion, whereas further reduction to 4KD leads to virtual elimination of secretion. This tends to suggest that the signal motif encompasses at least the final 39 amino acids, but that other upstream residues are required either as an inherent part of the signal or to ensure that the signal domain is allowed to fold independently of the rest of the molecule. Another possibility that has been investigated is that the failure of the final 4KD (39 residues) of HlyA to promote efficient secretion of an N-terminal portion of the Pcm polypeptide is due to occlusion of the signal by the large N-terminally located heterologous domain. However, the inability to detect secretion of the C-terminal 4KD HlyA peptide alone (unlike the control 15KD HlyA peptide) even under conditions of presumed high level expression also suggests that the last 39 residues do not encode all the information required for recognition and secretion by the HlyB,D system.

Experiments to test the ability of a small 4KD C-terminal peptide of HlyA to be secreted proved to be quite difficult to interpret given the extreme lability of such small peptides which prevents their accumulation inside cells when secretion is prevented. Thus, we cannot rule out the possibility that the 4KD peptide is too unstable within the cells to permit it to be

secreted, or that it is unstable even after being secreted, as appears to be the case with the 15KD HlyA peptide where secretion can be detected (see Figure 6b; lane 11). Another possibility is that the 4KD peptide is running aberrantly in SDS-acrylamide gels (even beyond the dye front and out of the gel), may not cross react sufficiently to be detected by the anti-23KDHlyA antibodies or may not stain with dyes as has been documented for other low molecular weight peptides (see Section 4.6.4).

In contrast to results with the HlyA molecule very recently it has been reported that the final 39 residues of an *Erwinia* protease is sufficient to allow the secretion of this peptide (fused to 12 residues of beta-galactosidase), via the Erwinia HlyB,D homologues (Letoffe *et al.*, 1990; C. Wandersman; personal communication). Indeed, this secreted 6KD peptide was not detected on stained gels (as has been documented for other small peptides; see Section 4.6.4), but was identified using antibodies specific to the protease C-terminal domain. Thus, it is possible that the HlyA targeting signal also resides entirely within the final 39 residues but that we did not detect the secreted small peptide due solely to the poor antigenic nature of the HlyA C-terminal portion (Mackman *et al.*, 1987; unpublished observations this laboratory). However, as noted elsewhere in this thesis (see Chapter 6; Section 6.1) the targeting signals are not conserved at the primary sequence level and therefore direct comparisons cannot be made. The Wandersman group in fact also report that the Hly export proteins, HlyB,D, can promote the secretion of the *Erwinia* protease only at approximately 1% of HlyA levels, while the Protease HlyB,D homologues cannot translocate either the HlyA toxin or the C-terminal 23KD HlyA peptide. This suggests that although the C-terminal targeting signals are related they also differ markedly in their specificity, implying distinctive recognition features. It must also be considered, in relation to the apparent localisation of the *Erwinia* signal motif, that in the secretion of the final 39 residues these were fused to 12 residues of beta-galactosidase which might fortuitously have restored some structural feature required for secretion, absent from the final 39 residues. Therefore, it is still possible that the HlyA targeting signal does not indeed reside within the final 39 residues but requires other upstream sequences for efficient recognition.

Thus, we can not conclude whether the HlyA signal motif resides within the final 39 residues, with many plausible reasons for our inability to detect the small secreted peptide, or whether sequences upstream are truely required not only for efficient secretion but also for recognition.

However, preliminary results from probing the signal region by mutagenesis (see Chapter 8) suggest that, for the Hly system at least, other upstream residues are also required for efficient recognition and thus secretion.

Figure 1:
Diagrammatical representation of the *pcm-hlyA* gene fusions generated in this chapter.
Filled in boxes represent regions of the *pcm* gene.
Dotted lines represent regions deleted in the fusion.
Open boxes represent the *hlyA* region encoding the C-terminal 23KD segment.
Stippled box represents extent of readthrough due to mutation of stop codon in pLG617.

Pcm-HlyA hybrid fusion-site junctions:

	Pcm	C-terminal HlyA
pLG806	ACC CAG GAG CCC	AAA AAC TGG TTT
	Thr Gln Glu Pro	Lys Asn Trp Phe
pLG805	ACC CAG GAG CCC	GGT AAC TTC GAT
	Thr Gln Glu Pro	Gly Asn Leu Glu
pLG807	AAG CTG GTG G	CT GCA GGT AAC
	Lys Leu Val(Gly) Ala	(Ala)Ala Gly Asn
	C-terminal HlyA	Pcm
pLG803	GCA TCA GCA GAA	TTC GAC GGG ATC
	Ala Ser Ala Glu	Phe Asp Gly Ile

Underlining = errors introduced when reconstrucing 3' *hlyA* via oligonucleotides as synthesis according to incorrect sequence data (see text)
Double underlining = conversion of stop codon by site directed mutagenesis (see text)
Brackets = natural residue at this position prior to fusion

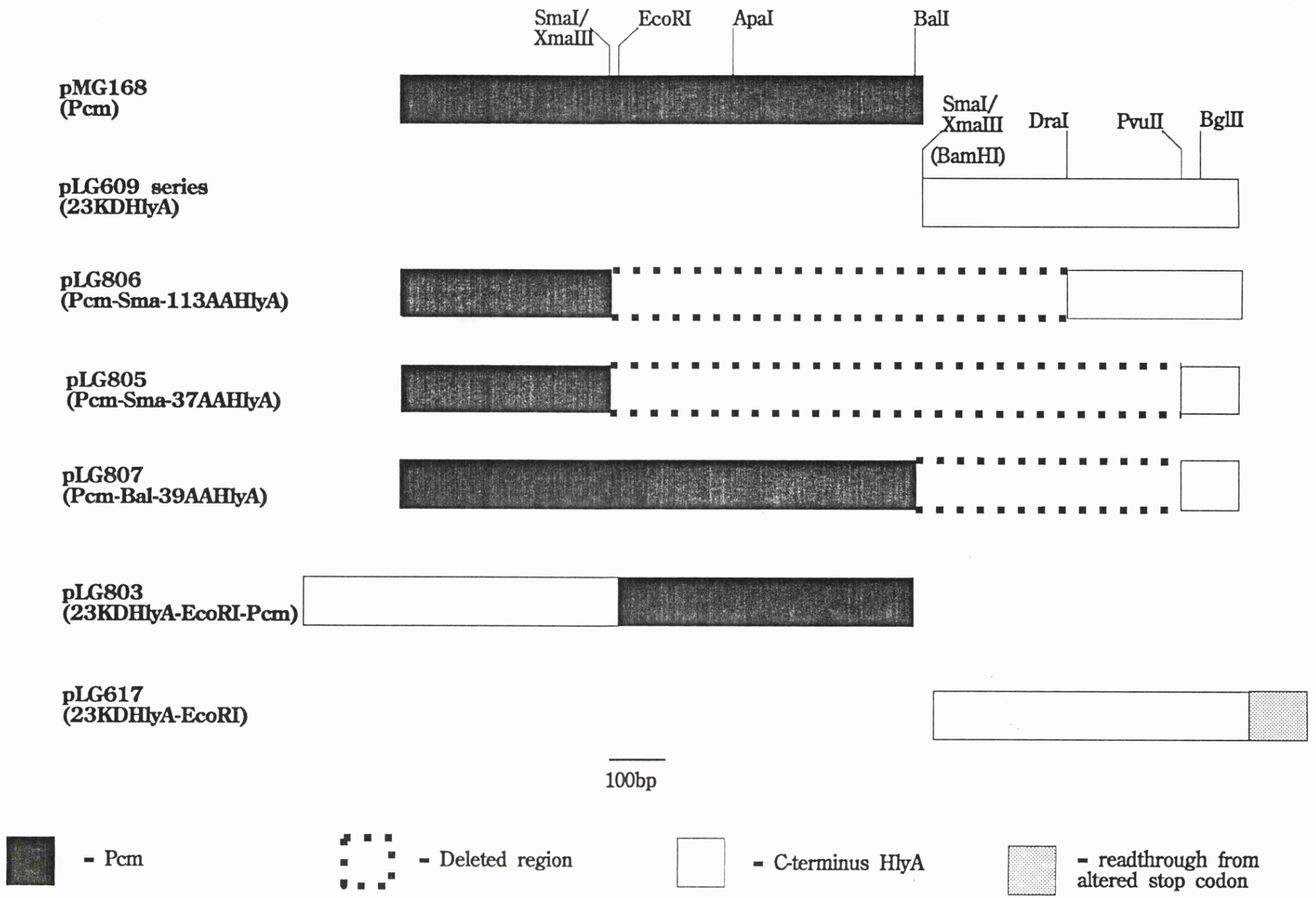

SmaI/
XmaIII
EcoRI
ApaI
BalI
pMG168
(Pcm)
SmaI/
XmaIII
(BamHI)
DraI
PvuII
BglII
pLG609 series
(23KDHlyA)
pLG806
(Pcm-Sma-113AAHlyA)
pLG805
(Pcm-Sma-37AAHlyA)
pLG807
(Pcm-Bal-39AAHlyA)
pLG803
(23KDHlyA-EcoRI-Pcm)
pLG617
(23KDHlyA-EcoRI)
100bp
- Pcm
- Deleted region
- C-terminus HlyA
- readthrough from
altered stop codon

Figure 2:

Secretion of a 32KD Pcm-Sma-12KDHlyA hybrid protein from *E. coli* in an HlyB,D dependent manner.

E. coli SE5000 containing pLG806 (*pcm*-Sma-12KD*hlyA*) +/- pLG575 (*hlyB,D*) was grown in Luria broth (with appropriate antibiotics) to an A^{600} of 0.66 before heat inducing high level expression. Samples were taken before and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.1 unit of cells (where 1 unit is defined as 1ml at A^{600} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were analysed on a mini 11% SDS-acrylamide gel and visualised by Coomassie blue staining.

Cellular samples: lane 1 and 2, time = 0 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 3 and 4, time = 0 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 5 and 6, time = 0 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 7 and 8, time = 0 and 3 hours, respectively, from the culture **plus** pLG575 (*hlyB,D*).

Arrows indicate the 2 forms of the secreted 32KD hybrid protein with molecular weight markers also shown.

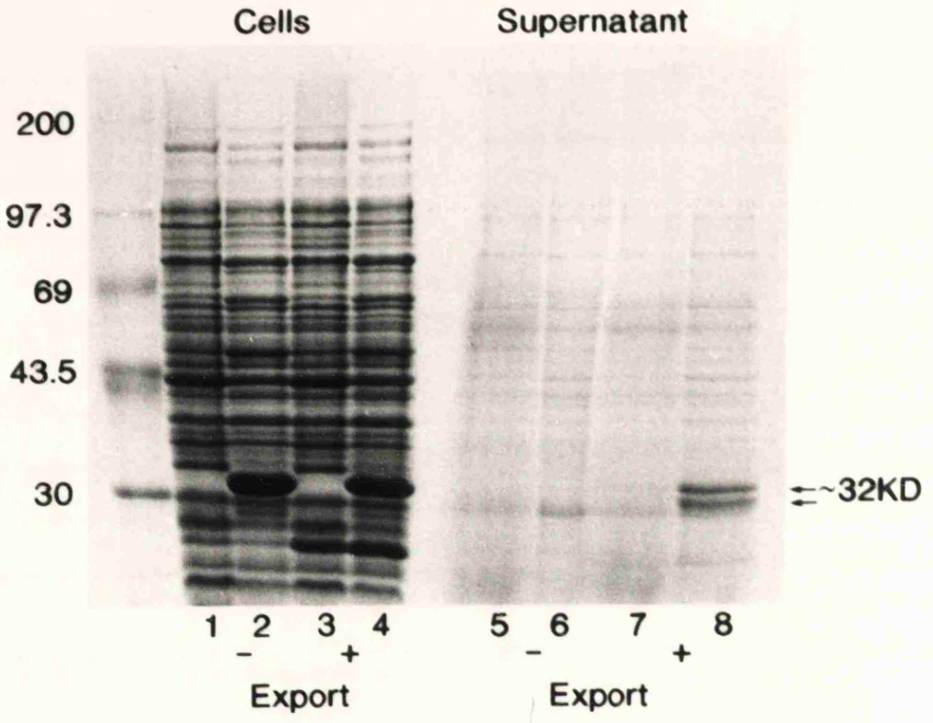

Cells
Supernatant
200
97.3
69
43.5
30
~32KD
1 2 3 4
5 6 7 8
-
+
-
+
Export
Export

Figure 3:

Non-secretion of a 23KD Pcm-Sma-37residue HlyA hybrid protein from *E. coli*.

E. coli SE5000 containing pLG805 (*pcm*-Sma-37residue *hlyA*) +/- pLG575 (*hlyB,D*) was grown in nutrient broth (with appropriate antibiotics) to an A^{600} of 0.4 before heat inducing high level expression. Samples were taken before, 1.5 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{600} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were separated on a 15% SDS-acrylamide gel and visualised by Coomassie blue staining.

Cellular samples: lane 1-3, time = 0, 1.5 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 4-6, time = 0, 1.5 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 7-9, time = 0, 1.5 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 10-12, time = 0, 1.5 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Arrow marks the position of the 23KD Pcm-HlyA hybrid protein with molecular weight markers also shown. The highly expressed band of ~25KD in lanes 4-6 represents the CAT gene product expressed from pLG575 (*hlyB,D*).

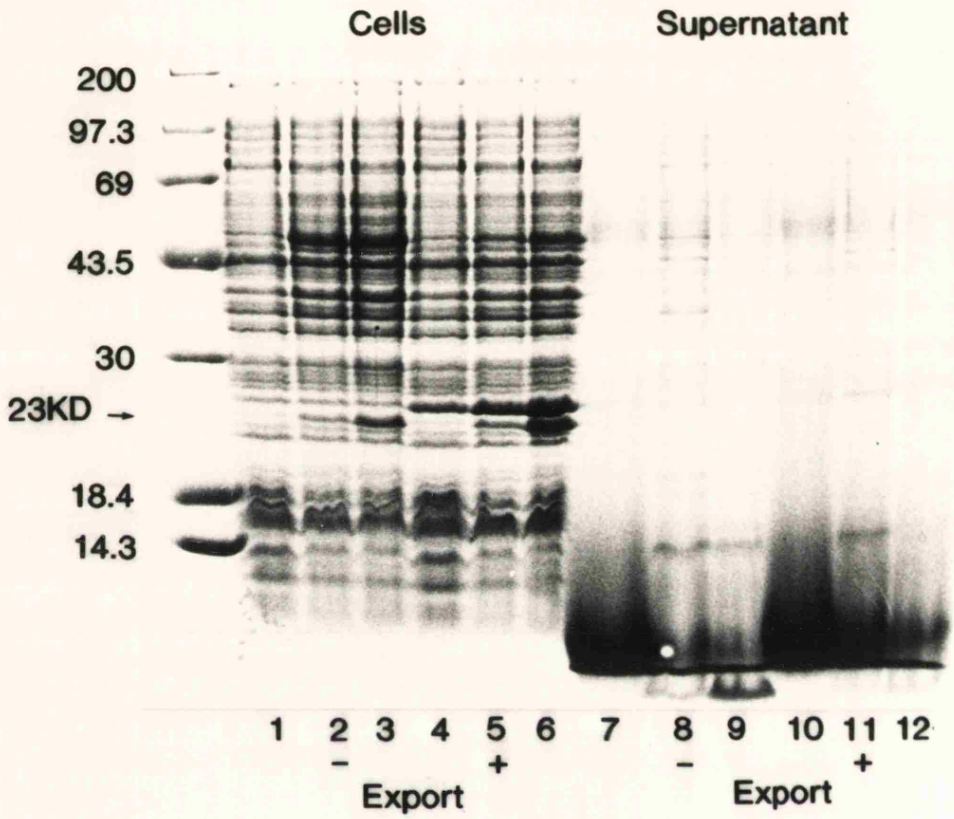

Cells
Supernatant
200
97.3
69
43.5
30
23KD
18.4
14.3
1 2 3 4 5 6 7 8 9 10 11 12
- + - +
Export
Export

Figure 4:

Secretion of a 44KD Pcm-Bal-39residueHlyA hybrid protein from *E. coli*.

E. coli NM522 containing pLG807 (*pcm*-Bal-39residue*hlyA*) $^{+}/_{-}$ pLG575 (*hlyB,D*) was grown in nutrient broth (with appropriate antibiotics) to an A^{450} of 0.6 before heat inducing high level expression. Samples were taken before and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were separated on a 11% SDS-acrylamide gel and visualised by Coomassie blue staining.

Cellular samples: lane 1 and 2, time = 0 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 3 and 4, time = 0 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 5 and 6, time = 0 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 7 and 8, time = 0 and 3 hours, respectively, from the culture **plus** pLG575 (*hlyB,D*).

Lane 9 represents the peptide profile obtained when TCA is added to sterile nutrient broth and the resulting pellet run on SDS-polyacrylamide gels.

Arrow marks the position of the 44KD hybrid protein with molecular weight markers also shown.

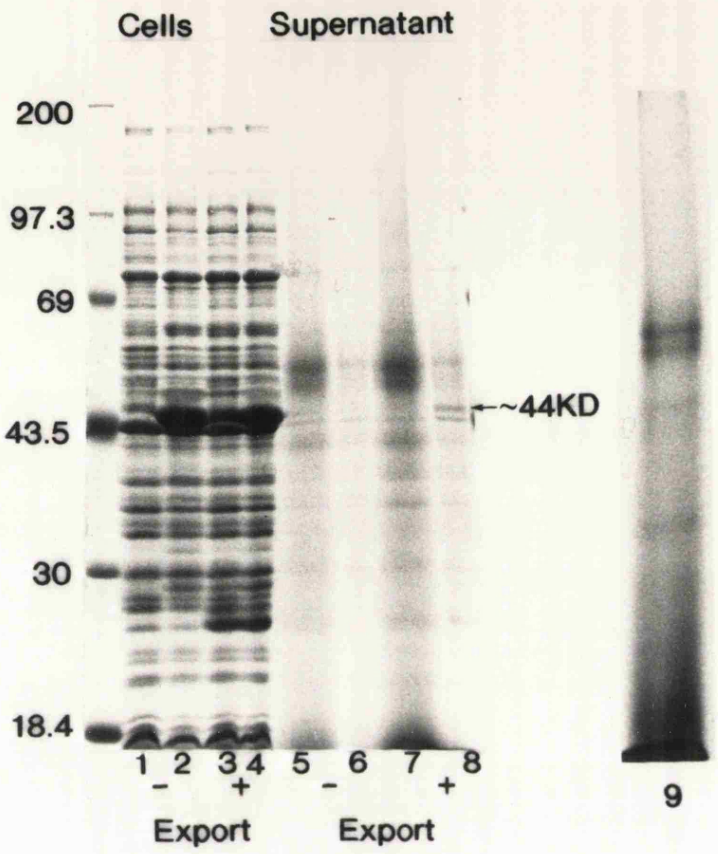

Cells
Supernatant
200
97.3
69
43.5
30
18.4
~44KD
1 2 3 4 5 6 7 8
9
- + - +
Export
Export

Figure 5:
Construction of plasmids permitting the controlled expression of the C-terminal 4 and 15KD domains of HlyA.
See text for details - Section 4.5.2
pPLEX: Sczakiel *et al.*, 1987

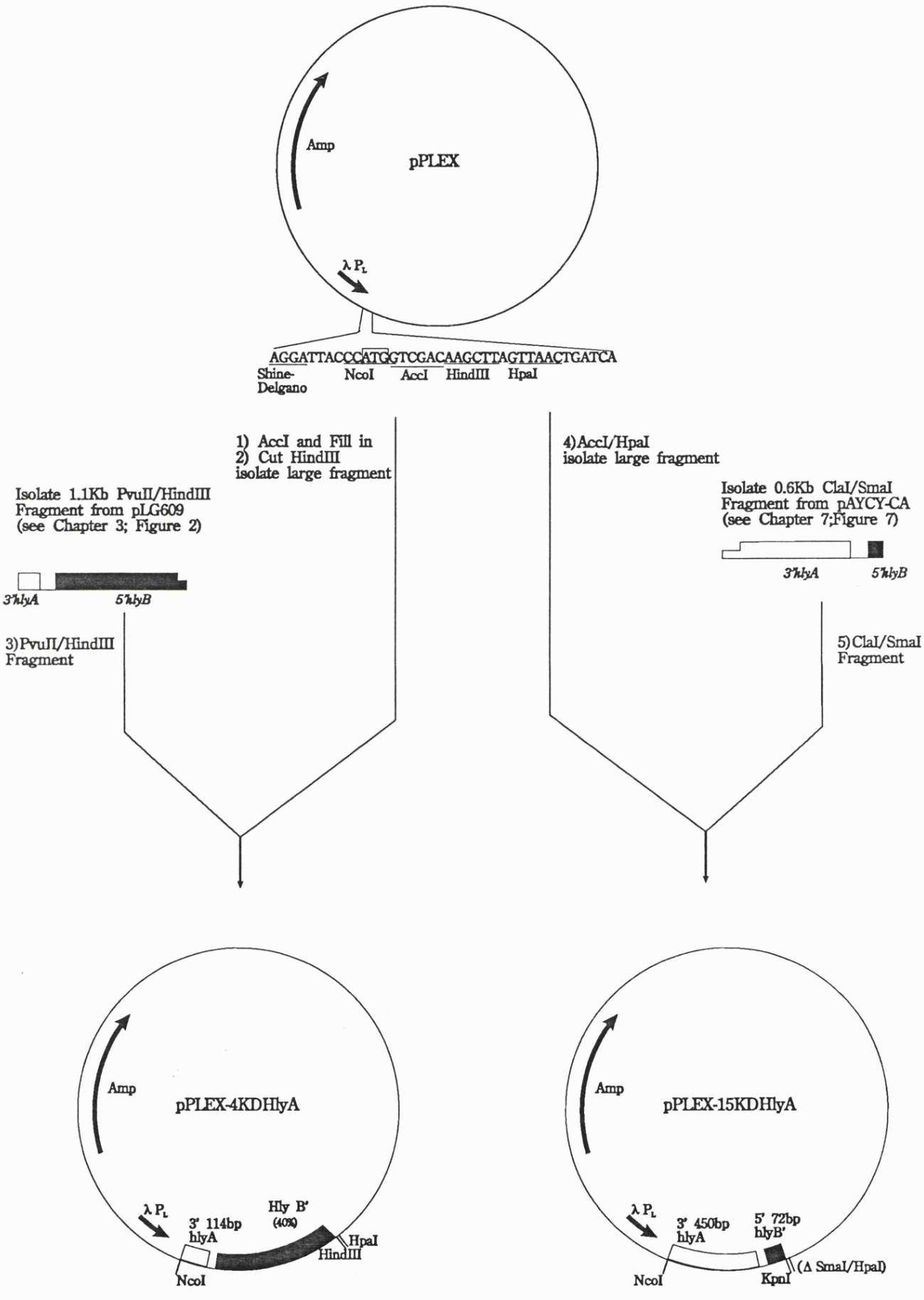

Amp
pPLEX
λ P_L
AGGATTACCCATGGTCGACAAGCTTAGTTAACTGATCA
Shine-Delgano
NcoI
AccI
HindIII
HpaI
1) AccI and Fill in
2) Cut HindIII
isolate large fragment
4) AccI/HpaI
isolate large fragment
Isolate 1.1Kb PvuII/HindIII
Fragment from pLG609
(see Chapter 3; Figure 2)
3'hlyA
5'hlyB
Isolate 0.6Kb ClaI/SmaI
Fragment from pAYCY-CA
(see Chapter 7;Figure 7)
3'hlyA
5'hlyB
3) PvuII/HindIII
Fragment
5) ClaI/SmaI
Fragment
Amp
pPLEX-4KDHlyA
λ P_L
3' 114bp
hlyA
Hly B'
(40%)
HpaI
HindIII
NcoI
Amp
pPLEX-15KDHlyA
λ P_L
3' 450bp
hlyA
5' 72bp
hlyB'
NcoI
KpnI
(Δ SmaI/HpaI)

Figure 6:

Western analysis of SDS-PAGE protein profiles from cultures expressing the C-terminal 4 and 15KD segments of HlyA in the presence / absence of the Hly export proteins.

E. coli ΔH1 cultures containing pPLEX-4KDHlyA $^{+}/_{-}$ pLG575 (*hlyB,D*) and pPLEX-15KDHlyA $^{+}/_{-}$ pLG575 (*hlyB,D*) were grown at 30°C in nutrient broth (with appropriate antibiotics) to an A^{450} of 0.7 before heat inducing expression. Cell and culture supernatant samples were taken before and 3 hours post-induction and prepared as described in Materials and Methods. 0.1 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were separated on a 20% mini SDS-acrylamide gel, Western blotted and probed with anti-23KDHlyA antibodies.

a) pPLEX-15KDHlyA $^{+}/_{-}$ pLG575 (*hlyB,D*) cultures

Cellular samples: lane 1 and 2, time = 0 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 3 and 4, time = 0 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 5 and 6, time = 0 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 7 and 8, time = 0 and 3 hours, respectively, from the culture **plus** pLG575 (*hlyB,D*).

Controls: lanes 9 and 10, cellular samples from previous experiments induced for high level expression of a hybrid protein encoding the C-terminal 23KD and 4KD of HlyA in both cases fused to 99% of the Pcm molecule, respectively.

Arrow marks the position of the 15KD HlyA protein with molecular weight markers also shown.

b) pPLEX-4KDHlyA $^{+}/_{-}$ pLG575 (*hlyB,D*) cultures

Cellular samples: lane 1 and 2, time = 0 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 3 and 4, time = 0 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 5 and 6, time = 0 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 7 and 8, time = 0 and 3 hours, respectively, from the culture **plus** pLG575 (*hlyB,D*).

Controls: lanes 9 and 10, cellular samples from previous experiments induced for high level expression of a hybrid protein encoding the C-terminal 23KD and 4KD of HlyA fused to 99% of the Pcm molecule, respectively.

Lane 11 is identical to lane 8 in Figure 6a above and run as an extra control and size marker

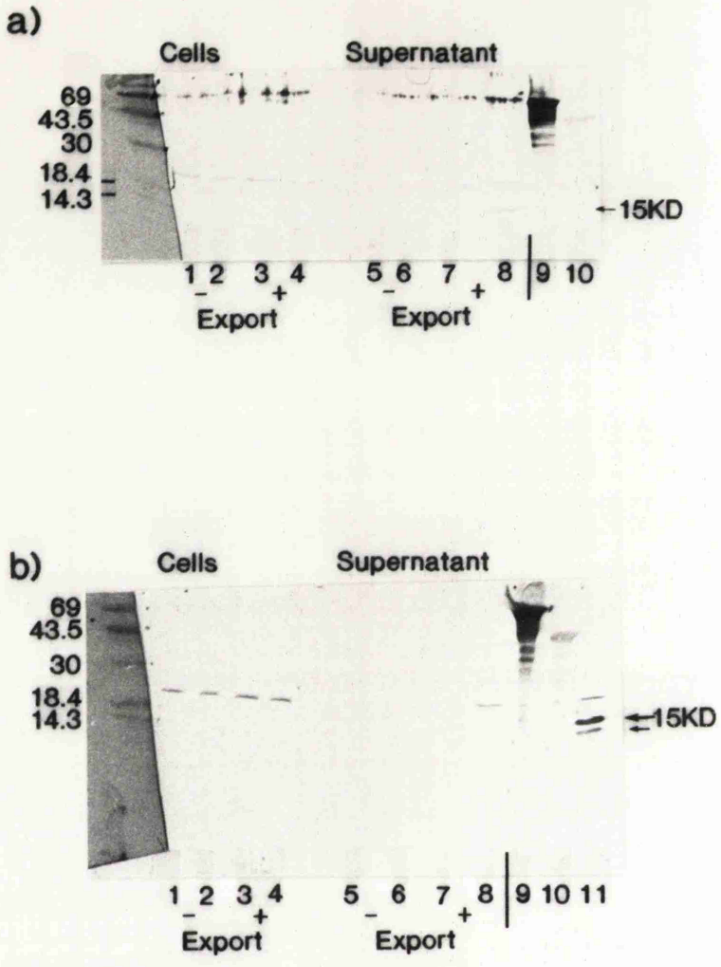
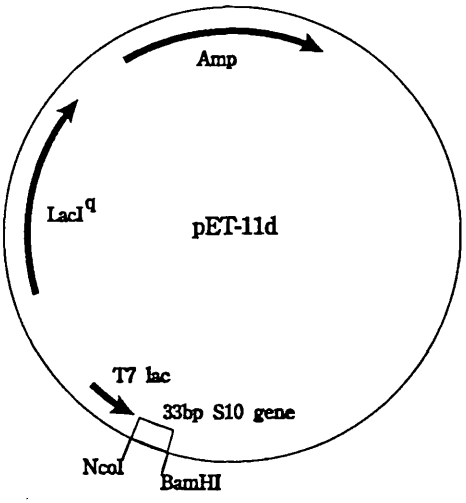
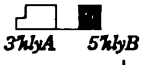

a)
Cells
Supernatant
69
43.5
30
18.4
14.3
15KD
1 2 3 4
5 6 7 8 9 10
Export
Export
b)
Cells
Supernatant
69
43.5
30
18.4
14.3
15KD
1 2 3 4
5 6 7 8 9 10 11
Export
Export

Figure 7:
Construction of plasmids with the expression of the C-terminal 4 and 15KD domains of HlyA under the control of T_7 RNA polymerase.
See text for details - Section 4.6.1
pET11d: Studier *et al.*, 1990
pPLEX-4KDHlyA and pPLEX-15KDHlyA see Figure 5.



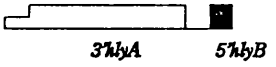
1)Cut BamHI and Fill in
2)Cut NcoI
isolate large fragment

Isolate 275bp NcoI/HpaI
Fragment from pPLEX-4KDHlyA
(see Fig 5)

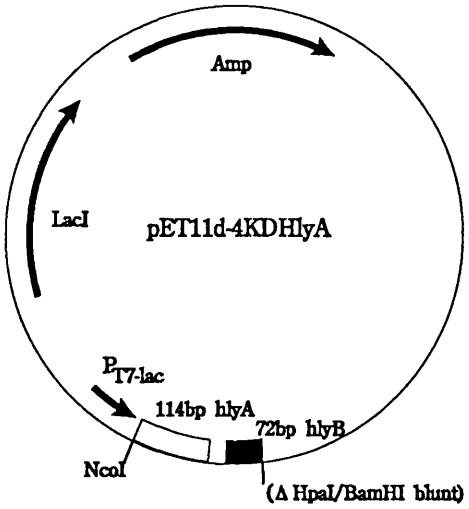


3)NcoI/HpaI
Fragment

Isolate 0.6Kb NcoI/KpnI(blunted)
Fragment from pPLEX-15KDHlyA
(see Fig 5)



4)NcoI/KpnI(blunt)
Fragment



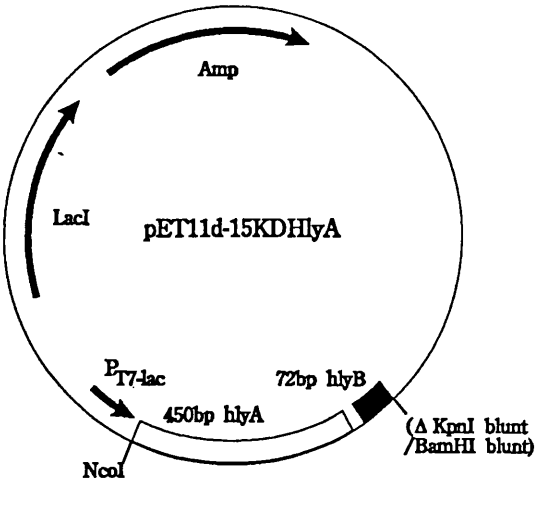


Figure 8:

SDS-PAGE gradient gel to try to resolve the 4 and 15KDHlyA peptides.

E. coli BL21 (DE3- prophage encoding inducible T7RNA polymerase) containing either pET11d-4KDHlyA or pET11d-15KDHlyA +/- pLG575 (*hlyB,D*) were grown in nutrient broth (with appropriate antibiotics) to an A^{450} of 0.7 before inducing with IPTG (0.5mM final concentration). RNA polymerase and thus the HlyA 4 and 15KD peptides should then be expressed. Samples were taken before and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were separated on a 10-18% SDS-acrylamide gel containing 0-10% sucrose (as described by Hashimoto *et al.*, 1983) and visualised by Coomassie blue staining.

pET11d-4KDHlyA +/- pLG575 (*hlyB,D*)

Cellular samples: lane 1 and 2, time = 0 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 3 and 4, time = 0 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 5 and 6, time = 0 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 7 and 8, time = 0 and 3 hours, respectively, from the culture **plus** pLG575 (*hlyB,D*).

pET11d-15KDHlyA +/- pLG575 (*hlyB,D*)

Cellular samples: lane 9 and 10, time = 0 and 3 hours, respectively from the culture **minus** pLG575 (*hlyB,D*) with lane 11 and 12, time = 0 and 3 hours respectively, from the culture **plus** pLG575 (*hlyB,D*).

Supernatant samples: lane 13 and 14, time = 0 and 3 hours, respectively from the cultures **minus** pLG575 (*hlyB,D*) with lane 15 and 16, time = 0 and 3 hours, respectively, from the culture **plus** pLG575 (*hlyB,D*). The profile evident in lane 14 represents some degree of cellular lysis or whole cell contamination.

Arrow marks the position of the 15KDHlyA peptide with molecular weight markers also shown.

Note the apparently specific release of two proteins of ~46 and 30KD into the supernatants after IPTG induction and independently of HlyB,D. These are probably prophage (DE3) encoded proteins expressed after induction of T7 RNA polymerase.

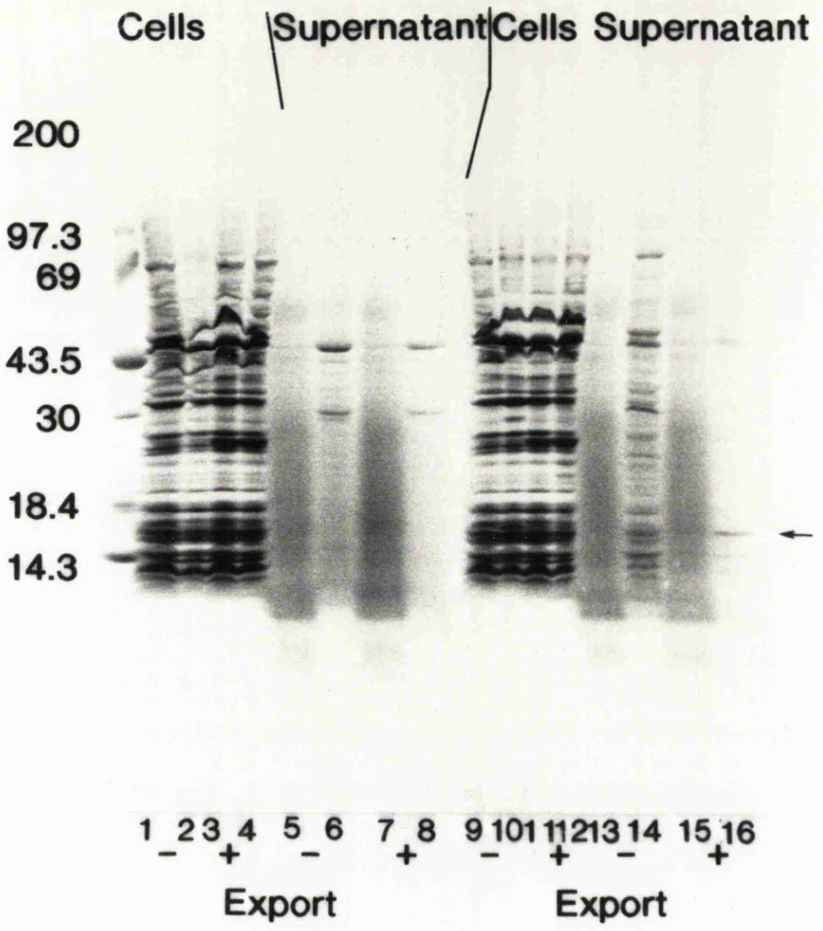
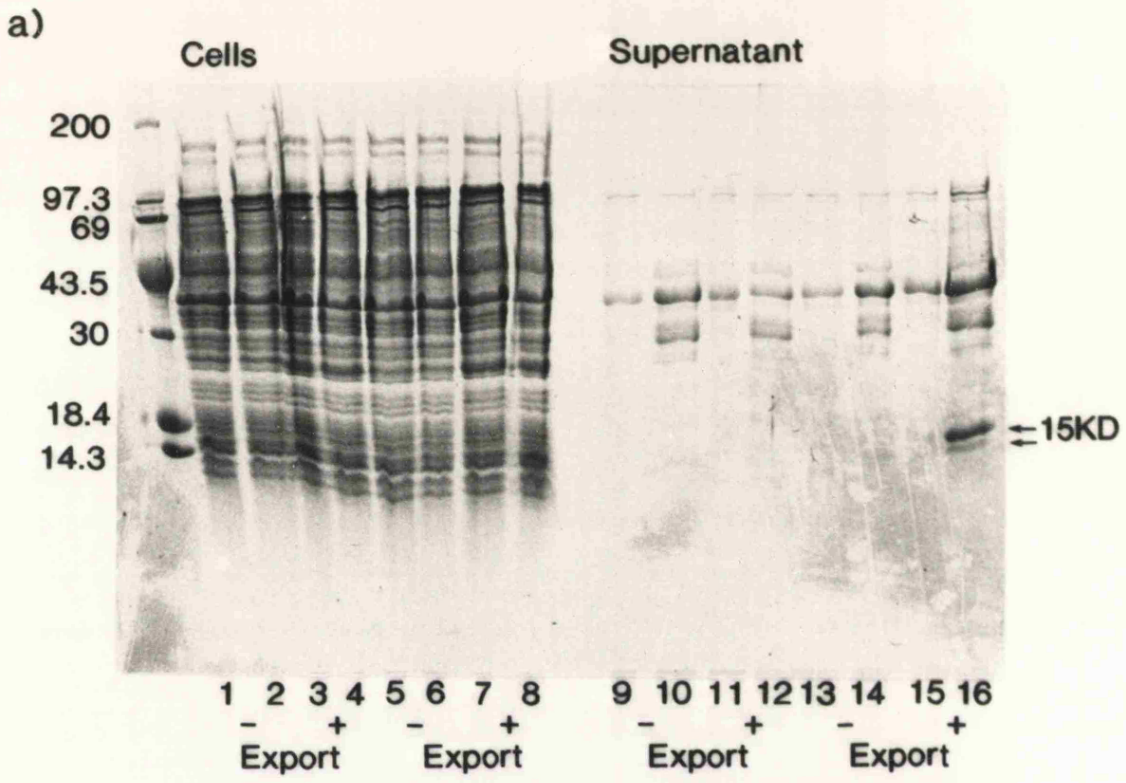

Cells
Supernatant
Cells
Supernatant
200
97.3
69
43.5
30
18.4
14.3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
− + − + − + − +
Export
Export

Figure 9:
Detectable secretion of the 15KD but not the 4KD HlyA peptide in the presence of HlyB,D.
E. coli BL21 (DE3- prophage encoding inducible T7 RNA polymerase) containing either pET11d-4KDHlyA or pET11d-15KDHlyA +/- pLG575 (*hlyB,D*) were grown in M9-minimal medium (with appropriate antibiotics) to an A^{450} of 1 before inducing with IPTG (0.4mM final concentration)T7 RNA polymerase and thus the HlyA 4 and 15KD peptide expression. Samples were taken before and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.1 unit of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were separated on a) a 20% SDS-acrylamide gel and visualised by Coomassie blue staining and b) on a 25% SDS-PAGE gel (Fling and Gregerson, 1986) and visualised by Western blot analysis using anti-23KDHlyA antibodies.

a) Cell samples:
Lane 1: pET11d-4KDHlyA **minus** pLG575, t = 0 hours
Lane 2: pET11d-4KDHlyA **minus** pLG575, t = 3 hours
Lane 3: pET11d-4KDHlyA **plus** pLG575, t = 0 hours
Lane 4: pET11d-4KDHlyA **plus** pLG575, t = 3 hours
Lane 5: pET11d-15KDHlyA **minus** pLG575, t = 0 hours
Lane 6: pET11d-15KDHlyA **minus** pLG575, t = 3 hours
Lane 7: pET11d-15KDHlyA **plus** pLG575, t = 0 hours
Lane 8: pET11d-15KDHlyA **plus** pLG575, t = 3 hours
Supernatant samples:
Lane 9: pET11d-4KDHlyA **minus** pLG575, t = 0 hours
Lane 10: pET11d-4KDHlyA **minus** pLG575, t = 3 hours
Lane 11: pET11d-4KDHlyA **plus** pLG575, t = 0 hours
Lane 12: pET11d-4KDHlyA **plus** pLG575, t = 3 hours
Lane 13: pET11d-15KDHlyA **minus** pLG575, t = 0 hours
Lane 14: pET11d-15KDHlyA **minus** pLG575, t = 3 hours
Lane 15: pET11d-15KDHlyA **plus** pLG575, t = 0 hours
Lane 16: pET11d-15KDHlyA **plus** pLG575, t = 3 hours
Arrows mark the position of the two forms of the 15KDHlyA peptide with high molecular weight markers also shown.
Note the apparently specific release of two proteins of ~46 and 30KD into the supernatants with this time the 46KD protein present even before IPTG induction (also see Figure 8 above).

b) Cell and supernatant sample loading are the same as above.
Positions of the molecular weight markers are shown with the question mark indicating the inability to identify both small molecular weight markers. Arrows indicating the position of the two forms of the secreted 15KDHlyA peptide.


a)
Cells
Supernatant
200
97.3
69
43.5
30
18.4
14.3
15KD
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
- + - + - + - +
Export Export Export Export

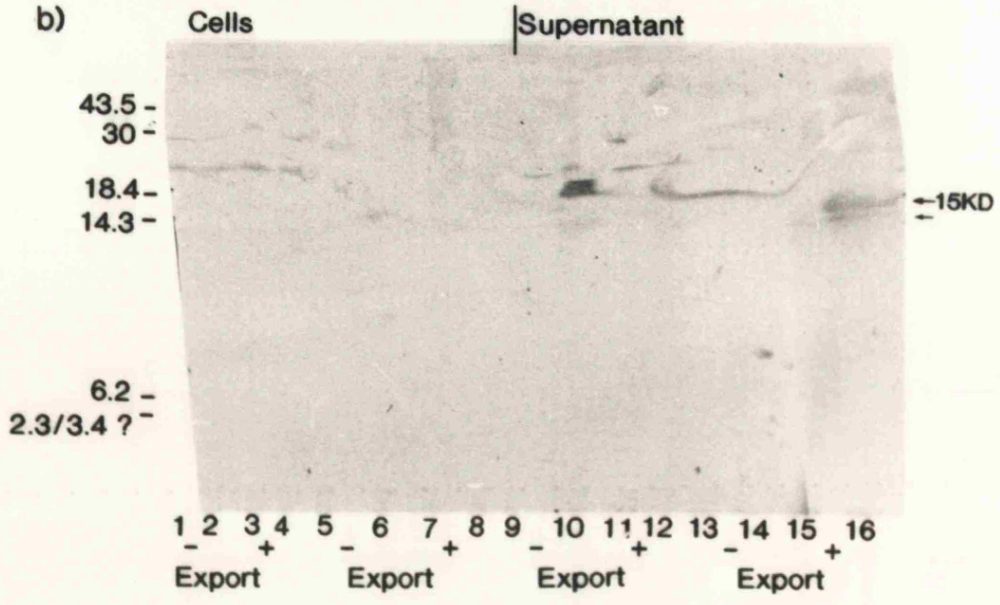
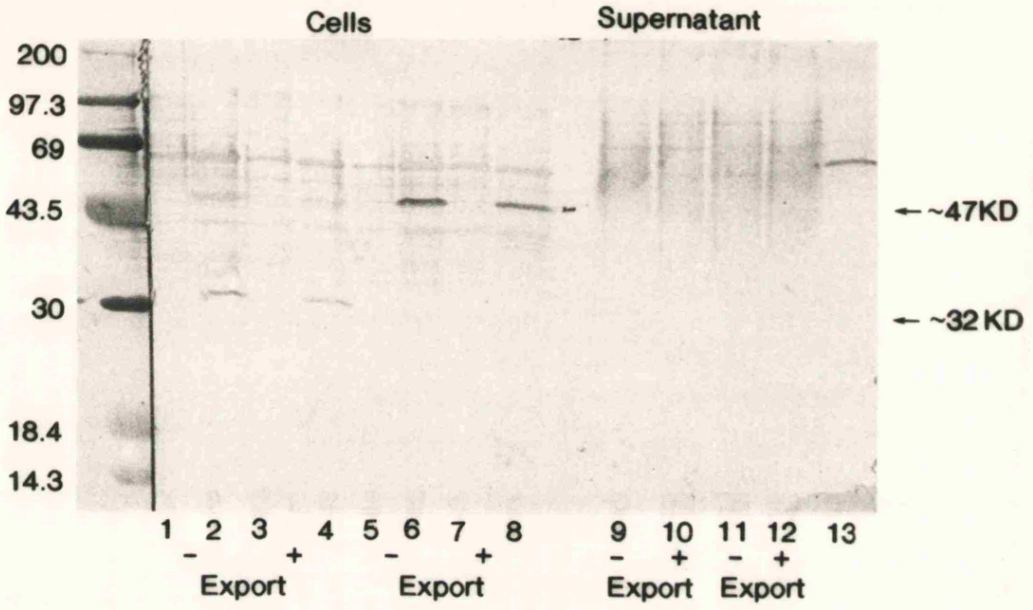

b)
Cells
Supernatant
43.5
30
18.4
14.3
15KD
6.2
2.3/3.4 ?
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
- + - + - + - +
Export Export Export Export

Figure 10:
Non-secretion of the 23KDHlyA peptide carrying C-terminal extensions:
E. coli JM101 (F',*lacI*q) containing either pLG618 (pUC23KD*hlyA*EcoRI) or pLG803 (pUC23KD*hlyA*EcoRI-*pcm*) +/- pLG575 (*hlyB,D*) were grown in nutrient broth (with appropriate antibiotics) to an A^{450} of 0.5 before IPTG induction (0.5mM final concentration) of expression. Cell samples were taken before and 2 hours post-induction with supernatant samples only obtained after 2 hours induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.1 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were separated on a 15% SDS-acrylamide gel, transferred to Nitrocellulose and probed by Western blot analysis using anti-HlyA antibodies.
Cellular samples:
Lane 1: pLG618 **minus** pLG575, t = 0 hours
Lane 2: pLG618 **minus** pLG575, t = 2 hours
Lane 3: pLG618 **plus** pLG575, t = 0 hours
Lane 4: pLG618 **plus** pLG575, t = 2 hours
Lane 5: pLG803 pLG575, t = 0 hours
Lane 6: pLG803 **minus** pLG575, t = 2 hours
Lane 7: pLG803 **plus** pLG575, t = 0 hours
Lane 8: pLG803 **plus** pLG575, t = 2 hours
Lane 13: pLG802 plus pLG575, t = 3 hours (see Chapter 3: Section 3.3.1)
Supernatant samples:
Lane 9: pLG618 **minus** pLG575, t = 2 hours
Lane 10: pLG618 **plus** pLG575, t = 2 hours
Lane 11: pLG803 **minus** pLG575, t = 2 hours
Lane 12: pLG803 **plus** pLG575, t = 2 hours

Arrows mark the position of the 32 and 47KD hybrid proteins with molecular weight markers also shown.


Cells
Supernatant
200
97.3
69
43.5
30
18.4
14.3
~47KD
~32 KD
1 2 3 4 5 6 7 8 9 10 11 12 13
- + - + - + - +
Export
Export
Export
Export

CHAPTER 5

Estimation of the relative level of secretion of the 23KD C-terminus of HlyA and Pcm-HlyA hybrids:

5.1 Introduction:

This chapter deals with the strategies undertaken to estimate the level of secretion of the various Pcm-HlyA hybrids, constructed in the previous two chapters, and that of the 23KD HlyA signal domain itself. The 23KD HlyA peptide does not accumulate intracellularly, presumably as a result of proteolysis, either in the presence or absence of export functions, consequently, it is not possible to estimate its absolute efficiency of secretion. Similarly, as the over-expressed Pcm-HlyA fusion proteins form insoluble inclusion bodies (Chapter 3; Section 3.2.1), it is also not possible to estimate the efficiency of secretion based upon ratios of secreted and non-secreted forms. However, it was hoped that comparison of the relative level of secretion of the various Pcm-HlyA hybrid proteins (see Chapters 3 and 4), with that of the 23KD peptide, might generate an insight into some of the structural factors which might affect the efficiency of secretion by the haemolysin export machinery. An indirect method of estimating the relative level of secretion would be to radioactively label a culture synthesising and secreting the polypeptide of interest, comparing the amount of radioactivity incorporated in the "secreted" polypeptide to that incorporated into the total cellular protein. This would give a relative value as a percentage of total labelled cellular protein. For this type of experiment, an intracellular protein, known to be present at a constant proportion of cellular protein, should be monitored to check the accuracy of the experiment. A good internal standard protein is EFTu, known to represent approximately 5.5% of total cellular protein throughout steady state growth in *E. coli* and *Salmonella typhimurium* growing in minimal medium (Pedersen *et al.*, 1978; Bremer and Dennis, 1987) and running as an almost homogeneous 44KD protein band on SDS-acrylamide gels (Boyd and Holland, 1979).

5.1.2 Radioactive ^{3}H-labelling of secreted proteins:

The radioactive isotope routinely used to label proteins is ^{35}S-methionine, but the 23KD C-terminal peptide of HlyA contains only a single internal methionine residue reducing the ability to detect this protein making it an unfavourable choice. In contrast, the C-terminal 23KD of HlyA, Pcm and EFTu molecules all contain leucine as approximately 8% of total residues. This is also the proportion reported for leucine as a percentage of total cellular proteins, making it an ideal residue for these quantification studies. ^{3}H is a very weak radioactive source (beta-emitter), and thus very susceptible to quenching, while the detection capabilities of a scintillation counter for ^{3}H-compounds varies between 10-60% depending on the machine used (Neame and Homewood,

1974). Therefore, it was decided that several million counts of incorporated activity would be required in order to obtain a detectable signal. 1μci of radioactivity emits 2.2 x 10^6 disintegrations per minute (dpm), with a minimum of 2.2 x 10^5 (10%) counts per minute (cpm) expected to be detected by the scintillation counter. In order to obtain an accurate, reproducible count it is recommended that each vial containing a radioactive sample should emit at least 10,000 cpm. Therefore, I decided to use 50μCi of ^{3}H-leucine (1.1 X 10^8 dpm) in the labelling experiments which should ensure the incorporation of sufficient counts for detection.

5.1.3 ^{3}H-leucine labelling of cultures secreting the 23KD peptide:

A pUC18 based plasmid, pLG612-1, expressing the 23KD of HlyA from an inducible *lac* promoter (R.Haigh this laboratory; see Figure 1) was transformed into the *E. coli* JM101 strain (carries the *lacI*q gene to control expression from the *lac* promoter) together with pLG575 (*hlyB,D*). All cultures for labelling were grown in M9-glucose minimal medium supplemented with all amino acids except leucine (Leucine Assay Mix - LAM; see Materials and Methods). An overnight culture was used to inoculate fresh media to an A^{450} of about 0.1 and grown at 37°C with shaking. At an A^{450} of 0.36 the culture was induced for 23KD HlyA peptide expression by the addition of IPTG (0.5mM final concentration; see Materials and Methods). After a further 30 minutes shaking, 10ml of culture was transferred to a new flask and 50μCi of ^{3}H-leucine was added together with 1μg/ml of cold leucine, in order to ensure constant uptake of ^{3}H-leucine throughout the labelling period. Only cold leucine was added to the remaining culture for monitoring cell growth. After 30 minutes, protein synthesis and labelling was stopped in both cultures by the addition of 250μg/ml of chloramphenicol and 100μg/ml of cold leucine. The cells were separated from the medium by spinning at 15K rpm for 10 minutes, and the proteins in the supernatant fraction were precipitated by the addition of TCA to a final concentration of 10% w/v as usual (see Materials and Methods). Various cell and medium samples were taken in duplicate for scintillation counting to give an indication of the level of incorporation of the ^{3}H-leucine into acid precipitated material, and counted using a Packard Scintillation counter (see Materials and Methods). The results obtained are given in Table 1.

5.1.4 Interpretation of results obtained from Scintillation counting:

Only 8.2 x 10^6 cellular counts per minute (Table 1a; B) and 2.2 x 10^6 supernatant cpms (Table 1a; D) were detected in the whole culture out of a total of 1.1 x 10^8 disintegration per minute (dpm) introduced into the experiment, suggesting that the scintillation counter used only detected about 10% of the radioactive disintegrations emitted. Of the 2.2 x 10^6 (Table 1a;D) counts detected in the culture supernatant only 0.13 x 10^6 (~9%) remained in the TCA precipitated pellet (Table 1a, E) indicating that the remaining counts represented unincorporated radioactivity (~20% of total introduced into experiment), indicating that the level of isotope was not limiting. This also suggested, that if this TCA precipitated material in the culture supernatant (0.13 X 10^6 cpm, Table 1a; E) consists only of the secreted 23KD peptide, that is, no significant levels of protein present due to cell lysis, then the 23KD protein constitutes approximately 1.7% of total detected labelled (Table 1a) protein of the cell.

5.1.5 Estimation of the 23KD peptide secreted as percentage of total cellular labelled protein:

Cell and supernatant samples equivalent to an A^{450} of 0.5 and 5 units, respectively, were separated on a 15% SDS-PAGE gel, Coomassie blue stained, dried and exposed to ^{3}H-sensitive X-ray film as described in Materials and Methods. A 10 day exposure of this gel is shown in Figure 2. It is evident that the 23KD protein is the major supernatant protein (lane 2), the other minor protein bands presumed to have emanated from a small percentage of unavoidable cell lysis, or from contamination of culture supernatant with whole cells. However, since there is only a minor intracellular band corresponding in size to the extracellular 23KD band, this implies that most, if not all, of this band is composed of truly secreted protein. Subsequent experiments have revealed that these extra supernatant bands are probably associated with the novel secretion of proteins under the culture conditions used (see below 5.3 and Figure 5).

Before excising the appropriate bands for scintillation counting to estimate relative levels of incorporation, it was first necessary to determine the efficiency of eluting and counting radioactivity from gel slices in liquid scintillant. This was necessary to allow comparison of counts incorporated in the secreted bands (estimated by eluting counts into liquid scintillant) to those in the total cellular lysate (estimated by solid phase counting of cellular sample immobilised on filters), as the characteristics of solid phase and liquid phase scintillation counting are

different. To do this, a cellular sample of a known radioactive content (estimated by solid phase scintillation counting), was separated by SDS-PAGE, Coomassie blue stained and dried. Then the whole lane was excised, rehydrated in 0.5ml water at room temperature for 10 minutes, 50ml NCS scintillation fluid added (see Material and Methods), vortexed and left overnight at 37°C. The sample was re-vortexed before two 5ml samples were taken and counted. Out of the 6.6 x 10^4 cpm estimated by solid phase counting to be loaded, only 4.5 x 10^4 cpm were detected. Thus the efficiency of elution and counting by liquid scintillant appeared to be ~70%.

5.1.6 Counts detected in the EFTu and the 23KD secreted bands:

The protein bands (see Figure 2) corresponding to the extracellular 23KD and intracellular EFTu (as an internal control) were excised and treated as above, except a tenth volume was used to elute the sample. The total radioactivity incorporated into the EFTu band during the experiment was 1.58 x 10^6cpm while the 23KD band, contained 0.19 x10^6 cpm, both corrected for the estimated 30% error encountered due to liquid phase counting (Table 1b). This suggested that the level of the 23KD protein secreted was equivalent to about 2.5% of the total labelled cellular protein (7.67 x10^6: Table 1a; lane A). However, the internal control, EFTu, band appeared to represent approximately 20% of the total cellular labelled protein instead of the expected 5.5% (see Section 5.1). The results by the two methods (1.7 and 2.5%) for secreted levels of the 23KD protein were in quite good agreement. Nevertheless, it appears that these values are over-estimates as the value obtained for the internal control (EFTu) is over-estimated by 2-3 fold from the documented value. One source of error might be due to different levels of quenching encountered with counting radioactivity emitted from cellular and TCA precipitated material or due to comparison of counts estimated by solid phase versus liquid phase scintillation counting. The high values of EFTu observed could also be due to other experimental errors that we are not aware of, including non-steady state labelling. However, if the value for the secreted 23KD protein is related to EFTu as approximately 5.5% of total protein under these conditions, we now obtain a value of between 0.47 -0.68%.

In view of the unsatisfactory results obtained above a duplicate experiment was carried out, except that 100μg/ml of cold leucine was added at the start of the labelling period to ensure a slow, constant uptake of the labelled residue. This culture, induced for expression of the

23KD peptide at an A^{450} equal to 0·32, gave results consistent with the previous experiment. Thus, the secreted 23KD polypeptide band when excised and counted was calculated to represent 2.4% of total cellular protein, although in this case EFTu represented 10.6% of the total labelled cellular protein (data not shown). EFTu, in the 30 minute labelling window, still therefore appeared to constitute substantially more than the 5 to 6% of the total labelled. In the above two experiments the eluted secreted 23KD peptide apparently represented approximately 2.4% of the total labelled proteins under the conditions of the experiment described. Nevertheless, these values must be considered high and to be viewed with some caution given the high and variable EFTu values obtained. This point will be returned to in the following section and the discussion.

5.2 Quantification of the level of secretion of the 60KD Pcm-23KDHlyA hybrid:

The 60KD Pcm-23KDHlyA fusion protein is expressed from the dual origin based vector pLG802 (see Chapter 3; Section 3.3.1). The protein consists of 99% of the Pcm molecule fused to the C-terminal 23KD of HlyA. The plasmid encoding this hybrid was transformed, as before, into *E. coli* JM101 together with pLG575, which encodes the *hlyB,D* export genes. Cultures were grown as described for the 23KD labelling experiments, but at 30°C for repression of plasmid copy number amplification. In this experiment the culture was heat induced at an A^{450} of 0.62, as described in Materials and Methods, for copy number amplification and thus high level expression of the hybrid protein from the constitutive *trp* promoter. Growth was continued for 30 minutes before ^{3}H-leucine was added to 10ml of culture, and after a further 30 minutes labelling, protein synthesis was arrested, as described earlier (also see Materials and Methods). Again cell and supernatant samples were taken for SDS-PAGE analysis and for scintillation counting. Comparison of the Coomassie blue stained and ^{3}H-labelled protein profiles (Figure 3; a versus b) identified several Coomassie stained bands from the culture supernatant (Figure 3a; Lane 3). However, only one of these bands appeared to emit a significant radioactive signal, the 60KD Pcm-23KDHlyA hybrid (Figure 3b; Lane 2). This indicated that during the labelling period this protein was actively synthesised and secreted to the medium and not released by lysis/whole cell contamination. The other bands appears to be associated with the unexpected release, when grown in minimal medium, of an *E. coli* protein band of about 52KD in size (see Figure 3a, Lane 3 and Section 5.3 below).

The cell samples did not show the presence of significant levels of the intracellular 60KD hybrid (Figure 3a, Lanes 1-2), which usually accumulates under conditions of high level expression (see Chapter 3), suggesting that during the labelling period the majority of this synthesised protein was secreted and therefore the export pathway may not have been saturated. Surprisingly, in this experiment only 1-2% of the total ^{3}H-leucine introduced was incorporated into cellular material (data not shown), resulting in reduced counts for analysis and thus for accuracy. Comparison of the ratio of counts eluted from the secreted, Pcm-23KDHlyA 60KD, band with that of the total cellular counts (estimated by solid phase counting of cellular samples immobolised on filters) in equivalent sample loadings indicated that in this case the hybrid represented approximately 0.14% of the total labelled cellular protein. As in the previous experiment the eluted EFTu constituted approximately 9% of the total cellular labelled protein in these calculations. Again therefore this is an unexpectedly high value for the intracellular EFTu protein and might indicate an overestimation of the secreted hybrid. If we take EFTu to represent 5.5% of cellular protein and adjust the 60KD Pcm-HlyA hybrid figure accordingly then a value of 0.08% of total cellular protein is obtained for this secreted peptide.

In case these apparently high values of EFTu were due to some experimental error in counting radioactivity I chose an alternative method of measuring both EFTu and the secreted hybrid by using Laser densitometry to scan the relevant exposed X-ray films. This method produced different results from those obtained by scintillation counting, in that now, the values for EFTu in all the labelling experiments experiments described so far were 9.9%, 5.5% and 4.9% compared to 20, 10.6 and 9% respectively. Thus, in two of the three cases the EFTu value now corresponds closely to that reported previously, implying that in the labelling experiments an incorrect estimation of the efficiency of protein elution or scintillant counting of ^{3}H-labelled material was obtained. The higher than expected EFTu value of 9.5% obtained in the first experiment still remains to be explained. Interestingly, IPTG treatment of comparable cultures to those used in these experiments appears to induce to variable (high) levels a protein band of ~46KD (data not shown) in total cell lysates which could account for this high EFTu value if those two proteins are not well separated on the scanned 11% SDS-acrylamide gel run.

The results obtained from the scanning of the radioactive protein profiles now allow new calculations of the level of secreted 23KD. This was done relative to EFTu assuming the latter to be 5.5% of total cellular protein. By this measurement the secreted 23KD peptide and the 60KD Pcm-HlyA hybrid now constitute approximately 0.49% and 0.07%, respectively, of total cellular protein. These values agree well with those obtained from the scintillation counting when recalculated and related to EFTu assuming that EFTu represents 5.5% of total cellular protein (0.47-0.68% and 0.08% for the 23KD and 60KD peptides respectively; see above and Section 5.1.6). Taking into account the difference in protein size (same percentage leu content) the Pcm-HlyA23KD hybrid was secreted at only 5.5% the level of the 23KD alone.

Although the results obtained by laser scanning of autoradiographs appeared more reliable and in particular as both cell and supernatant protein bands were quantified in the same manner, the absolute values obtained for secretion seemed to be unexpectedly low from previous, visual, estimation of the level of secretion of the 23KD HlyA peptide and Pcm-HlyA hybrids, for example in rich nutrient media (data not shown; see below). This apparent difference, if real, might be due to labelling before optimal levels of secretion had been obtained, or differences in the induction conditions used, heat shock (for the Pcm-HlyA hybrids) compared to IPTG induction (for the 23KDHlyA peptide) or the composition of the growth medium. To investigate possible effects of some of these conditions on secretion, I compared directly the level of the 23KD polypeptide in the medium in rich nutrient versus minimal media and the effect of simulated heat induction conditions.

5.3 Secretion of the 23KD HlyA peptide in minimal versus rich media:

To simplify this experiment the DNA fragment coding for the C-terminal 23KD of HlyA and HlyB,D were cloned onto the same plasmid as shown in Figure 4. This pUC based plasmid was designated pLG614, the *hly* genes having been cloned downstream of the *lac* promoter. This plasmid was transformed into the F' *lacI*^{q} containing strain NM522 and plated out on M9-glucose minimal or rich nutrient agar plates. Overnight cultures were grown in either M9-glucose minimal medium or rich nutrient broth and used to inoculate fresh cultures to an A^{450} of 0.2. Plasmid gene expression was subsequently induced by addition of IPTG (0.5mM final concentration) at an A^{450} of 0.7 (see Materials and Methods) and growth maintained at 37^{o}C. Both cellular and supernatant samples were taken

at the time of induction and at 1,2 and 3 hours after induction. The samples were prepared as usual (see Materials and Methods) and a 15% SDS acrylamide gel loaded with 0.5 cell and 5 A^{450} equivalent units of supernatant sample, respectively. The proteins were separated by SDS-PAGE and the resulting Coomassie stained gel is shown in Figure 5. Figure 5 clearly demonstrates the specific secretion of the 23KD polypeptide, following IPTG induction of expression, in both media (Figure 5; lanes 9-12 and 13-16) and as expected there is no visible accumulation of the 23KD hybrid within the cellular samples (Lane 1-4 and 5-8). Visual comparison suggests that approximately 4-5 times more 23KD polypeptide was ultimately secreted in rich media compared with M9-glucose minimal medium. In M9-glucose minimal medium the amount of the 23KD protein in the supernatant fraction appears to be constant, after induction, throughout the course of the experiment (equal A^{450} samples loaded). This indicated that the 23KD peptide, under these conditions, was secreted at a constant rate throughout the experiment assuming that the protein is stable in the medium. In contrast, in nutrient broth the rate of secretion appeared to increase during the experiment, accumulating more 23KD peptide per A^{450} unit of cells with time. The molecular basis for this difference or whether it is due to increased synthesis and/or secretion of the peptide are not known.

Comparison of the cellular profiles in Figure 5, showed the presence of a highly stained Coomassie stained band of approximately 52KD, visible only during growth in the minimal medium culture (Figure 5; lanes 1-4 versus 5-8). More, surprisingly, this presumed protein band was found in substantial amounts in the medium of the minimal media grown cells (Figure 5; lanes 9-12). However, the synthesis of this protein was not induced with IPTG and appeared to be released with a subset of cellular proteins. Nevertheless, the high levels of this band appear to represent another novel polypeptide released from *E. coli*, under these specific growth conditions. This protein is also secreted from the strain NM522 lacking any plasmid vectors, except F', when grown in either M9-glucose or glycerol minimal media (data not shown). The identity of this material however remains unclear.

5.4 Effect of growth temperature on the secretion of the 23KD HlyA peptide:

In this experiment the 23KD HlyA peptide was expressed from pLG612-1 under the control of the *lac* promoter (see Figure 1), with the genes for *hlyB,D* present on a second plasmid pLG575 (*hlyB,D*). Both of these

plasmids were transformed into the F' *lacI*^{q} containing *E. coli* strain JM101 and maintained at 30°C. An overnight culture, grown at 30°C, was used to inoculate 200ml of fresh nutrient broth to an A^{450} of 0.06 and grown at 30°C to an A^{450} of 0.15. At this point 60ml were transferred into a separate culture flask and grown at 37°C. Both the 30°C and 37°C cultures were induced for 23KD expression with IPTG (0.5mM final concentration) at an A^{450} equal to 0.74. Upon IPTG induction of the 30°C culture, half was transferred to another culture flask and placed in the 37°C shaker. Cell and supernatant samples were taken from each culture before, 1.5 and 3 hours after the addition of IPTG. The samples were prepared for SDS-PAGE as before (see Materials and Methods) and 0.5 and 5 OD of A^{450} equivalent units of cell and supernatant samples, respectively, separated on a 15% gel. The resulting Coomassie blue stained gel is shown in Figure 6. The gel shows no detectable difference in the profiles of the total cell protein grown at different temperatures, noting the absence of intracellular accumulation of the 23DK peptide (Figure 6: lanes 1-8). However, there is a striking difference in the relative amounts of the 23KD peptide in the supernatant samples derived from cultures grown at the different temperatures (Figure 6; lanes 9-16). Repression of 23KD synthesis and thus non-secretion of this peptide until after IPTG induction of expression of the controlling promoter is evident in both 30 and 37°C grown cultures (Lanes 9 and 12). However, induction of expression with IPTG (0.5mM final concentration) of the 23KD HlyA peptide in the cells grown at 30°C, resulted in an estimated 5-fold reduction (by visual comparison) in the quantity ultimately secreted compared to the identical culture grown at 37°C (lanes 13,14 versus 10,11). Moreover, transferring the induced culture grown at 30°C to 37°C, resulted in a gradual recovery of the level of secretion to approximately that of the culture grown throughout the experiment at 37°C (lanes 15,16 versus 13,14 compared to 10,11). These results clearly indicated that the secretion process is temperature dependent. This might reflect the evolution of this toxin secretion system in pathogenic bacteria at body temperature (37°C) at a level other than transcription as these subcloned *hly* genes are apparently deleted for the major promoter region. This result is also consistent with the participation of induced heat shock proteins in facilitating the translocation process as has been shown to be the case in other export systems (see Chapter 1: Section 1.4.5). It is worth noting that rate of secretion at 30°C appears to remain constant while at 37°C the level of secretion of the 23KD protein appeared to increase with time throughout the experiment.

5.5 Laser scanning of SDS-PAGE Coomassie blue stained profiles to determine the relative levels of secretion of the 23KD and Pcm-HlyA hybrid proteins:

5.5.1 Background:

Due to the observed difference in secreted protein levels under different conditions it was decided to try to quantify the relative level of secretion for cultures grown under the standard laboratory conditions of growth in nutrient broth at 37°C. In this case radioactive labelling was not convenient and the alternative method chosen was to again use laser densitometry (LKB Ultrascanner XL) to scan Coomassie blue stained SDS-PAGE profiles of the relevant cell and supernatant samples. The relative level of secretion of the various peptides could then be expressed as a percentage of total cellular protein, by comparing the secreted band to an intracellular protein band whose value as a percentage of cellular protein can be obtained directly from the laser densitometer data output. This approach was required as the output from the laser scanning when applied to a total protein gel profile does not give an accumulative relative absorbance value for all the cellular bands detected.

5.5.2 Estimation of the relative amount of secreted 23KD HlyA peptide expressed from a *lac* inducible promoter by laser scanning densitometry:

Secretion of the 23KD peptide was estimated when expressed from the *lac* promoter on the pUC based plasmids pLG614 or pLG612-1 (see below and Figures 1 and 4) in the *E. coli* strain NM522 (F' *lacI*q). The two plasmid systems differ in the presence of the haemolysin export genes, *hlyB,D*, provided on the same plasmid (pLG614) or in the case of pLG612-1 on a separate plasmid, pLG575 (*hlyB,D*). Cultures, grown in nutrient broth at 37°C, were initiated at an A^{450} of about 0.1 and expression of the 23KD HlyA peptide induced with IPTG (0.5mM final concentration) at an A^{450} of approximately 0.45. Cell and supernatant samples were taken prior to and 3 hours post-induction and samples prepared as usual (see Materials and Methods). O.5 and 1 equivalent A^{450} unit of cell and supernatant sample, respectively, were separated by SDS-PAGE. The resulting Coomassie blue stained gel, shown in Figure 7 was scanned using the LKB ultrascanning densitometer. The resulting data output expressed a particular band as a relative absorbance value and as a percentage of the total bands detected in that lane. The levels of the secreted 23KD band could then be expressed relative to an intracellular band, EFTu chosen as it constitutes the major intracellular band and would also serve to monitor the accuracy of the calculations as it is

known to represent about 5.5% of total cellular protein. The results obtained are summarised in Table 2a, Lines 1 and 2. From the table it is evident that EFTu was measured at roughly the expected 5.5% of total cellular protein suggesting that this method of quantification was more reliable than the labelling experiments. Comparison of the relative level of 23KD peptide secretion shows that it represents about 1.8% of total cellular protein when the C-terminal 23KD of HlyA and HlyB,D are expressed from the same plasmid as compared to 2.95% when expressed on different plasmids. Interestingly, the level of the secreted 23KD protein measured in these experiments was approximately 6 fold higher than that obtained by scanning the autoradiographs from the labelling experiments (0.49% see Section 5.2) using identical constructs. This appears to confirm the experiments, described above, where the level of the secreted 23KD HlyA peptide was visually compared in M-9 minimal media versus rich nutrient broth, revealing an approximate 4-5 fold difference in staining (see Section 5.3).

It is also worth noting in Table 2a that the amount of 23KD HlyA peptide secreted actually fell rather that increased, when the export proteins were expressed from the same plasmid. This reduction might reflect some form of restored regulation of *hlyB* and or *hlyD* due to their now normal position downstream of *hlyA* and the putative rho independent transcriptional terminator between *hlyA* and *hlyB* (see Chapter 1: Section 1.6.4). This extra regulation might be lost when *hlyB,D* are subcloned on to another plasmid, even with the putative rho independent transcriptional terminator present. However, no further studies were carried out to confirm these effects, which might equally reflect differences due to copy number, stability and or compatibility interactions of the different plasmids. In this respect it is important to note the extensive duplication of the *hlyB* gene in pLG614 (see Figure 4) enabling the export genes to be readily deleted if the secretion system perturbs the cells.

5.5.3 Measurement of the relative levels of secretion of various Pcm-HlyA chimeric proteins by laser scanning densitometry:

All the Pcm-HlyA hybrid proteins are expressed from constitutive trp promoters on the copy number amplifiable, dual origin based vector, derived from pMG168 (see Chapter 3 and 4). These plasmids, pLG800-1, pLG801, and pLG802 carry approximately 40, 70 and 99% of Pcm fused to the C-terminal 23KD of HlyA, respectively. These plasmids and in

addition pLG806 (40% of Pcm fused to the C-terminal 12KD of HlyA) and pLG807 (99% of Pcm fused to the C-terminal 4KD of HlyA; see Table 2b) were transformed into *E.coli* SE5000 containing pLG575 (*hlyB,D*). Cultures were grown in Luria broth at 30°C to an A^{450} of approximately 0.78, before heat induction for plasmid copy number amplification, and thus high level expression as usual (see Materials and Methods). Cell and supernatant samples were again taken before heat induction and 3 hours post-induction and the samples separated by SDS-PAGE (see Figure 8), before laser scanning of the Coomassie blue stained profiles. The results are summarised in Table 2a. Encouragingly, the EFTu band continued to represent under these conditions approximately 5.5% of total cellular protein. Table 2a, lines 3-7 show the relative level of secretion of the various Pcm-HlyA chimeric proteins as a percentage of the internal accumulated hybrid band, EFTu and total cellular protein. These figures are not directly comparable due to the difference in the sizes of the various hybrids, consequently Table 2a also shows the relative level of each hybrid secreted as a percentage of the smallest Pcm-23KDHlyA hybrid (encoded by pLG800-1) or the 23KD peptide itself. These results are calculated taking into account the size difference, and making the assumption that all the hybrid bands stain with Coomassie blue to the same degree. Taking the secretion value of the smallest Pcm-23KDHlyA hybrid as 100%, it appears that when the portion of the Pcm molecule fused to the 23KD of HlyA increases from 40% to approximately 70% and 99%, the level of secreted hybrid falls by 25% and 50%, respectively. More dramatically, if the 12KD, C-terminal portion of HlyA is fused to the first 40% of Pcm then the amount of hybrid secreted is reduced to only 30% of that when it is fused to the 23KD peptide in an otherwise identical construct. This result suggests that the additional residues present in the 23KD but not in the 12KD peptide contribute, in some manner, to the efficiency of secretion. Comparison of the relative level of secretion of the smallest Pcm(40%)-23KDHlyA fusion compared to the 23KD peptide itself demonstrated that under the conditions used, noting the use of pUC and dual origin based vectors for expression respectively, that this fusion was only secreted at approximately 16% of that of the 23KD peptide. This suggested that the presence of this foreign peptide fused to the C-terminal 23KD of HlyA dramatically reduces the secretion process. This might be due presumably to some inhibiting sequences or structures, but not size per se, since the 23KD signal peptide is designed to secrete the 84KD N-terminal toxin domain. From these experiments it was not, however, possible to determine whether such "inhibitory" sequences

affected secretion directly or whether this was due to the insoluble nature of the intracellular hybrid limiting its availability for translocation.

5.6 Discussion:

This chapter described the attempts made to estimate the level of secretion of the C-terminal peptides of HlyA, when expressed independently or fused to portions of the Pcm molecule. These results also generated an insight into some of the factors that alter the efficiency of the secretion process. It is important to emphasise when attempting to measure secretion efficiencies that the C-terminus of HlyA is intrinsically unstable within the cells, while over-expression of the Pcm-HlyA fusions led to the rapid incorporation into insoluble inclusion bodies, making it impossible to estimate the efficiency of secretion by comparing intracellular to extracellular levels. As an alternative we attempted initially to use radioactive labelling experiments to obtain a relative secretion value, by comparing directly the amount of label incorporated in the secreted protein to that of total cellular protein, monitoring EFTu (the most abundantly expressed *E. coli* protein under normal conditions) as an internal control. The evaluation of ^{3}H-leucine radiolabelling experiment was a relatively complex procedure, giving variable and in some cases unexpectedly high values for the intracellular EFTu protein. This may be due to either errors in estimating the efficiency of elution of bands from gels (or perhaps differences in eluting gel slices versus total gel lanes) or errors involved in detecting weak ^{3}H counts very susceptible to quenching. Errors may also have been compounded by the requirement to compare counts detected after elution from a gel band with total cellular counts (estimated by counting cellular samples immobilised on filters), presumably involving different quenching factors (see Section 5.1.6).

In an attempt to eliminate some of these variables the radioactive profiles (on X-ray autoradiographic film) were scanned using a laser densitometer and the relative values of EFTu and secreted bands expressed as a percentage of the total cellular labelled protein profile. Indeed, in 2 of the 3 experiments using this method, the value for EFTu was close to that expected from the literature values, with the secreted 23KD HlyA and a Pcm hybrid now representing 0.49 and 0.07% of total cellular labelled protein respectively. The laser densitometry method of estimation appeared, to be more reliable as the same method is used to determine the relative level of protein bands in both the cell and supernatant samples, avoiding the comparison of liquid phase eluted counts from gel

slices to solid phase scintillation counting of total protein. These values compare with those obtained from the scintillation counting experiments when the values were adjusted taking EFTu to represent 5.5% of total cellular protein and not the estimated elevated value (See Section 5.2).

The level of 23KD protein secretion detected in these experiments, surprisingly, was low compared with many previous studies and this suggested that the conditions used in the labelling method were themselves affecting the secretion process. Indeed, this was borne out by subsequently studying the effects of both growth in rich versus poor media and at low versus higher temperature (see below), conditions used in the labelling experiments. Nevertheless, the results obtained from scanning the autoradiographic profiles suggested that the largest Pcm-23KDHlyA hybrid (60KD) was secreted at about 6% the level of the 23KD peptide itself (taking into account molecular weight differences). Indeed, a low value was confirmed in further quantification experiments (see below), involving scanning laser densitometry of stained gel profiles. This indicates that the conditions used for the labelling experiments (for example minimal medium) reduced the absolute amount of all the HlyA secreted proteins but not the relative levels of the different secreted proteins.

Following these observations of low level secretion during radiolabelling the analysis of the effect of cultural conditions revealed substantial changes in secretion levels. Thus, an approximately 4-5 fold increase in the secreted 23KD HlyA was obtained when the cells were grown in rich medium (see Sections 5.3 and 5.5.2). The rates of secretion in the different media also appeared to differ with that in minimal media being apparently constant while growth in rich broth leads to the progressive increase in the rate of secretion of HlyA polypeptides during the growth phase. The basis of these differences are unknown, though possible feedback regulation at the major *hly* promoter can be ruled out as it is not present in these constructs. However, it is possible that there is some "sensory" regulation acting at the level of expression of the export proteins, either via the rho independent terminator present upstream of *hlyB,D* or on the *hlyD* promoter reported to exist at the 3' end of the *hlyB* gene. Such sensory regulation via environmental signals in these experiments appears to exist for the wild-type *hly* determinant (M. Blight, 1990). A surprising observation was the specific induction and secretion of an apparently novel 52KD protein band when grown in M9-glucose minimal medium. This band appears to represent another rare, hitherto

not observed unpublished example of a protein released from *E. coli*. Indeed, the release of two other novel proteins (migrating with indicated sizes of ~30 and 46KD) was also observed during these studies from an *E. coli* B strain carrying the prophage DE3 (see Chapter 4; Section 4.6.3)

The growth temperature also appeared to have a marked effect on the rate of secretion of the 23KD HlyA polypeptide, with secretion apparently at a constant rate at 30°C, but increasing substantially during the growth phase at 37°C. Similarly, transferring cultures from 30°C to 37°C resulted in the marked increase in the quantity of protein secreted. This shows that temperature is a significant factor in the haemolysin secretion system, and may be a result of the evolution of this system at body temperature, 37°C. If we speculate, for example, that HlyB has an unfoldase activity then perhaps less energy is required at higher temperature to unfold molecules prior to secretion. Other possibilities are differential effects of DNA supercoiling on transcription (see Pruss and Drlica, 1989 for mini-review) or the presence of induced heat shock proteins promoting more efficient translocation as has been detected in other export systems (see Chapter 1: Section 1.4.5). Again, it is also possible that there is some form of regulation acting on the *hlyD* promoter or decreasing transcription through the rho independent terminator as speculated previously.

As growth in minimal media clearly leads to reduce levels of secretion of HlyA derivatives another method of quantifying secretion in rich broth was required. Thus, Coomassie blue stained protein gel profiles were scanned by laser densitometry making the assumption that all proteins absorb the stain to the same extent. With this method EFTu was calculated to be ~5.5% of total cellular protein while the secreted 23KD peptide (expressed from the *lac* promoter on a pUC based plasmid) represented ~3% of total cellular protein, that is about 6-fold higher than levels in minimal media. The observed reduction in secretion levels from cells grown in minimal media appears at least in part to be due to the apparently constant rate of secretion observed in this medium, whereas the rate of secretion appears to increase with time during the growth phase, of nutrient broth grown cells as reported in Section 5.3.

Analysis of the relative level of secretion obtained by scanning stained bands of the fusion proteins consisting of portions of the Pcm molecule fused to C-terminal segments of HlyA provided some important results. Firstly, even with the smallest Pcm-23KDHlyA fusion protein the level of

secretion was at least 6-fold less than with the 23KD alone. Moreover, increasing the amount of the Pcm molecule fused to the 23KD HlyA signal peptide appeared to reduce the relative level of secretion by about two fold (taking into account differences in molecular weight). This, effect might be due to increased occlusion of the signal peptide by the increasing tendency of the passenger peptide to fold into its native structure. Similarly, the increased ability of the larger Pcm fragments to form a more stable folded structure might lead to more rapid sequestration into inclusion bodies reducing the proportion of hybrid available to the secretion machinery. Finally, increasing folding and or stability of the Pcm moiety might result in features hindering its actual translocation through the export "channel". Interestingly, a similar effect of size of the passenger protein has been obtained by Richard Haigh in this laboratory (Kenny *et al.*, Submitted). In this case, increasing the proportion of the *E. coli* cytoplasmic protein, chloramphenicol transacetylase (CAT), fused to the 23KD peptide of HlyA dramatically reduced the efficiency of secretion. Neither of these results can be put down to size limitations *per se* since the wild-type toxin is much larger than either of these families of secreted hybrids. Moreover, in the case of the CAT or CAT-HlyA fusions no insoluble aggregates appear to form within the cells. We therefore favour the possibility that increasing the size of the passenger polypeptide leads to its increased ability to undergo some degree of folding, interfering with secretion either at the recognition step (via the targeting sequence of HlyA) or during the translocation step itself.

A dramatic 70% reduction in secretion levels was also observed when the N-terminal 40% of the Pcm molecule, was fused to the C-terminal 12KD of the HlyA signal peptide, rather than the 23KD fragment. The reduced distance between passenger and the targeting region in the smaller hybrid might lead to greater occlusion of the signal sequence or simply interfere with the formation of specific secondary/tertiary structures required for efficient recognition by the export proteins. Such results again appear to indicate that the *hly* secretion system is critically dependent upon the overall structure (folding ?) of the polypeptide to which the signal motif is attached, with other sequences possibly required to allow the targeting domain to function independently of upstream domains.

As indicated above all the Pcm fusion proteins were poorly secreted when compared to the 23KD HlyA protein alone. This might be due to a number of factors, as I have argued, occlusion of the signal domain or

sequestration of such proteins into insoluble aggregates. However, when comparing the levels of secretion of the 23KD peptide with the Pcm-HlyA hybrids it must also be noted that for technical reasons the hybrids and the 23KD protein were expressed from different plasmids and promoters making absolute conclusions difficult.

In conclusion the levels of secretion of the Pcm polypeptides achieved via the HlyA-system were far in excess of those obtained with Pcm fused to an N-terminal signal sequence (Little *et al*, 1989; see Chapter 3; Section 3.6). Thus, although the present results point to the importance in secretion efficiency of the tertiary structure of the protein attached to the HlyA or N-terminal targeting signal, the results also indicate some important, as yet undefined, differences in the basic mechanism of the HlyA and N-terminal signal transport pathways.

Figure 1:

Structure of pLG612-1 (pUC-23KD*hlyA*).

pLG612-1 is a pUC18 based vector (pUC18; Norrander et al., 1983), constructed by R. Haigh this laboratory

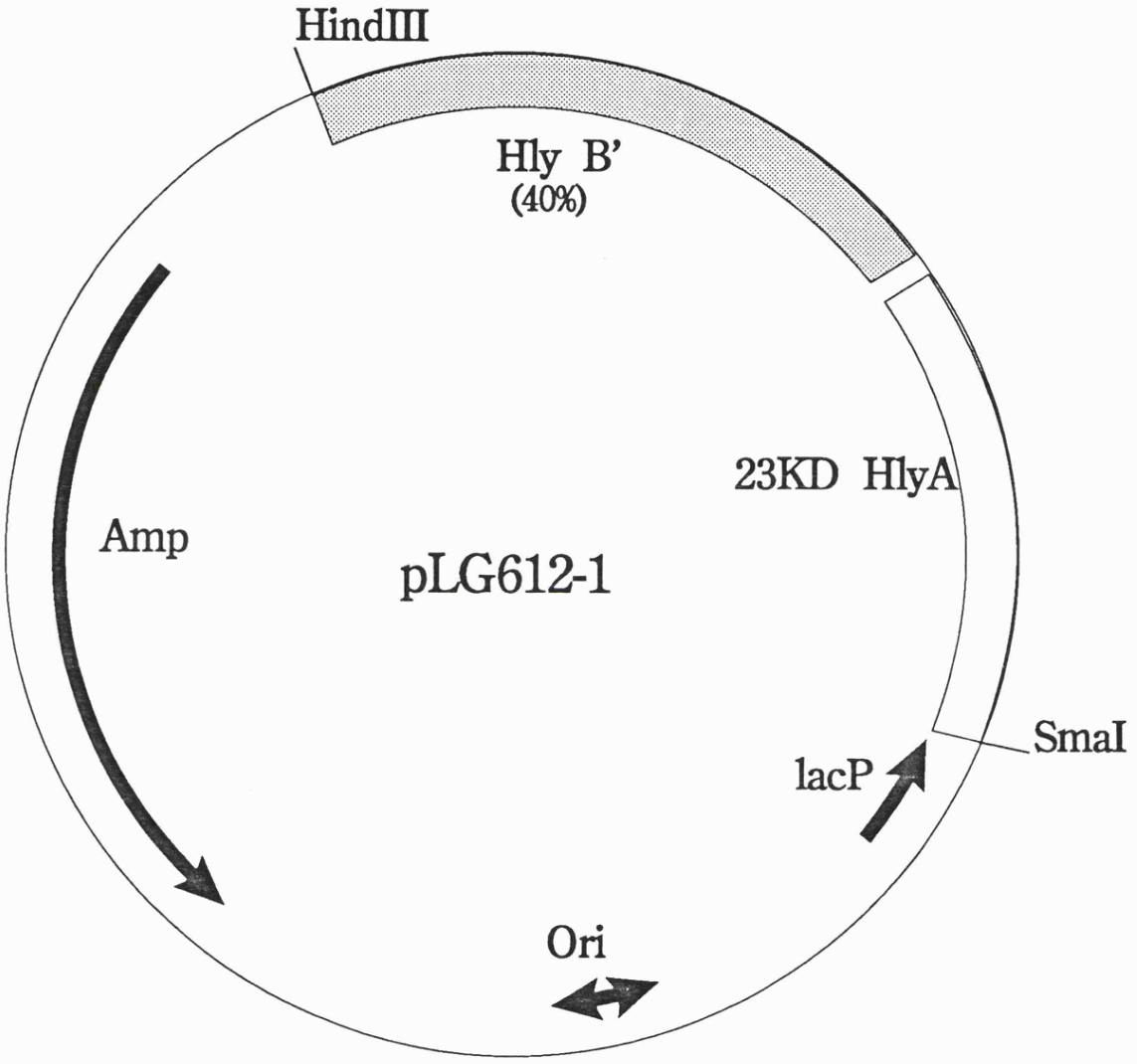

HindIII
Hly B'
(40%)
23KD HlyA
Amp
pLG612-1
SmaI
lacP
Ori

Table 1:

a) Results of scintillation counting culture fractions in secretion of 23KD HlyA ^{3}H labelling experiment.

A culture containing pLG612-1 (pUC-23KD*hlyA*) and pLG575 (*HlyB,D*) was initiated at an A^{450} of 0.1 in M-9 minimal media supplemented with leucine assay mix (see Materials and Methods) and induced for 23KD HlyA expression with IPTG (0.5mM final concentration) at an A^{450} of 0.36. The culture was split into two and growth maintained at 37°C for 30 minutes. At this stage 50μCi of ^{3}H-leucine together with cold leucine (1μg/ml final - to ensure constant uptake of radioactivity) was added to one 10ml culture. Only cold leucine (1μg/ml final) was added to the other culture to monitor growth. After a further 30 minutes protein synthesis and labelling was arrested by the addition of chloramphenicol (250μg/ml final concentration) and leucine (100μg/ml final concentration). The cells were isolated (15K, 10min) and washed 4 times in fresh minimal medium and resuspended in 400μl of minimal medium. Four $^{1}/_{100}{}^{th}$ samples, by volume, were removed and placed on filters for scintillation counting, two of which were washed with 10% w/v cold TCA (to remove any remaining unincorporated counts) - Sample A - the other two counted directly without washing, Sample B. The remaining cellular sample was resuspended in 2X sample buffer for analysis by SDS-PAGE. TCA (10% w/v final concentration) was added to a sample of the original culture supernatant and left on ice for 1 hour to precipitate any proteins. The TCA precipitated protein from the culture supernatant was spun down (10K, 15min) and resuspended in 0.4ml minimal medium. Four $1/90^{th}$ volume samples were removed and placed on filters for scintillation counting, two of which were washed with 10% w/v cold TCA (to remove unincorporated counts) - Sample E - and the other two counted directly, Sample F. The remaining supernatant sample was resuspended in sample buffer for further analysis by SDS-PAGE. $1/100^{th}$ of the culture supernatant was also taken and placed directly on filters (in duplicate) two of which were washed with 10% w/v cold TCA (to remove unincorporated counts) - Sample C - and the other two were counted directly (Sample D) to obtain an estimation of the counts remaining free in the culture supernatant. The table shows the total number of counts detected in each fraction for that experiment and as a percentage of total cellular detected counts.

b) Results of scintillation counting of ^{3}H-leucine eluted from gel slices.

0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample were separated on a 15% SDS-acrylamide gel. The gel was stained with Coomassie blue and dried down, and the secreted 23KD HlyA and the intracellular EFTu bands excised. The excised bands were rehydrated, separately, in 50μl of distilled water and placed into 5ml of NCS scintillant fluid, vortexed and left at 37°C overnight. The samples were then re-vortexed before counting. Table 2b shows the counts extrapolated to have been incorporated into both bands during the labelling period, adjusting for the 30% error apparently involved in the counting material in liquid scintillant rather than immobolised on solid filters (see text). The total counts obtained are also given as a percentage of total cellular labelled protein (% T.C.L.P).

a)

Samples counted on Solid Phase Support	Counts/Minute (X 10^6) during Labelling	% T.C.L.P
A) Cells: TW	7.67 +/- 2%	100
B) Cells: NW	8.2 +/- 5%	107
C) Medium: TW	0.58 +/- 25%	8
D) Medium: NW	2.2 +/- 3.6%	29
E) TCA Pellet: TW	0.13 +/- 9.2%	1.7
F) TCA Pellet: NW	0.2 +/- 8%	2.6

TW = Sample on filter washed with 10% TCA prior to counting
NW = Sample not washed with 10% TCA prior to counting
+/- = error margins from duplicate samples counted twice
% T.C.P.L = Percentage of total cellular labelled protein

b)

Samples counted in Liquid Phase	Counts/Minute (X 10^6) during Labelling	% T.C.L.P
Eluted Secreted 23KD HlyA Band	0.19	2.5
Eluted Cellular EFTu Band	1.58	20

Figure 2:

0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample (isolation of samples described in Table 1a above) were separated on a 15% SDS-acrylamide gel together with ^{14}C labelled markers. The resulting gel was first Coomassie stained and then dried down before exposing to ^{3}H sensitive autoradiograph film. This Figure shows a 10 day exposure with the molecular weight size markers on either side.

Lane 1: Cell sample after 30 minute ^{3}H-labelling.

Lane 2: Supernatant sample after 30 minute ^{3}H-labelling.

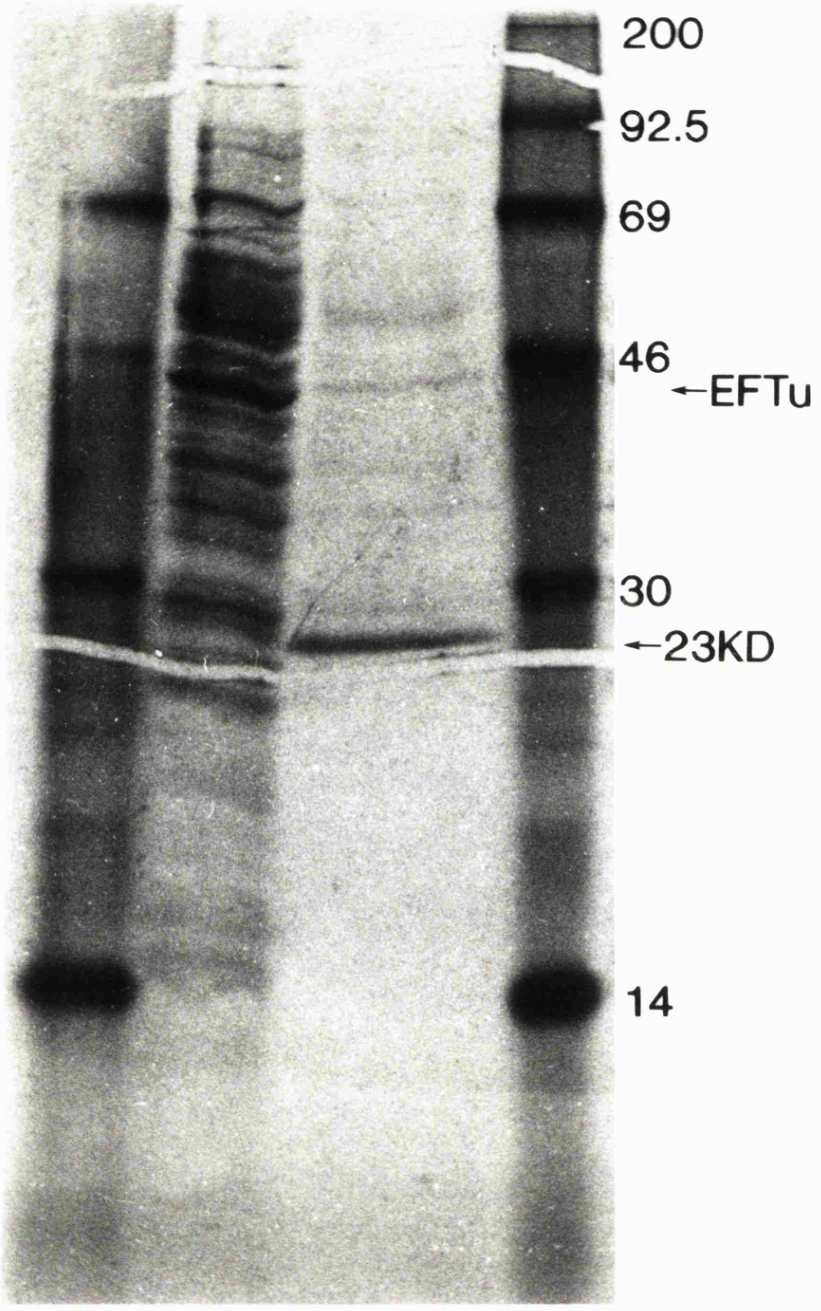
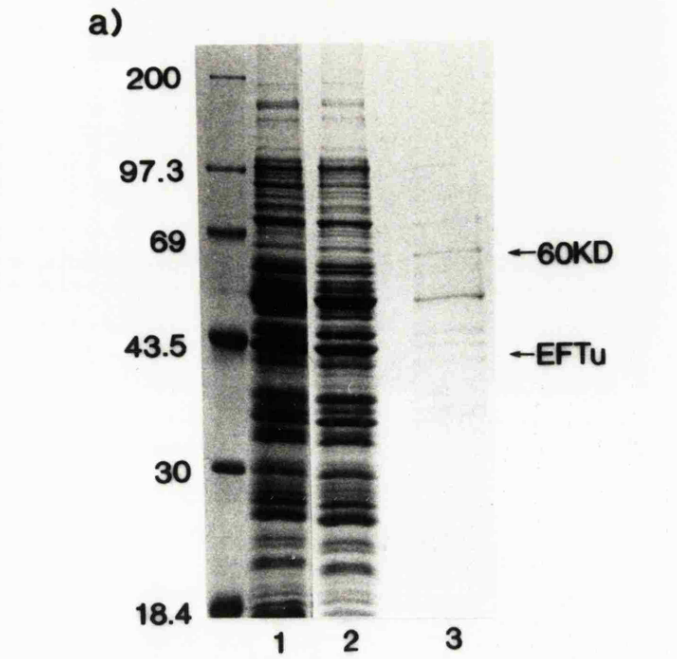

200
92.5
69
46
←EFTu
30
←23KD
14
1
2

Figure 3:

E. coli JM101 (F',*lacI*q) carrying pLG802 (Pcm-Bal-23KDHlyA) plus pLG575 (*hlyB,D*) was innoculated at an A^{450} of 0.2, at 30°C, in M-9 minimal media (with appropriate antibiotics) supplemented with leucine assay mix (see Materials and Methods) to an A^{450} of 0.62 before heat inducing high level expression. Growth was maintained at 37°C for a further 30 minutes before the culture was split into two 10ml samples. 50μCi of ^{3}H leucine together with cold leucine (1μg/ml final - to ensure constant uptake of radioactivity) was added, as before, to one culture. Only cold leucine (1μg/ml final) was added to the other culture to monitor growth. After a further 30 minutes protein synthesis and labelling was arrested by the addition of chloramphenicol (250μg/ml final concentration) and leucine (100μg/ml final concentration). The cells and culture supernatant samples were separated and prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were analysed by SDS-PAGE (11%) and visualised by Coomassie blue (a) and dried down and exposed to ^{3}H-sensitive autoradiograph film (b).

Lane 1: Cell sample before induction and labelling.
Lane 2: Cell sample after 30 minute ^{3}H-labelling.
Lane 3: Supernatant sample after 30 minute ^{3}H-labelling.

Arrow marks the position of the 60KD hybrid protein and EFTu with molecular weight markers also indicated.


a)
200
97.3
69
43.5
30
18.4
60KD
EFTu
1
2
3

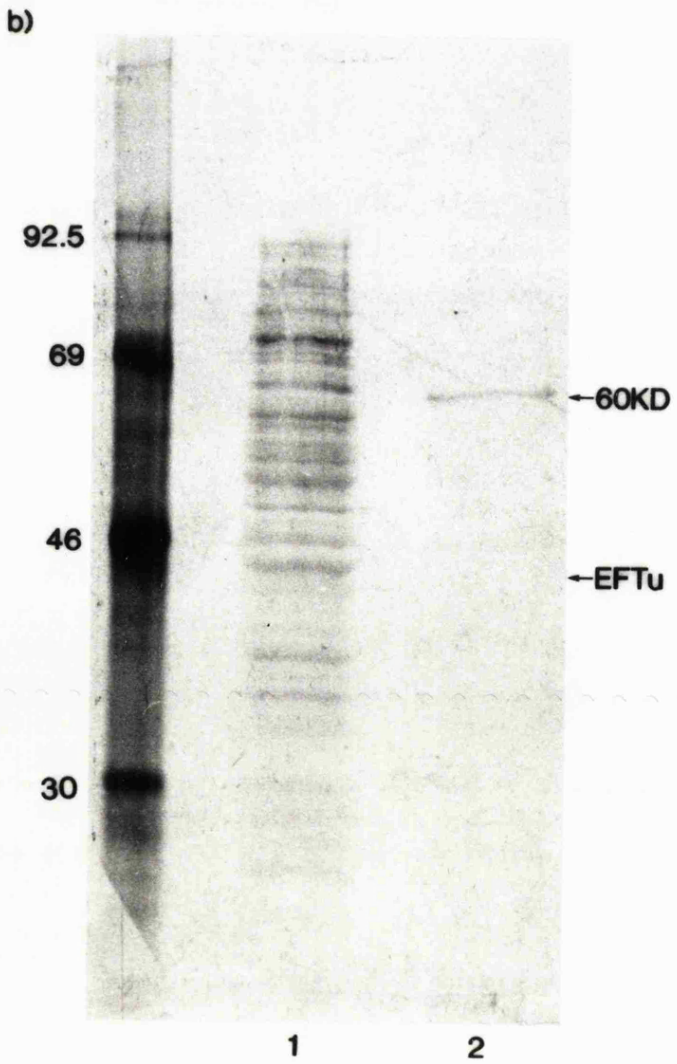

b)
92.5
69
46
30
60KD
EFTu
1
2

Figure 4:

Construction of pLG613 (pUC-23KD*hlyA* Km^R) and pLG614 (pUC-23KD*hlyA*,*hlyB*,*D* Km^R)

pLG612-1: see Figure 1; R. Haigh this laboratory

pUC4K: Vieira and Messing, 1982

pLG570: Mackman and Holland, 1984b

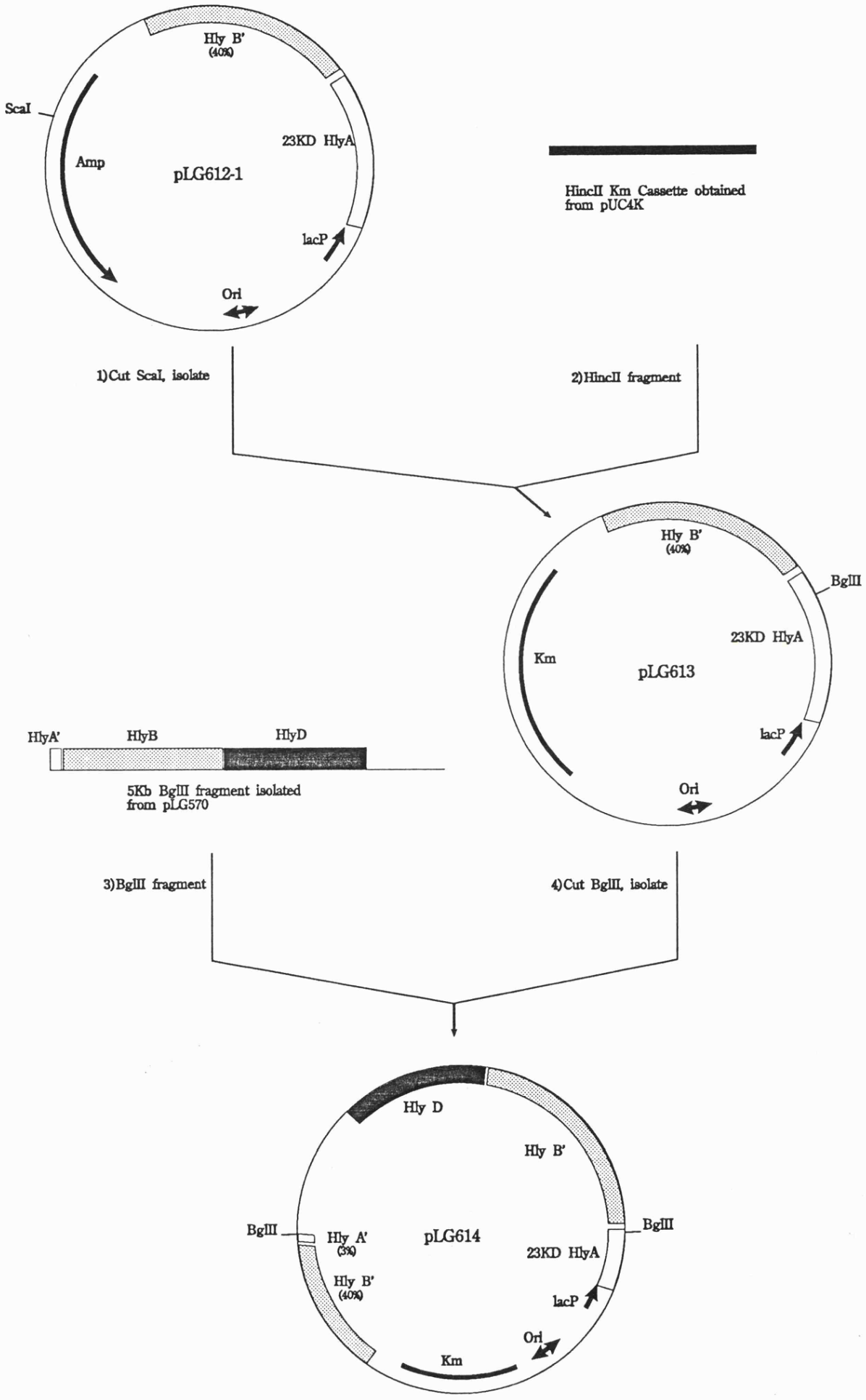

Hly B'
(40%)
ScaI
Amp
pLG612-1
23KD HlyA
lacP
Ori
HincII Km Cassette obtained
from pUC4K
1)Cut ScaI, isolate
2)HincII fragment
Hly B'
(40%)
BglII
23KD HlyA
Km
pLG613
lacP
Ori
HlyA'
HlyB
HlyD
5Kb BglII fragment isolated
from pLG570
3)BglII fragment
4)Cut BglII, isolate
Hly D
Hly B'
BglII
BglII
Hly A'
(3%)
pLG614
23KD HlyA
Hly B'
(40%)
lacP
Ori
Km

Figure 5:

Secretion of 23KD HlyA in M-9 minimal glucose versus rich medium.

E. coli NM522 (F',*lacI*q) containing pLG614 (pUC-23KD*hlyA*,*hlyB*,*D* Km^R) plus pLG575 (*hlyB*,*D*) was grown in M-9 minimal medium or nutrient broth (with appropriate antibiotics) from an A^{450} of 0.2 to 0.7 before inducing 23KD HlyA expression with IPTG (0.5mM final concentration). Samples were taken before, 1, 2 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were separated on a 15% SDS-acrylamide gel and visualised by Coomassie blue staining.

Cellular samples:

lane 1, 2, 3 and 4 time = 0, 1, 2 and 3 hours following addition of IPTG, respectively from the minimal medium grown culture with lane 5, 6, 7 and 8 time = 0, 1, 2 and 3 hours respectively after IPTG induction, in the nutrient broth grown culture.

Supernatant samples:

lane 9, 10, 11 and 12 time = 0, 1, 2 and 3 hours following addition of IPTG, respectively from the minimal medium grown culture with lane 13, 14, 15 and 16 time = 0, 1, 2 and 3 hours respectively after IPTG induction, from the nutrient broth grown culture.

Arrows mark the position of the secreted 23KD HlyA peptide and the unidentified band secreted specifically from minimal medium grown cells. The molecular weight markers are also indicated.

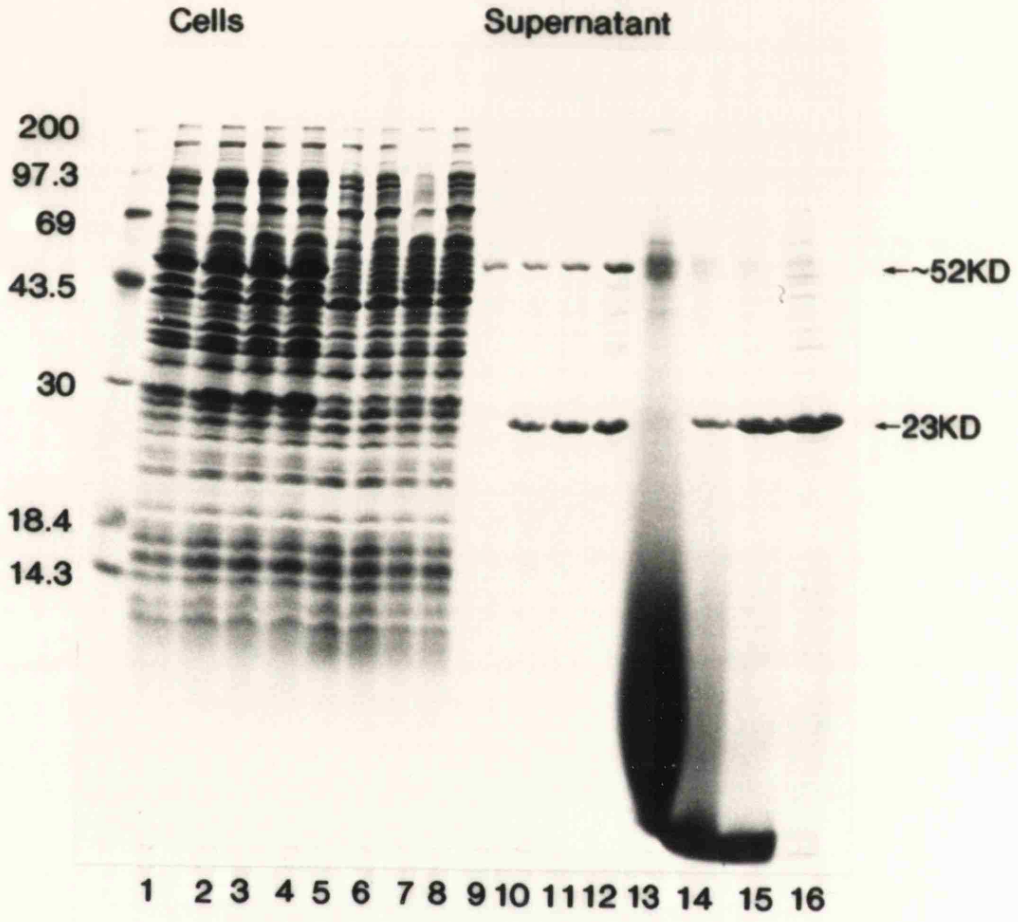

Cells
Supernatant
200
97.3
69
43.5
30
18.4
14.3
←~52KD
←23KD
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

Figure 6:

Secretion of 23KD HlyA in nutrient broth at 30 versus 37°C.

E. coli JM101 (F',*lacI*q) containing pLG612-1 (pUC-23KD*hlyA* Km^R) plus pLG575 (*hlyB,D*) was grown in nutrient broth (with appropriate antibiotics) from an A^{450} of 0.05 to 0.15 at 30°C. At this stage the culture was split into two and one maintained at 30°C and the other at 37°C. Upon reaching an A^{450} of 0.74 both cultures were induced for 23KD HlyA expression by the addition of IPTG (0.5mM final concentration). The culture maintained at 30°C was immediately split again into two and one was transferred to a 37°C waterbath and the other maintained at 30°C. Samples were taken before, 1.5 and 3 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 5 equivalent units of supernatant sample were loaded. Proteins were separated on a 15% SDS-acrylamide gel and visualised by Coomassie blue staining.

Cellular samples:

Lane 1, 2 and 3 time = 0, 1.5 and 3 hours following IPTG addition, from the culture grown throughout at 37°C, with lane 4, 5 and 6 time = 0, 1.5 and 3 hours respectively, from the culture grown and induced at 30°C. Lanes 7 and 8 time = 1.5 and 3 hours respectively, from the culture grown and induced at 30°C and then transferred to 37°C for the final 3 hours of growth.

Supernatant samples:

Lane 9, 10 and 11 time = 0, 1.5 and 3 hours following IPTG addition, from the culture grown throughout at 37°C, with lane 12, 13 and 14 time = 0, 1.5 and 3 hours respectively, from the culture grown and induced at 30°C. Lanes 15 and 16 time = 1.5 and 3 hours respectively, from the culture grown and induced at 30°C and then transferred to 37°C for the final 3 hours of growth.

The molecular weight markers are indicated.

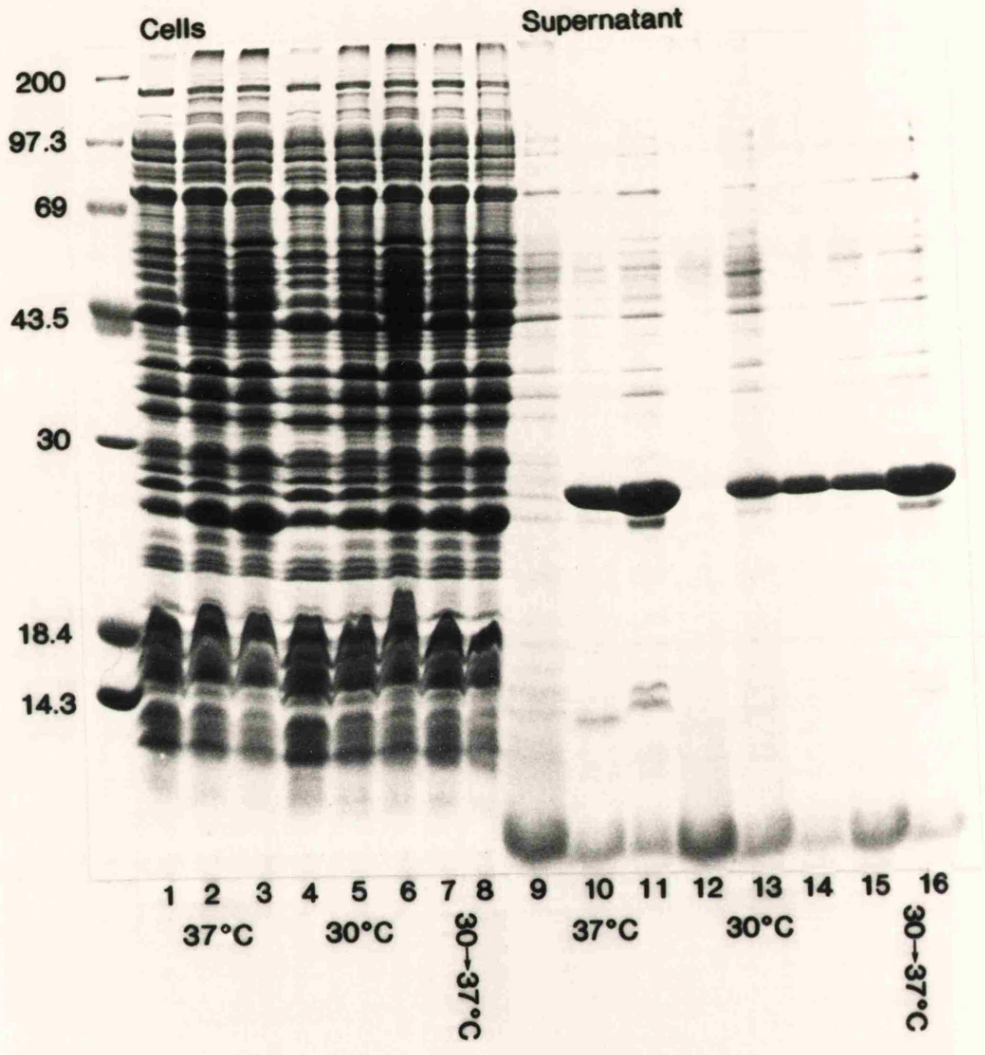

Cells
Supernatant
200
97.3
69
43.5
30
18.4
14.3
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
37°C
30°C
30→37°C
37°C
30°C
30→37°C

Figure 7:

Secretion of the 23KD HlyA peptide in nutrient broth at 37°C for quantification by laser densitometry.

E. coli NM522 (F',*lacI*q) containing either pLG612-1 (pUC-23KD*hlyA* Km^R) plus pLG575 (*hlyB,D*) or pLG614 (pUC-23KD*hlyA,hlyB,D* Km^R) were grown in nutrient broth (with appropriate antibiotics) from an A^{450} of 0.1 to 0.45 at 37°C before being induced for 23KD HlyA expression by the addition of IPTG.(0.5mM final concentration). Growth was continued for 3 hours before cell and culture supernatant samples were taken and prepared as described in Materials and Methods. 0.5 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were separated on an 11% SDS-acrylamide gel and visualised by Coomassie blue staining.

Lane 1: Cellular sample: t = 3 hours pLG612-1 (pUC-23KD*hlyA* Km^R) plus pLG575 (*hlyB,D*)

Lane 2: Supernatant Sample: t = 3 hours pLG612-1 (pUC-23KD*hlyA* Km^R) plus pLG575 (*hlyB,D*)

Lane 3: Cellular sample: t = 3 hours pLG614 (pUC-23KD*hlyA,hlyB,D* Km^R)

Lane 4: Supernatant Sample: t = 3 hours pLG614 (pUC-23KD*hlyA,hlyB,D* Km^R)

The position of the 23KD HlyA band is indicated together with the molecular weight markers. The position of a highly expressed band of ~25KD is highlighted in Lane 1 and represents the CAT gene product expressed from pLG575 (*hlyB,D*) see Chapter 4: Figure 3

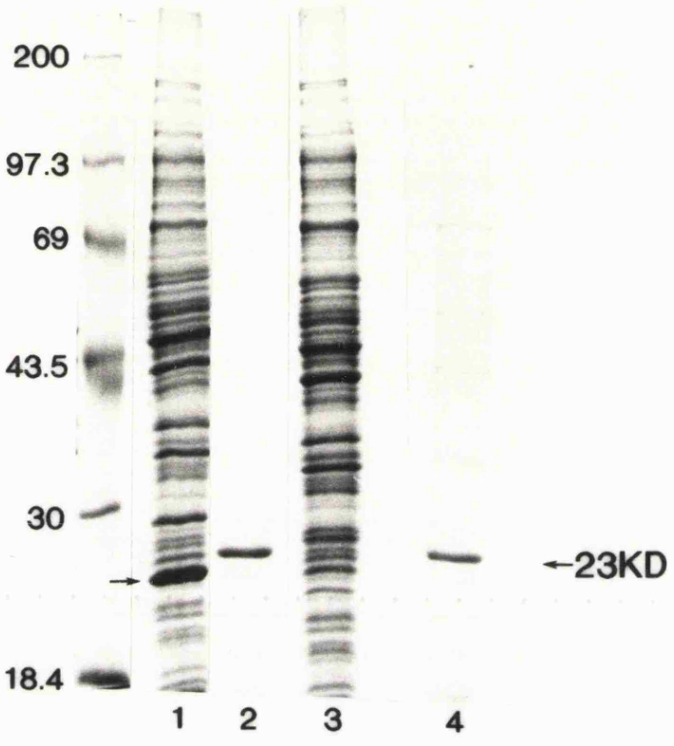

200
97.3
69
43.5
30
18.4
1
2
3
4
←23KD

Figure 8:

Secretion of the Pcm-HlyA hybrid peptide from cells grown in rich broth at 37°C: Quantification by laser densitometry.

E. coli SE5000 containing either pLG800-1 (38KD *pcm*-Xma-23KD*hlyA*), pLG801 (50KD *pcm*-Apa-23KD*hlyA*), pLG802 (60KD *pcm*-Bal-23KD*hlyA*), pLG806 (32KD *pcm*-Sma-12KD*hlyA*), pLG807 (44KD *pcm*-Bal-120bp*hlyA*) - see Table 2b below - plus pLG575 (*hlyB,D*) were grown at 30°C in Luria broth (with appropriate antibiotics). Cultures were initiated at an A^{450} of 0.26 and grown to A^{450} of 0.78 before heat inducing high level expression. Growth was continued for 3 hours at 37°C before cell and culture supernatant samples were taken and prepared as described in Materials and Methods. Various amounts of cell and supernatant samples were loaded to aid densitometry, separated on a 15% SDS-acrylamide gel and visualised by Coomassie blue staining.

Loading: 1 O.D unit of cells is defined as that equivalent to 1ml at A^{450} = 1 or that derived from an equivalent supernatant sample.

Cellular samples:

Lane 1: pLG800-1 (38KD *pcm*-Xma-23KD*hlyA*) 0.1 O.D loading

Lane 2: pLG801 (50KD *pcm*-Apa-23KD*hlyA*) 0.5 O.D loading

Lane 3: pLG802 (60KD *pcm*-Bal-23KD*hlyA*) 0.3 O.D loading

Lane 4: pLG806 (32KD *pcm*-Sma-12KD*hlyA*) 0.2 O.D loading

Lane 5: pLG807 (44 *pcm*-Bal-120bp*hlyA*) 0.3 O.D loading

Supernatant samples:

Lane 6: pLG800-1 (38KD *pcm*-Xma-23KD*hlyA*) 2.5 O.D loading

Lane 7: pLG801 (50KD *pcm*-Apa-23KD*hlyA*) 2.5 O.D loading

Lane 8: pLG802 (60KD *pcm*-Bal-23KD*hlyA*) 2.5 O.D loading

Lane 9: pLG806 (32KD *pcm*-Sma-12KD*hlyA*) 5.0 O.D loading

Lane 10: pLG807 (44KD *pcm*-Bal-120bp*hlyA*) 10 O.D loading

Arrows indicate predicted positions of fusions in total cell or supernatant profiles. The molecular weight markers are also indicated.

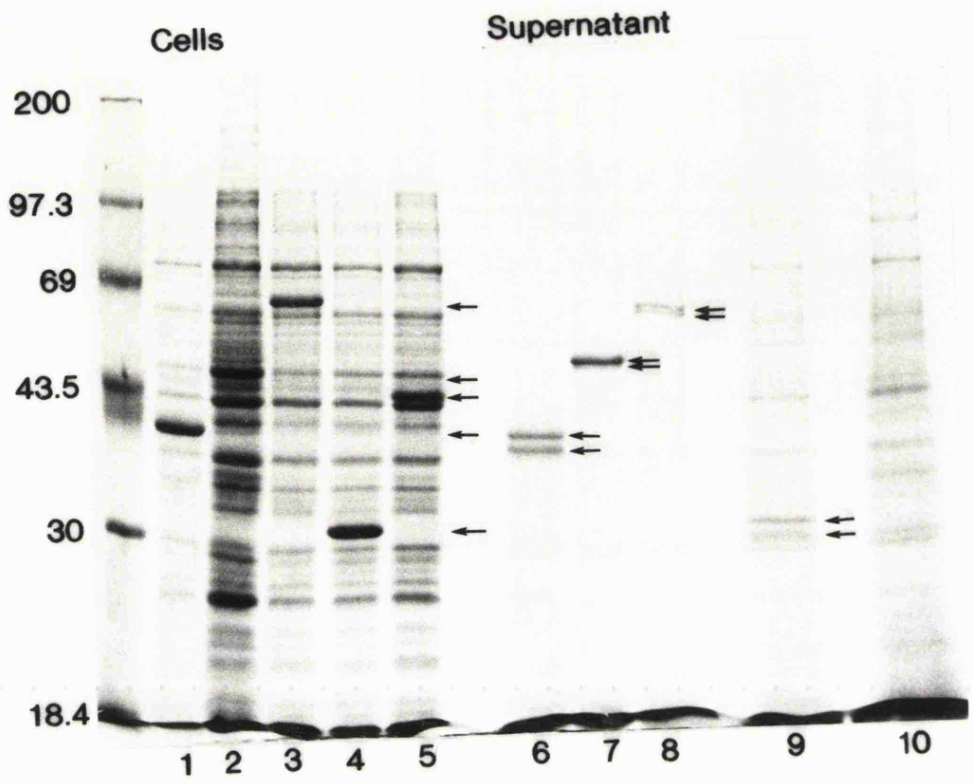

Cells
Supernatant
200
97.3
69
43.5
30
18.4
1
2
3
4
5
6
7
8
9
10

Table 2:

Results of laser scanning densitometry.

a) The SDS-acrylamide gels described above in Figures 7 and 8 were scanned using an LKB densitometer and the relevant data summarised in this table. The output from the scanned lanes expressed each individual band as a percentage of total bands detected in that lane and also as a relative absorbance value. Thus, the percentage of any particular intracellular band (EFTu or hybrid bands) is given directly whereas a secreted band could be expressed, indirectly, as a percentage of the corresponding total cellular profile. Thus, the absolute absorbance of a secreted band can be compared to that of an intracellular band and thus permit its estimation as a percentage of total cellular protein.

The table presents

A) the vectors expressing the different HlyA fusion peptides and the 23KD HlyA peptide,

B) apparent molecular weight of HlyA derivatives,

C) the amount of EFTu as a percentage of total cellular protein,

D) the amount of the intracellular HlyA hybrid protein as a percentage of total cellular protein,

E) the secreted protein as a percentage of

- i) its intracellular form
- ii) EFTu
- iii) total cellular protein

F) Level of secreted protein as a percentage of the secreted 23KD HlyA peptide, compensating for differences in molecular weight sizes.

G) Level of secreted fusion proteins as a percentage of the smallest secreted fusion protein, compensating for differences in molecular weight sizes.

b) Diagrammatical representation of the constructions encoding the various Pcm-HlyA hybrids quantified above.

a)

	A VECTOR		B Mr KD	C EFTU as % T.C.P	D HYBRID BAND as % T.C.P	E SECRETED BAND as % INTERNAL HYBRID BAND	E SECRETED BAND as % EFTU	E SECRETED BAND as % T.C.P	F RELATIVE SECRETION as % 23KD	G RELATIVE SECRETION as % 800-1
1)	612-1	(a)	23	5.0	–	–	59	2.95	100	–
2)	614	(b)	23	6.7	–	–	27	1.82	62	–
3)	800-1	(c)	38	4.4	25.9	2.84	18	0.79	16	100
4)	801	(c)	50	5.3	6.3	11.9	14.3	0.76	12	77
5)	802	(c)	60	6.0	9.6	6.42	9.9	0.6	7.6	50
6)	806	(c)	32	5.2	18.6	0.9	3.35	0.174	4.2	29
7)	807	(c)	44	6.3	9.3	–	–	–	–	–

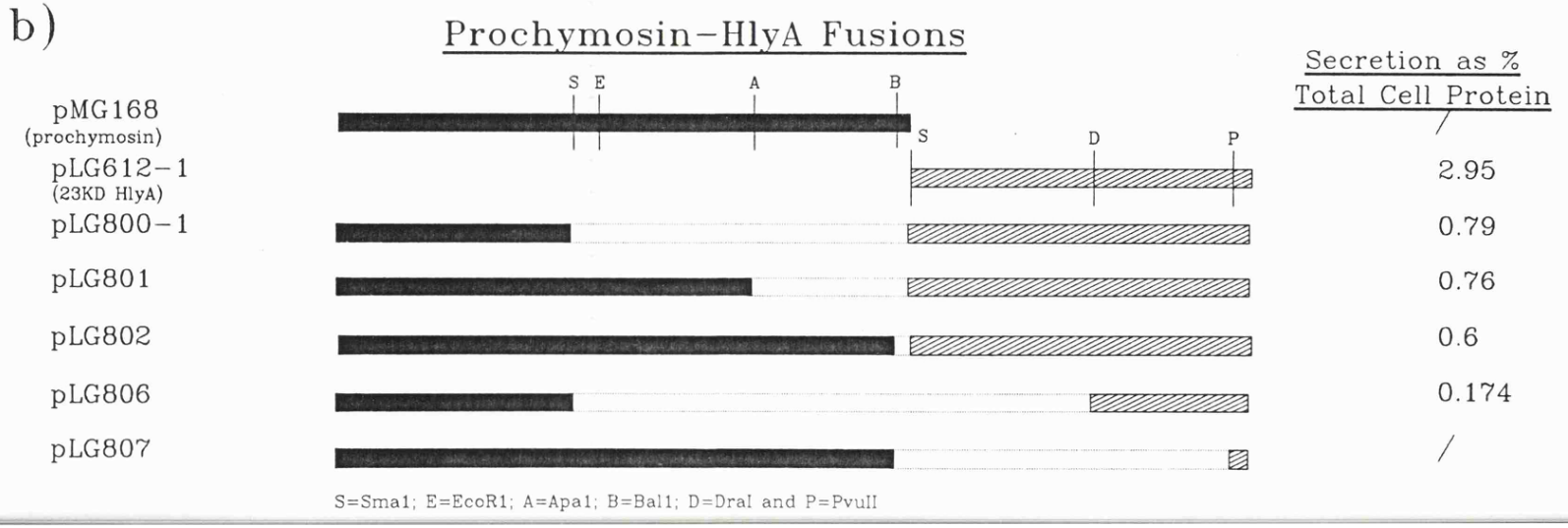
(a) = pUC-23KDHlyA HlyB,D provided on pLG575

(b) = pUC-23KDHlyA and HlyB/D

(c) = DOV-PcmHlyA Fusions (see below)

T.C.P = Total Cellular Protein

b)



CHAPTER 6

Analysis of the HlyA targeting signal using site directed mutagenesis (SDM)

6.1 Introduction:

This and the subsequent chapters are primarily concerned with attempts to identify those residues important in the targeting of the HlyA molecule. The position and character of such residues it was anticipated would allow us to deduce something of the basis of the recognition of HlyA by the export proteins. From previous work involving deletions and fusions of the signal domain of HlyA (see Chapter 1; Section 1.6.5) and studies described in Chapters 3 and 4 (summarised in Figure 1a, Lines 1 to 3 and Figure 1b), we may only conclude with confidence that the signal motif resides somewhere within the final 113 residues of HlyA (*Mackman* et al., 1987). Deletion and fusion analysis (Gray *et al.*, 1986; *Mackman* et al., 1987; see also Figure 1a; Lines 1-3) initially provided some evidence that the final 27 residues were important and perhaps largely sufficient for recognition. However, the inability to efficiently secrete fusion proteins fused to the final 39 residues of HlyA, together with the apparent inability to secrete the HlyA C-terminal 4KD peptide (final 39 residues) when expressed independently (see Chapter 4), suggested that although an important portion of the signal motif resides within the last 27 amino acids, residues extending up to and beyond the final 39 residues are also apparently required for recognition. The importance of the C-terminal region in secretion was also supported by the abolition of secretion by the addition of C-terminal peptide extensions (see Chapter 4; Section 4.7.3).

Deletion studies carried out during the course of this project by Hughes' group (Koronakis *et al.*, 1989) using a related *E. coli* determinant, demonstrated that fusion of the N-terminal 831 residues of HlyA to the final 53 residues (deleting the intervening portion) still permitted efficient secretion (Figure 1a, Line 5), whereas fusion to the final 38 residues reduced secretion by at least 90% (Figure 1a, Line 4). Thus, this finding is consistent with our observation that the C-terminal 39 residues of HlyA do not apparently encode all the information required for efficient recognition by the Hly export proteins. The results of Koronakis *et al.*, (1989) in turn suggested that all the relevant information for secretion might reside within the final 53 residues. However, it is necessary to be cautious with results obtained from fusion/deletion experiments as novel constructs may fortuitously recreate structures, required on the one hand for recognition, or on the other hand, create structural changes which result in occlusion of an otherwise intact signal domain. In either case the result could lead to incorrect conclusions.

The above type of problem was emphasised from the contradictory results obtained from two mutants with overlapping deletions within the *hlyA* portion of an *ompF*-23KD*hlyA* hybrid (see Mackman *et al.*, 1987; see Figure 1a, Lines 2,3). Internally deleting the HlyA region from Ser^{996} (27 residues from the C-terminus) to His^{906} still permitted substantial levels of hybrid secretion, suggesting that the signal lay entirely within the final 27 residues. However, an internal deletion from Ser^{996} further upstream to His^{857} was found unexpectedly to abolish efficient secretion although the upstream region deleted was previously shown to be dispensable for the secretion of the 12KD HlyA peptide (Mackman *et al.*, 1987). As a result of such ambiguous findings from deletion/hybrid analysis, it was decided to take an alternative approach in order to identify the residues constituting the HlyA targeting signal. The approach chosen was to identify single residues involved in targeting through induced mutagenesis.

Two strategies for this approach were designed. The first, involves the identification of any apparently conserved features within the C-terminal regions of related haemolysins and other proteins encoding C-terminal targeting signals, followed by the use of site directed mutagenesis (SDM) to probe/disrupt those features. This type of analysis should allow us to determine whether these putative "conserved" features constitute an essential part of the targeting signal or are required for efficient secretion at some other level. The second strategy (See Chapters 7 and 8) was to introduce random point mutations, into the specific 3' region of *hlyA* encoding the targeting signal. This strategy should, with time, lead to the saturation of this region with mutations generating a wide variety of amino acid substitutions. The progressive isolation and analysis of these mutants should pin-point those changes which do /do not affect secretion, highlighting those residues essential for recognition and hopefully permitting us to determine the nature of the targeting signal.

6.2 SDM of some of the conserved features within the HlyA signal domain:

6.2.1 Background:

In order to limit the scale of the task the decision was taken on the basis of all the fusion-deletion data, including those of Koronakis *et al.*, (1989) that the signal motif most likely resided within the final 53 C-terminal residues of HlyA. Accordingly, this region of the LE2001 HlyA toxin, together with the closely related haemolysin (Felmlee *et al.*, 1985a; Hess

et al., 1986), was examined for potential secondary structures (using Chou-Fasman prediction programme based on Chou and Fasman, 1978) which might be conserved between the homologues. The results indicated that two stretches of residues had the potential to form putative weak and strong alpha-helices, respectively (see Chapter 1; Section 1.6.5; Gray, 1987; and Figure 2). Indeed, helical structures have been postulated to be required for recognition of normal N-terminal and for mitochondrial signal sequences (see Chapter 1; Sections 1.2.3 and 1.4.2), warranting the investigation of the possible role of this feature in targeting.

Comparison of the final 53 residues of HlyA with the closely related toxins of *Proteus* and *Morganella* strains (Koronakis *et al.*, 1987) revealed extensive homology with a high background of amino acid substitutions in this C-terminal region highlighting conserved features (Koronakis *et al.*, 1989). Koronakis *et al.*, (1989) pointed out that the nature of the final 25 residues was weakly hydrophobic, rich in hydroxylated residues, that there was a cluster of charged residues (992-996) and a potential amphiphilic alpha-helical region (residues 972-989) conserved in these haemolysins (see Figure 2).

Another apparently conserved feature was noted (see Figure 3) when the HlyA signal domain was compared with more distantly related secreted toxins, such as leukotoxin from *Pasteurella haemolytica* (Lo *et al.*, 1987), adenylate cyclase toxin from *Bordetella pertussis* (Glaser *et al.*, 1988), *Erwinia chrysanthemi* protease (Delepelaire and Wandersman, 1989) or the closely related protease from *Serratia marcescens* (Nakahama *et al.*, 1986). This possibly conserved motif is what we have termed an "Aspartate box" being defined as a block of 12-14 mainly small uncharged residues, flanked by negatively charged residues usually aspartic acid (see Figure 3). It has also been reported that some peroxisomal proteins carry a small C-terminal extension which apparently directs import (Swinkels *et al.*, 1988) and also appears to feature an "Aspartate box" (see Figure 3).

Interestingly, there are no other apparent features, at the primary or secondary level, conserved in the C-terminal region of these molecules and a number of these features, for example the charged cluster or the amphiphilic helix are not conserved across the whole range of molecules, suggesting that the targeting signal might be composed of a more subtle, more complicated structure. Despite this lack of a readily identifiable, common, signal it is important to note that both leukotoxin and adenylate

cyclase are apparently recognised and translocated by HlyB,D in *E. coli* (Strathdee and Lo, 1989; Masure *et al.*, 1990) although at a low level. Indeed, recent studies also indicate that both a Psuedomonas protease (A. Lazdunski, personal communication) and an Erwinia protease (C. Wandersman, personal communication) are also apparently secreted by the HlyB,D export functions suggesting some clearly conserved similar recognition mechanism. However, the Erwinia protease is very inefficiently secreted (about 1% of the level of the homologous HlyA toxin), whilst the protease homologues of the HlyB,D export proteins cannot apparently export either HlyA or the 23KD C-terminal HlyA peptide, suggesting that the respective targeting signals although related are distinct and highly specific.

Thus, within this small 53 residue region of HlyA a number of features (both at a primary and secondary level) have been identified, although apparently not conserved in different C-terminal targeting signals. In order to evaluate the possible role of some of these features in the *E. coli* haemolysin system it was decided to specifically alter/disrupt them by specific mutagenesis and analyse the effect on HlyA secretion.

6.2.2 Disruption of the predicted strong Alpha-Helix:

As seen from Figure 3 the predicted strong alpha-helix in HlyA LE2001 consists of 15 amino acids, overlapping the cluster of charged residues. One way to disrupt this structure was to substitute a central residue of the helix (Ser^{996}) for a proline (a small non-polar amino acid known to be incompatible with alpha-helical structures; Creighton, 1984). Thus, a synthetic 17mer oligonucleotide was obtained which carried a single base alteration encoding the substitution of Ser^{996} for Pro. This change was to be introduced into the 3' of *hlyA* using the Amersham site directed mutagenesis kit as described in the next section.

6.2.3 Ser^{996} to Pro substitution by SDM:

The SDM kit necessitated in the first place the cloning of the relevant fragment to be altered into the bacteriophage M13 vector, in order to produce single stranded DNA, the template for the mutagenesis reaction. Therefore, the 1.6Kb SmaI/HindIII fragment from pLG609-1 (encoding both the 23KD C-terminus of HlyA and the first 40% of HlyB; See Chapter 3: Figure 2b) was cloned, as described previously, into the multiple cloning site of M13mp18 to generate M13mp18-23KDHlyA (see Appendix 7a). This double stranded vector was then transformed into *E. coli* JM101 (carries F' required for infection of M13) and used to isolate single

stranded DNA as a template for the SDM reaction (see Materials and Methods). The SDM kit was then used to introduce the required single base change, which was also designed to destroy the unique BglII site in *hlyA*, and transformants screened by restriction analysis for loss of the BglII site. The mutant sequence was then verified by single stranded sequencing (see Materials and Methods; data not shown).

6.2.4 Effect of the Ser^{996} to Pro substitution on the secretion of the 23KD HlyA peptide:

In order to determine whether the disruption of the putative strong alpha-helix structure affected secretion of the C-terminal 23KD HlyA signal domain, the altered fragment was transferred into a plasmid expression vector. Thus, the same 1.6Kb SmaI/HindIII fragment, from M13mp18-23KDHlyA, carrying the specific mutation was cloned into the multiple cloning site of pUC12 (as described before; see Chapter 4: Section 4.7.2 and Appendix 7a). This placed the *hlyA* fragment inframe with the *lacZ* initiation codon and also under the control of the *lacZ* promoter generating the plasmid pLG617. pLG617 is identical to pLG618 (see Appendix 7a) except that they carry different 3' *hlyA* point mutations. Double stranded pLG617 DNA was isolated, purified on CsCl gradients and double strand sequenced to verify the presence of the mutation in the plasmid construct (see Materials and Methods; data not shown).

pLG617 (pUC-23KDHlyA Pro) and pLG609 (pTTQ-23KDHlyA Chapter 3; Section 3: Figure 2) as a positive control, were transformed into *E. coli* JM101 (carries F', *lacI*q to repress expression from *lac/tac* promoters) in the presence or absence (+/-) of pLG575 (*hlyB/hlyD*). Overnight cultures were used to inoculate fresh nutrient broth to an A^{450} of 0.1 and then growth continued until an A^{450} of 0.4. At this point both cellular and supernatant samples were taken for later analysis, and IPTG (0.5mM final concentration) added to the remaining cultures to induce expression from the *lac/tac* promoters. After a further 2 hours of growth additional samples were taken and prepared as usual (see Materials and Methods) for analysis by SDS-PAGE. The resulting 15% gel is shown in Figure 4.

Lanes 1-4 and 5-8 do not show any intracellular accumulation of either the mutant or the wild-type 23KD peptides, respectively, after induction in the presence or absence of HlyB,D. The secretion of both peptides in an HlyB,D dependent manner was however clearly demonstrated after induction of expression (Lanes 12 and 16). Unfortunately, in this experiment the supernatant fraction contained many cellular bands,

presumably due to lysis or contamination. However, the absence of the 23KD HlyA peptide from the sample equivalent to the supernatant of the mutant peptide lacking the export function (Lane 14) was confirmed by Western blotting analysis with anti-HlyA antibodies (data not shown). However, the anti-HlyA antibodies did cross-react with both the secreted 23KD HlyA mutant and wild-type protein bands. In contrast, as previously noted, the intracellular form of the 23KD peptide is apparently rapidly degraded and not detected with antibodies (data not shown). It was not possible to directly compare the level of secretion of the wild-type and the mutant 23KD HlyA molecules as each is expressed from different promoters on different plasmids. Thus, to determine whether the introduction of this Pro residue into the middle of the predicted strong alpha-helix alters the efficiency of secretion, the mutation was transferred into the intact *hlyA* gene in order to monitor the secretion of the haemolytically active toxin compared with the wild-type, in this case from isogenic constructs and strains.

6.3 Effect of the Pro^{996} substitution on the efficiency of secretion of the active HlyA toxin:

Due to the absence of any unique or usable restrictions sites within the *hlyA* fragment carrying the SDM introduced point mutation, or indeed within available plasmids carrying the entire *hlyA* gene, this mutated fragment could not be easily re-introduced into the *hlyA* gene. However, other experiments had required the construction of a plasmid carrying both an intact *hlyC* and *hlyA* gene together with a number of flanking unique cloning sites to aid manipulation (see Chapter 7; Section 7.5.1). One of these plasmids pACYC-CA,ClaI (Chapter 7; Section 7.5.1.: Figure 7) carries both the *hlyC/A* genes deleted for the primary *hly* promoter region, but both gene products are still expressed *in vivo* presumably from an alternative upstream promoter. This construct has novel unique KpnI,SmaI and SalI restriction sites downstream of the *hlyA* gene, and also a unique ClaI site close to but within the 3' end of the *hlyA* gene. This ClaI site, introduced earlier by site directed mutagenesis (see Chapter 7; Section 7.2.3), is 447bp from the 3' end of *hlyA* and permits the precise exchange of the 3' end of the *hlyA* gene on a 0.6Kb ClaI/KpnI, SmaI or SalI fragment. Unfortunately, the *hlyA* 3' fragment encoding the Pro^{996} substitution did not carry this unique ClaI site so it first had to be introduced, as described in the next section, to permit the transfer of the mutant fragment.

6.3.1 Transferring the "Pro^{996}" mutation into an intact *hlyA* gene:

The strategy taken to transfer the Pro^{996} mutation into the intact *hlyA* gene is depicted in Appendix 8a/b. This was a long and complicated procedure requiring the construction of a number of intermediate plasmids, in order to avoid the need to repeat the SDM reaction to introduce the ClaI site. First of all, the 0.8Kb EcoRI/SalI *hlyA* fragment from pACYC-CA,ClaI (carries the unique ClaI site) was cloned into the multiple cloning site within pUC18 (deleted for its PstI site; see below) in order to aid its subsequent manipulation. The next step was to swap the approximately 250bp PstI/SmaI 3' *hlyA* fragment for the PstI/HpaI fragment carrying the Pro mutation from pLG617, thereby placing it downstream of the unique *hlyA* ClaI site. The pUC18 vector also carries a PstI site and so this was first deleted using T4 DNA polymerase (see Materials and Methods) before cloning in the initial *hlyA* fragment (see Appendix 8a). After these steps, it was now possible to clone the 0.6Kb ClaI/SalI *hlyA* fragment, now encoding the Ser^{996} to Pro substitution, into the same sites in pACYC-CA,ClaI recreating the entire *hlyA* gene, pACYC-CA,Pro^{996} (see Appendix 8b). The insertion of the mutant fragment was confirmed by restriction analysis, as the point mutation destroys the unique BglII restriction site in *hlyA*.

6.3.2 Haemolytic assay of Pro^{996} mutant versus wild-type HlyA:

Both pACYC-CA,ClaI and its derivative carrying the Ser^{996} to Pro substitution were transformed, using the $CaCl_2$ method (see Materials and Methods), into *E. coli* NM522 containing pLG339B/D (*hlyB*/*hlyD*; see Chapter 7; Section 7.5.2). Overnight cultures were used to inoculate fresh nutrient broth containing 10mM $CaCl_2$ (which appears to stabilise the secreted toxin activity; Mackman and Holland, 1984a) to an A^{450} = 0.1. Cultures were grown with shaking at 37°C and at intervals samples were taken to monitor optical density and haemolytic activity which was measured as described in Materials and Methods. A preliminary experiment was carried out, taking samples in duplicate, to monitor the growth and haemolytic activity of these cultures (see Materials and Methods). However, some of the duplicate samples taken near the end of the experiment differed significantly so the experiment was repeated, with modifications to the assay procedure. The results of this second experiment are depicted in Figure 5 with the average of both duplicate haemolytic assay samples expressed in arbitrary Haemolytic Units (H.U).

1 H.U. represents the A^{543} reading derived from 1 A450 unit of cells. This time there was no significant difference in the duplicate assay readings.

The growth curves of the two cultures were virtually identical until stationary phase when some lysis appeared to occur in the control strain. Comparison of the pattern of release of haemolysin throughout the growth phase revealed no significant difference between the two forms of the secreted toxin. Indeed, the Pro mutant showed higher levels of secreted activity in the first peak. These results therefore indicate that the introduction of a proline residue within the predicted strong alpha helix had no inhibitory affect upon secretion or the activity of the toxin.

Surprisingly, in these experiments the pattern of HlyA secreted toxin activity was quite different from that previously reported with the original wild-type *E. coli* strain LE2001 or K12 strains carrying the cloned determinant, pLG570, where expression is initiated in early exponential phase and switched off in late exponential phase generating one peak of secreted activity (Nicaud *et al.*, 1985b). With the complementation experiments carried out here a similar peak of release in the exponential phase was observed but was also followed by an additional broad peak of toxin release after cessation of exponential growth (A^{450}, 8-10). It is important to note that the primary *hly* promoter reported to be located upstream of the *hlyC* gene in the intact determinant (Hess *et al.*, 1986; Welch and Pellett, 1988; Koronakis and Hughes, 1988 and Koronakis *et al.*, 1988), is not present in these constructs. The expression of the *hlyB,D* genes is presumably controlled, at least in part, from the *tet* promoter upstream of *hlyB,D* (pLG339B,D see Chapter 7; Section 7.5.2), whereas *hlyC* and *hlyA* are expressed on the pACYC-CA,ClaI plasmid (Chapter 7; 7.5.1) from an unknown promoter. It has in fact been reported that both *hlyB* and *D* can be expressed from promoter-like sequences immediately upstream of the respective genes (Felmlee *et al.*, 1985a; Hess *et al.*, 1986 and this laboratory unpublished observations). Thus, the appearance of the second, delayed release of toxin activity might be due to regulation of expression at these "secondary" promoters. This point will be returned to in more detail in section 6.6.

6.4 Investigation of a role for the "Aspartate box" in signal recognition:

In order to determine whether the apparently conserved "aspartate box" (AB) is indeed an essential feature in HlyA targeting we decided to alter some of its features. The AB sequence is defined as two negatively charged residues (usually aspartic acid) separated by 12 to 14 mainly small uncharged residues. A strategy of site directed mutagenesis was therefore adopted to replace the two flanking negatively charged residues with large positive ones. In a second approach it was also decided to replace one of the small internal residues with a large bulky amino acid to test the requirement for a string of small residues in this region of the signal domain.

6.4.1 Specific amino acid replacements in the HlyA targeting region by SDM:

Glu^{994}

The 5' negative charge of the postulated aspartate box in HlyA is in fact a large Glutamic acid (Glu^{994}). Only a single base change (**G**AA to **A**AA) is required to generate a Lysine (Lys, a large positively charged residue) at this position and an appropriate 17mer oligonucleotide was utilised for this purpose. There are 81 Lys residues within the HlyA molecule and 73% are encoded by the AAA triplet, consequently this change was consistent with the normal codon usage for HlyA.

Asp^{1009}

It was simpler to convert the small Asp^{1009} encoding codon to that coding for the large positive Arginine residue (Arg) than Lys so a 21mer oligonucleotide carrying the two required alterations (**GA**T to **CG**T) was obtained. Arginine is only present 29 times in HlyA but the codon chosen is the most frequently used (28%).

Introduction of a large residue in the middle of the Aspartate Box

The Serine^{999} residue within the aspartate box was arbitrarily chosen for substitution for a large Tryptophan residue (Trp). This required the introduction of two base changes (T**CT** to T**GG**) incorporated into a 21mer oligonucleotide for *in vitro* mutagenesis. In fact, due to an error in transcribing the details of the oligonucleotide, the molecule obtained encoded the following changes: TCT GCC GCT T**GG** TTA T**G**G CAG. Fortuitously, the second codon alteration resulted in the conversion of Leu^{1001} to a second Trp residue, which should in fact more rigorously

probe the effect on secretion of large bulky residues in this region. However, this extra change might reduce the chances that the oligonucleotide can bind specifically to the template, due to possible bubbling effects from the two regions of mismatch in a relatively short oligonucleotide.

6.4.2 Facilitating the SDM reaction:

Again the SDM protocol required that the template for mutagenesis be cloned into the M13 bacteriophage vector in order to generate single stranded DNA templates. In addition, to aid the re-introduction of the mutations into the intact *hlyA* gene, it was necessary to clone the 3' fragment of *hlyA* which carries the unique ClaI site (introduced by SDM; see Chapter 7: Section 7.2.3) into M13 for use as the template for future SDM reactions. Thus, the 0.8Kb EcoRI/KpnI fragment (encoding the C-terminal 23KD of HlyA and the first 24 residues of HlyB) from pACYC-CA,ClaI was inserted into the multiple cloning site of M13mp18, generating M13mp18-23KDHlyACla (see Figure 6). This was then used to produce single stranded templates for the SDM reaction (see Materials and Methods).

The two oligonucleotides encoding the Glu^{994} to Lys and Asp^{1009} to Arg substitutions were used in the same SDM reaction (as described by the manufactures, see Materials and Methods), in an attempt to isolate single and double mutants to test the additive effect of altering both residues. A second reaction was carried out simultaneously with the oligonucleotides encoding the Asp^{1009} to Arg change and the Ser^{999} and Leu^{1001} to Trp changes. Three resulting M13 transformants from each of these reactions were used to generate single stranded DNA and sequenced using the sequenase kit (see Materials and Methods), in order to determine the success of the reactions. All six transformants carried a specific mutation (see Figure 7), but only the first (Figure 7, Number 1) carried the Glu^{994} to Lys change, the others all carried the Asp^{1009} to Arg change (see Figure 7). This bias in the frequency of mutants obtained could well be explained by better competition for the template (higher molarity or better binding ?) of the Asp^{1009} oligonucleotide compared to the other two. In an attempt to isolate a mutant carrying the Trp substitutions, I prepared single stranded DNA from 9 more transformants from the appropriate reaction and from the wild-type M13mp18-23KDHlyA,ClaI clone, sequencing only the track where all the expected changes should be evident (changing three C's to G's). However, I was still unable to detect any differences from the positive control (data not shown). I also carried

out a separate SDM reaction using only the oligonucleotide encoding the Trp substitutions but again failed to isolate the required mutant. From, these results it must be assumed that the 21mer oligonucleotide carrying 3 base changes (see Section 6.4.1) cannot bind to the specific *hlyA* region either because of the presence of too many mismatches in this 21mer oligonucleotide or the presence of secondary structure preventing its correct binding. It is also conceivable that the mutants might encode forms of the 23KD protein deleterious to the cells in some way, thereby preventing their isolation. However, this is unlikely as normally the intracellular 23KD HlyA is rapidly degraded unless rescued by secretion.

6.4.3 The effect of Glu^{994} and Asp^{1009} substitutions on secretion:

In order to assess the effect of the residue substitutions on secretion it was necessary first to transfer the mutations into the intact *hlyA* gene. This would then permit the monitoring of the secretion of the active toxin. This was easily achieved by isolating the 0.6Kb ClaI/KpnI fragment from the M13mp18-23KDHlyA,ClaI mutants and transferring this into the same sites within pACYC-CA,ClaI (see Figure 8). Double stranded plasmid DNA for both the mutants was isolated and purified on CsCl gradients and the presence of the mutations verified by double stranded sequencing, again using the sequenase kit (see Materials and Methods; data not shown)

The pACYC-CA,Glu^{994}, pACYC-CA,Asp^{1009} and pACYC-CA,ClaI (wild-type) plasmids were transformed into *E. coli* NM522 containing pLG339B/D (*hlyB/hlyD*) and used to inoculate fresh nutrient broth (supplemented with 10mM $CaCl_2$ to increase the stability of secreted toxin), to an A^{450} of 0.1. At intervals during growth at 37°C, samples were taken to monitor both optical density and haemolytic activity exactly as described previously (also see Materials and Methods). This experiment was also carried out on a separate day and virtually identical results were obtained, with the data from one such experiment depicted in Figure 9. The results demonstrated that changing the large negative residue, Glu^{994}, to a large positive Lys residue did not alter the activity profile significantly from that of the wild-type control. However, conversion of the small residue, Asp^{1009}, to a large positive Arg reduced the level of the active toxin detected in the supernatant during late stationary phase by about 70%. A similar reduction was not evident in the earlier short burst of activity. Indeed, the activity detected in this short window is variable from experiment to experiment, unlike the broad post exponential phase

activity (data not shown). The most probable reason for this is due to imprecise optical densities values obtained when measuring low numbers of cells, any inaccuracies of which are then magnified when the detected haemolytic activity is expressed per optical density unit.

It was important now to try to distinguish whether this single residue substitution (Asp^{1009} to Arg) was affecting secretion of the toxin or just the stability of the toxin. One way to examine this was to determine the half-life of the mutant toxin compared to the wild-type toxin.

6.4.4 Half-Life of the Asp^{1009} mutant toxin versus wild-type HlyA:

Overnight cultures of the above strains, expressing the wild-type and Asp^{1009} to Arg mutant, were used to inoculate fresh nutrient broth (containing 10mM $CaCl_2$) to an A^{450} of 0.1. Growth was monitored and at an A^{450} of about 4.5, 3ml of each culture was taken and the cells spun out. 1.5ml of the resulting supernatant was removed and Sodium Azide (final concentration of 5mM) was added to prevent the growth of any remaining contaminating cells. The supernatant was then incubated at 37°C and samples were removed at 15 minute intervals over a 105 minute period and monitored for haemolytic activity, as described above (see Materials and Methods). The results obtained are depicted in Figure 10 and reveal virtually identical decay patterns but again with an approximately 70% difference in the detectable haemolytic activities. The half-life of both of the toxin samples was approximately 25 minutes which contrasts with that of 2.5 hours previously reported for haemolysin in rich broth plus 10mM $CaCl_2$ (Nicaud *et al.*, 1985b). This greater instability of haemolysin presumably reflects some change in the culture conditions (nutrient broth used instead of luria broth, with the secreted toxin activity being assayed after the end of exponential growth in the present study compared to early exponential growth phase) leading to a much more rapid inactivation of the toxin in the supernatant samples.

6.5 Comparison of the level of 107KD polypeptide and haemolytic activity in the medium using a 2 plasmid *hly* complementation system:

Nicaud *et al.*, (1985b) previously reported that high level transcription of *hlyC* and *hlyA* is abruptly switched off before the end of the exponential growth phase apparently involving regulation of the promoter upstream of *hlyC*. Those workers also found that "switch off" was accompanied by rapid loss of haemolytic activity and proteolytic degradation of the 107KD

protein. Wagner *et al.*, (1988) have reported that this loss of haemolytic activity may not always be due to degradation of the secreted toxin but to reversal of the activation step promoted, in an as yet unknown way, by HlyC. Thus, I decided to test whether the two peaks of secreted haemolytic activity directed by the subcloned *hly* determinant, missing the reported primary promoter, reflected genuine changes in the levels of toxin secreted, degradation or possibly transient secretion of an inactive toxin.

The wild-type plasmid pACYC-CA,ClaI (*hlyC,A*) was transformed into NM522 carrying pLG339BD (*hlyB,D*) to monitor secreted toxin activity and 107KD polypeptide distribution between cells and supernatant during the growth cycle. As a control I monitored the secreted activity as well as the intracellular and extracellular accumulation of a C-terminally truncated HlyA molecule expressed from an identical vector and in an isogenic strain. In the construct chosen for this purpose the 3' *hlyA* ClaI/KpnI fragment (encoding the final 147 residues of HlyA) was substituted for a similar sized fragment from the M13mp19 vector. This truncated form of HlyA, deleted for the final 147 residues, was predicted to produce an approximately 94KD HlyA truncate, and expected to remain haemolytically active, since a C-terminal truncate of about 200 residues was reported by Felmlee and Welch (1988) to be active but not secreted. Cultures of both strains were initiated at an A^{450} = 0.21 in nutrient broth (supplemented with 2.5mM $CaCl_2$). Samples were taken every hour and the optical density recorded together with the supernatant haemolytic activity (in duplicate) as described before and in Materials and Methods (See Figure 11a). At the same time intervals as above, samples were taken, cellular and supernatant fractions were isolated and prepared, as usual, for analysis by SDS-PAGE (see Materials and Methods). O.1 A^{450} units of each cellular sample and 1 equivalent A^{450} unit of TCA concentrated supernatant sample were separated on a 11% SDS-acrylamide gel, transferred to nitrocellulose and probed with anti-23KDHlyA antibodies. The results obtained are shown in Figure 11b. From Figure 11a/b it is apparent that the growth of both cultures were identical, and as expected no secreted haemolytic activity or protein was detected from the strain encoding the truncated HlyA. In particular it is important to note the absence of the late burst of haemolytic activity from the strain encoding the truncated form of HlyA which confirmed that such activity in the wild-type, truely, represents secretion of the toxin and

is not due to some non-specific factor. The wild-type HlyA expressed from an identical vector gave the expected dual peak of secreted haemolytic activity.

In contrast to the two peaks of secreted haemolytic activity, surprisingly the 107KD protein detected by antibody was secreted at an apparently constant rate/ unit cell mass throughout the growth period and without detectable degradation (Figure 11b). This contrasts with the detected oscillating toxin activity, suggesting that either the toxicity of the HlyA molecule can be modulated after release from the cell or that the specific activity of the newly synthesised and secreted toxin can be regulated in a growth phase dependent manner. Western blot analysis of equivalent sample loadings of the intracellular protein detected both the expected "94KD" truncate and WT HlyA forms in the appropriate samples (Figure 11b). The intracellular levels of detected WT toxin appeared to increase with time. Whether this reflects some increase in transcription/ translation of HlyA, increased stability or reflects the gradual saturation of the secretion pathway is not known. The truncate, expressed from an identical vector gives similar results to that of wild-type HlyA 107KD polypeptide but at reduced levels (Figure 11b). Thus, the truncated HlyA molecule, due to its inability to interact with the export machinery might be more unstable and therefore degraded more rapidly. However, no degradation products were detected within the cellular samples suggesting as an alternative some form of feedback regulation of expression, other than at the level of the *hly* primary regulatory promoter, which is absent in these constructs.

6.6 Discussion:

Before discussing the results obtained from the site directed mutagenesis it is worth first of all commenting on the haemolytic assay system used to determine the affect of substitutions on the secretion process. As it was not possible to use the intact *hly* determinant in these experiments, due to its unsuitability for genetic manipulation, a more adaptable system was devised. This provided *hlyC* and *hlyA* on one plasmid (pACYC-CA; derived from the low copy number vector pACYC184; Chang and Cohen, 1978) and *hlyB/hlyD* on a second plasmid (pLG339B/D; derived from the low copy number vector pLG339; Stoker *et al.*, 1982). Consequently, all of these genes are now divorced from any regulation via the reported primary *hly* promoter upstream of the *hlyC* gene. Prior to the construction of this 2 plasmid *hly* system, I had also generated a 3 plasmid system with *hlyC* and *hlyA* provided on two separate pUC based vectors

(pLG591Tc and pLG585 respectively; both high copy number plasmids; construct not shown) and *hlyB,D* on the previously constructed low copy number plasmid pLG575 (Mackman *et al.*, 1985b). Again, in these constructs the regulatory region upstream of *hlyC* was absent. Interestingly, *hlyA* was found to be expressed when inserted in either direction relative to the pUC *lac* promoter (differing in expression levels by ~10 fold; data not shown), suggesting the existence of a weak promoter like sequence immediately upstream of the *hlyA* gene. There is also some evidence that *hlyC* deleted for its upstream promoter is also expressed independently by an alternative upstream region (this laboratory; unpublished observations). Indeed, cloning both the *hlyC* and *hlyA* genes, minus the upstream *hly* promoter region into the apparently promoterless region of the pACYC184 plasmid (see Chapter 7; Section 7.5.1) still allows both genes to be expressed, again indicating the presence of alternative promoter like regions. Other data obtained by Felmlee *et al.*, (1985a), Hess *et al.*, (1986) and this laboratory (unpublished observations) indicate that additional promoters may also exist immediately upstream of *hlyB* and *hlyD* respectively. The role of these promoters in normal regulation of Hly expression is unclear and may indeed change when the *hly* genes are cloned independently.

The haemolytic profiles and growth curves from a series of separate experiments involving the secretion of the HlyA toxin from the wild-type *E. coli* LE2001 strain and the toxin secreted from the 2 and 3 plasmid *hly* systems mentioned above are shown in Figure 12. It should be noted that the three *E. coli* strains used (LE2001, NM522 and SE5000 respectively) are unrelated. In addition, only the first two experiments involved the addition of 10mM $CaCl_2$, which appears to stabilise haemolytic activity and thus aid detection of the secreted toxin (Mackman and Holland, 1984a). The secreted haemolytic activity was monitored by assaying for the release of haemoglobin from sheep red blood cells at a wavelength A^{543} (see Gray, 1987). Activity measured in these experiments was expressed as that detected at A^{543} per A^{450} unit of cells. The haemolytic activity profile for the wild-type, pathogenic strain LE2001 was similar to that reported previously by Mackman and Holland (1984a) for this strain. Thus, the secreted toxin activity was detected early in exponential growth, peaking and rapidly decreasing as the cells enter the late exponential phase of growth. This loss of detectable HlyA activity co-incides with the switch off of *hly* transcription from the *hly* promoter (Nicaud *et al.*, 1985b). In many cases the loss of HlyA activity in the medium has been attributed to the accompanying proteolytic breakdown of the toxin (see

Nicaud *et al.*, 1985b). However, as reported by Wagner *et al.*, (1988) and Oropeza-Wekerle (1989) the loss of activity can still occur in the absence of degradation apparently by reversal of the activation process.

When the supernatant haemolytic profiles directed by the subcloned *hly* systems are examined (Figure 12), an burst of haemolytic activity early in the exponential phase which decreases in mid to late exponential phase was observed, as previously seen with the wild-type LE2001 strain. However, quite differently from the wild-type determinant, haemolytic activity directed from the subcloned determinants peaked again during late stationary phase before declining as growth ceased. The high copy number 3 plasmid *hly* complementation system (Figure 12c) in fact appeared to inhibit the growth of the cells shifting the haemolytic profiles slightly but the overall effect was the same as with the 2 plasmid *hly* system (Figure 12b). The high copy number, 3 plasmid system apparently secreted about twice the amount of active toxin as compared with that of the 2 plasmid system (low copy number), although this is probably an underestimate as the former experiment was not carried out in the presence of 10mM $CaCl_2$ which increases the stability and detection of the toxin activity. However, the level of secreted haemolytic activity detected from the 2 and 3-plasmid complementation systems during early exponential phase (first peak) was only about one-tenth that of the wild-type strain LE2001.

Analysis of the levels of secreted 107KD HlyA polypeptide directed by the 2 plasmid system (see Section 6.5) demonstrated that the amount remained fairly constant throughout the whole growth phase, whereas the toxin activity clearly oscillates. Therefore, the cells apparently secrete the HlyA polypeptide at a constant rate (without any signs of degradation) whilst toxin activity fluctuates. There might be two possible alternatives to explain these data. Firstly, that the specific activity of newly synthesised toxin is modulated by HlyC in a growth phase dependent manner prior to secretion, or secondly that the activity of the toxin can be both lowered or raised in some unknown manner once secreted. Presumably, the second peak of haemolytic activity in the wild-type strain (LE2001) is not detected since the upstream regulatory mechanism ensures that HlyC and HlyA synthesis is switched off in late exponential phase from the major *hly* promoter.

Turning now to the mutagenesis experiment a possible role for the predicted strong alpha-helix at the C-terminus of HlyA, as an essential

element in HlyA targeting, was tested by the substitution of a central Ser^{996} residue for a Pro. Proline is a unique amino acid in that the side chain group is covalently bonded to the nitrogen atom of the peptide group, and thereby has no amide hydrogen available for participation in hydrogen bonding. This cyclic five membered ring also imposes rigid constraints on rotation about the C-N bond of the peptide backbone. This inability to form hydrogen bonds and the rigid constraints of the ring structure of proline residues make them incompatible with alpha-helix formation (Creighton, 1984). The initial results demonstrated that the 23KD HlyA peptide carrying the Pro^{996} substitution was still secreted efficiently. Moreover, when the mutation was transferred into the wild-type HlyA molecule, the presence of the Pro residue did not apparently alter the efficiency of secretion of the toxin. Thus, these data indicate that this putative alpha-helix *per se*, is not essential for recognition by the export proteins or for maximal secretion efficiencies. However, examination by Chou-Fasman secondary structural predictions suggested that the introduction of this Pro residue did not abolish the potential for an alpha-helical structure in this region. Recent access to a suite of eight secondary structure prediction programs available at Daresbury (see Materials and Methods) suggest all the residues, except the single Pro, encompassing this putative alpha-helical structure are indeed still predicted, by more that 5 of the 8 programmes, to form a helical structure. Thus, although the introduction of a Pro residue should clearly disrupt or kink the helical structure it is possible that the remaining helical structure is still present, permitting the HlyA molecule to be recognised and secreted. We may nevertheless conclude that the presence of an undisrupted, alpha-helix is not a prerequisite for targeting to or efficient secretion by the HlyB,D complex.

Analysis of a role for the apparently conserved "aspartate box" in HlyA recognition by the export proteins was more informative in that an effect of an amino acid substitution upon secretion was observed. The conversion of the Asp^{1009} to Arg (from a small negative residue to a large positively charged residue) reduced the amount of total secreted haemolytic activity to about one third of the wild-type. Analysis of the half-life of this mutant toxin compared to the wild-type revealed no significant difference in decay pattern although the activities again differed by ~70%, suggesting that the substitution is truely altering the ability of the toxin to be recognised/secreted. However, the conversion of Glu^{994} to Lys had no significant effect on the level of secretion and thus

implies that the negative character of the residue at this position is not critical for targeting.

Attempts to alter the nature of the small string of residues separating the flanking "aspartate box" residues by introducing a large Trp residue was unsuccessful. This could have been a direct consequence of the erroneously designed synthetic oligonucleotide for SDM. This encoded the required adjacent two base change in the middle of a 21mer oligonucleotide but in addition also encoded a third mismatch, four bases from one end of the oligonucleotide. This might have prevented the efficient and correct binding of this oligonucleotide to the homologous sequence due to too many mismatches. Alternatively, although this appears more unlikely, the resulting mutant HlyA toxin may have been lethal to the cell and thus could not be recovered.

From these results the Aspartate Box as defined, appears now less likely to play any significant role (also see below), given that the replacement of the Glu^{994} has no effect on secretion. However, the effect of introducing additional large residues remains to be tested. Although the Asp^{1009} residue appears to play some role in the secretion of HlyA, its role is not yet clear, and the nature of the residue (charge/size/property) that is essential at this position also remains to be established. However, comparison of the HlyA C-terminal amino acid sequence with the related *Proteus* and *Morganella* species (not available at the time of investigation) shows that this 3' Asp residue is not in fact conserved, diminishing the possible role of this feature as defined in secretion. In these two strains the Asp residue is replaced by Gly and Asn respectively (Koronakis *et al.*, 1989) conserving the small polar nature of the Asp residue, suggesting that size might be the important characteristic of this residue required for efficient targeting.

Although probing the role of putative structures possibly involved in secretion by site directed mutagenesis has been instructive, this approach is intrincally slow and not always informative. Therefore, it was decided to try another approach, that of random mutagenesis in order to probe the entire region of the targeting sequence. Hopefully, this strategy would generate a bank of secretory mutants, the phenotypes and identity of which would help us to more precisely locate the residues essential for secretion and perhaps permit us to define the nature of the targeting signal.

Figure 1:
Investigation of the location of the HlyA signal sequence by fusion and deletion analysis:
a) Shows the primary structure of the C-terminal 53 residues of the HlyA LE2001 molecule together with deletions generated within this region. The level of secretion or inability/ability of such deletions mutants to be secreted is given on the right hand side.
1) Deletion of C-terminal 27 residues from Ser^{996}, replaced with 7 residues from the tetracycline gene during the construction generating pLG574 (Gray *et al.*, 1986).
2,3) Internal 23KD HlyA deletions generated in an OmpF-23KDHlyA hybrid (Mackman *et al.*, 1987). Line 2 represents the deletion from Ser^{996} to His^{906} within the HlyA domain, while line 3 depicts a larger HlyA deletion from Ser^{996} to His^{857}.
4,5) C-terminal deletions of HlyA (encoded by the pHLY152 determinant) constructed by the generation of stop codons by site directed mutagenesis (Koronakis *et al.*, 1989).

b) Diagrammatical representation of some of the HlyA peptides and fusion peptides generated to date (this study Chapters 3 and 4; Nicaud *et al.*,1985a; Mackman *et al.*, 1987) involving various portions of the HlyA, C-terminal domain, plus their ability/inability to be secreted by HlyB,D.
Filled boxes represent the C-terminal 23KD portion of the HlyA molecule, with the number of amino acids indicated at right.
Hatched boxes represent various heterologous polypeptides OmpF (Mackman *et al.*, 1987) or Pcm fused to different lengths of the HlyA, C-terminal domain (this study Chapters 3 and 4).

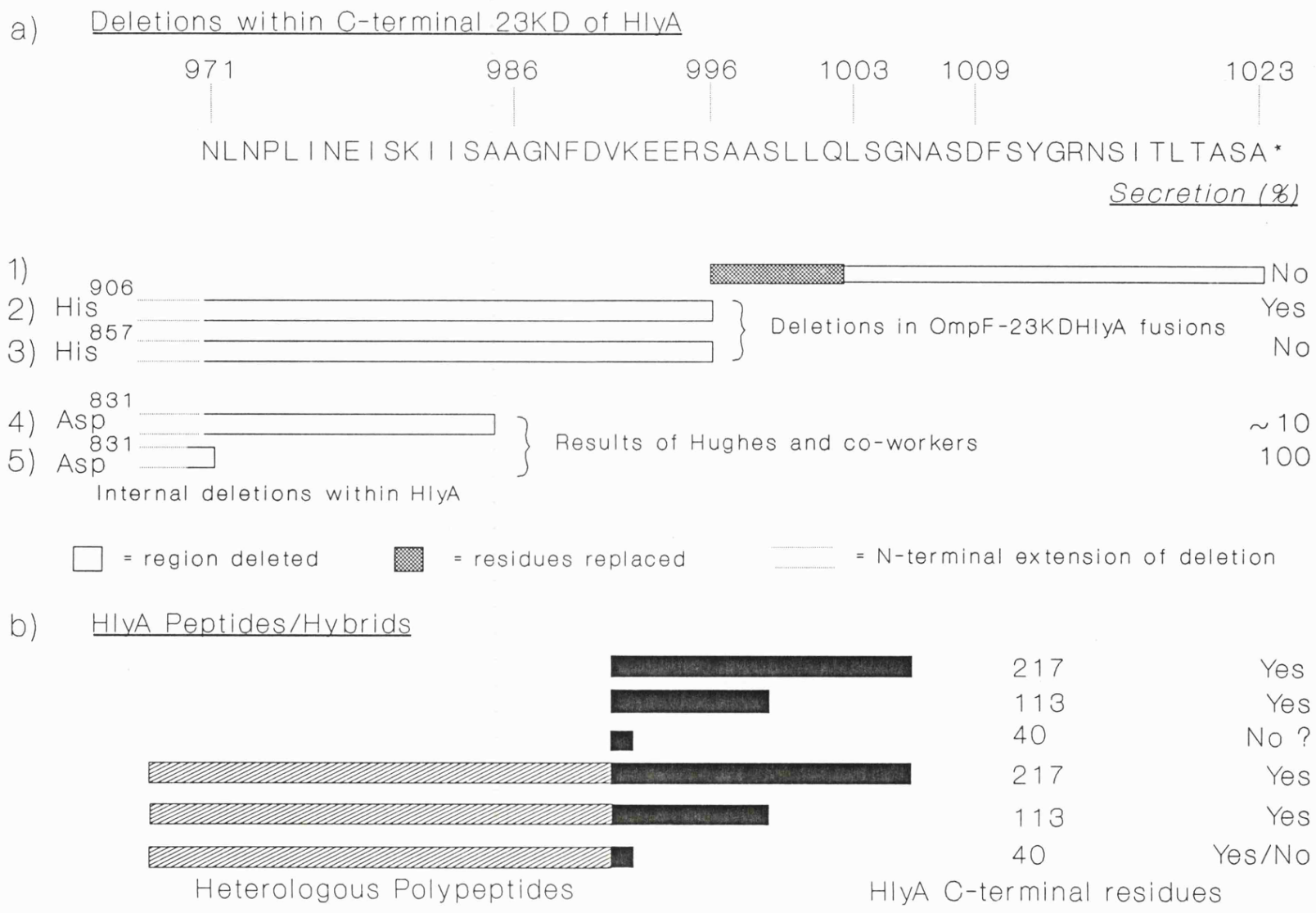

a) Deletions within C-terminal 23KD of HlyA
971
986
996
1003
1009
1023
NLNPLINEISKIISAAGNFDVKEERSAASLLQLSGNASDFSYGRNSITLTASA*
Secretion (%)
1)
No
906
2) His
Yes
857
3) His
No
Deletions in OmpF-23KDHlyA fusions
831
4) Asp
~10
831
5) Asp
100
Results of Hughes and co-workers
Internal deletions within HlyA
= region deleted
= residues replaced
= N-terminal extension of deletion
b) HlyA Peptides/Hybrids
217
Yes
113
Yes
40
No ?
217
Yes
113
Yes
40
Yes/No
Heterologous Polypeptides
HlyA C-terminal residues

Figure 2:

Predicted primary and secondary structural features of the C-terminal 53 residues of HlyA.

The primary sequence of the C-terminal 53 residues of the HlyA 2001 molecule from Asn^{971} to Ala^{1023} is shown with the predicted amphiphilic helix (Koronakis *et al.*, 1989) and "aspartate box" features also indicated. The Chou and Fasman (1978) secondary structure predictions are shown below, with t, h and b indicating predicted turn, alpha helix and beta sheet structures, respectively. The capital letters denote a strong potential for this feature with small letters a weaker potential.

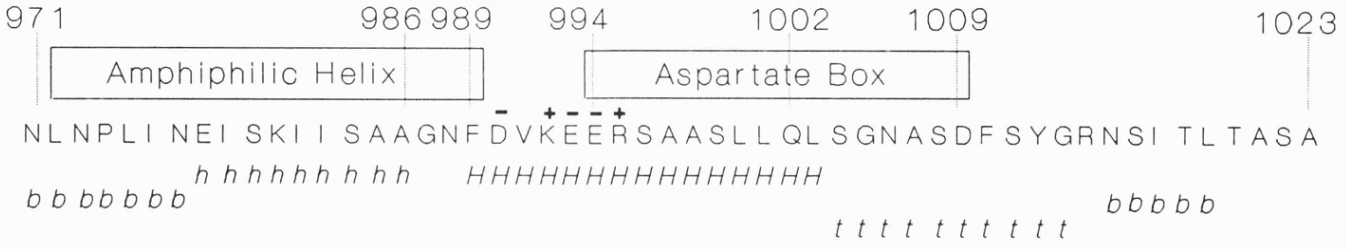

971
986
989
994
1002
1009
1023
Amphiphilic Helix
Aspartate Box
- + - +
NLNPLI NEI SKI I SAAGNF DVKEERSAASLL QL SGNASDF SYGRNSI TL TASA
h hhhhh h hh
HHHHHHHHHHHHHHH
b b bbbbb
bbbbb
t t t t t t t t t t

Figure 3:

Secondary structure predictions of the final 53 residues of molecules carrying C-terminal targeting sequence.

Each protein sequence was run through a suite of eight secondary structure prediction programmes, available on the vax computer system of Leeds university (see Materials and Methods) and the joint prediction for each sequence is given.

T, H and B indicating predicted turn, alpha helix and beta sheet by at least 5 out of the 8 prediction programmes used.

t, h and b indicating predicted turn, alpha helix and beta sheet by less than 5 out of the 8 prediction programmes.

HlyA: Haemolysins from *Escherichia coli* LE2001 (Gray *et al.*, 1986 and sequence this study), *Proteus vulgaris* and *Morganella morganii* (Koronakis *et al.*, 1989)

CyaA: Adenylate cyclase from *Bordetella pertussis* (Glaser *et al.*, 1988)

LtkA: Leukotoxin from *Pasteurella haemolytica* A1 (Lo et al., 1987)

PrtB: Protease from *Erwinia chrysanthemi* (Delepelaire and Wandersman, 1989)

Protease: Protease gene from *Serratia marcescens* (Nakahama *et al.*, 1986)

NodO: Nodulation gene from *Rhizobium leguminosarum* (Economou *et al.*, 1990)

PGK: Phosphoglycerate kinase of *Crithidia fasciculata* and *Trypanosoma brucei* (Swinkels *et al.*, 1988).

Broken lines indicate position of predicted amphiphilic helix.

Solid lines indicate position of "aspartate box" feature.

Charged residues are also indicated.

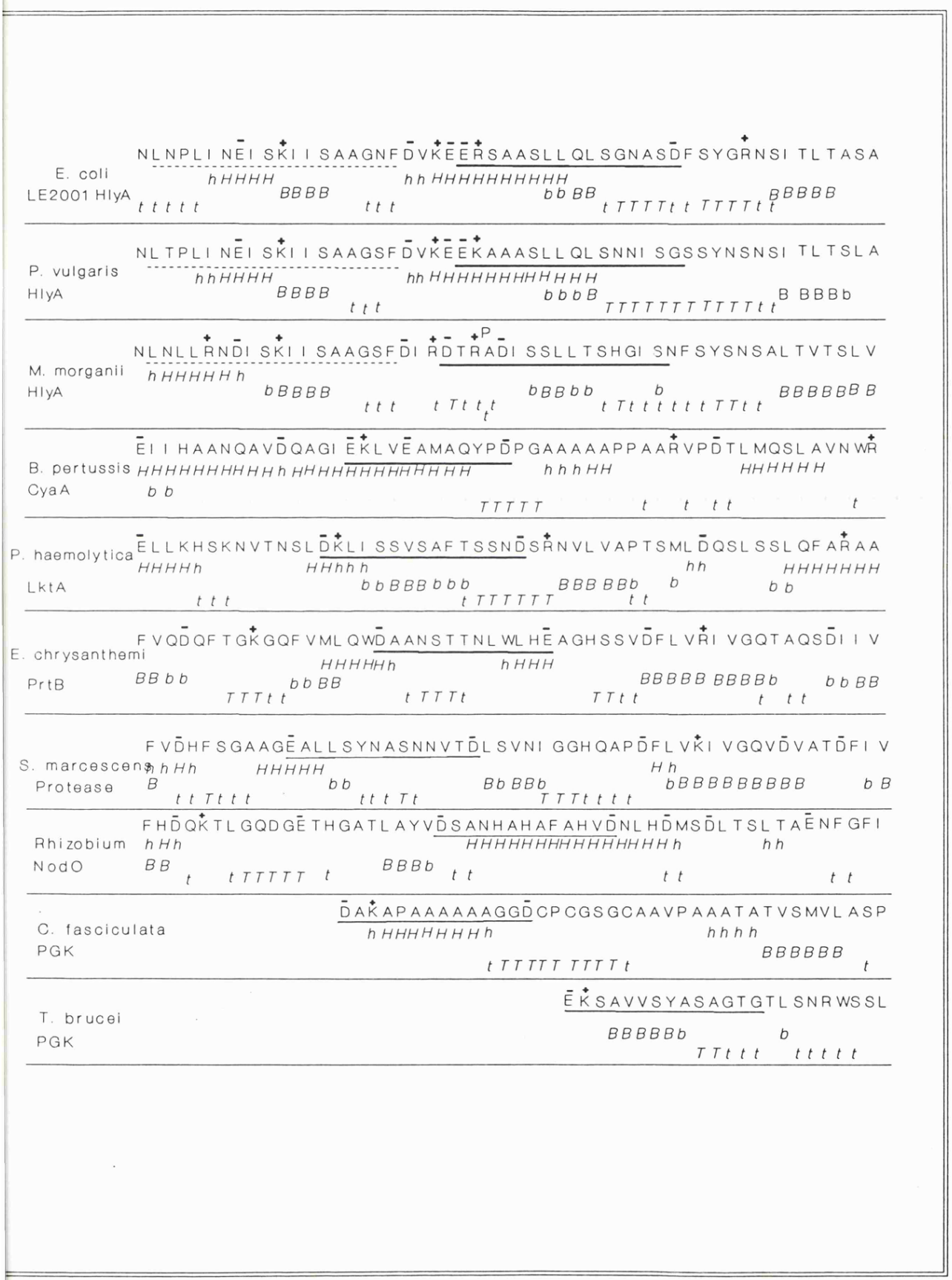

E. coli LE2001 HlyA
NLNPLI NEI SKI I SAAGNF DVKEERSAASLLQLSGNASDFSYGRNSI TLTASA
P. vulgaris HlyA
NLTPLI NEI SKI I SAAGSF DVKEEKAAASLLQLSNNI SGSSYNSNSI TLTSLA
M. morganii HlyA
NLNLLRNDI SKI I SAAGSF DI RDTRADI SSLLTSHGI SNF SYSNSALTVTSLV
B. pertussis CyaA
EI I HAANQAVDQAGI EKLVEAMAQYPDPGAAAAAPPAARVPDTLMQSLAVNWR
P. haemolytica LktA
ELLKHSKNVTNSLDKLI SSVSAFTSSNDSRNVLVAPTSMLDQSLSSLQFARAA
E. chrysanthemi PrtB
FVQDQFTGKGQFVMLQWDAANSTTNLWLHEAGHSSVDFLVRI VGQTAQSDI I V
S. marcescens Protease
FVDHFSGAAGEALLSYNASNNVTDLSVNI GGHQAPDFLVKI VGQVDVATDFI V
Rhizobium NodO
FHDQKTLGQDGETHGATLAYVDSANHAHAFAHVDNLHDMSDLTSLTAENFGFI
C. fasciculata PGK
DAKAPAAAAAAGGDCPCGSGCAAVPAAATATVSMVLASP
T. brucei PGK
EKSAVVSYASAGTGTLSNRWSSL

Figure 4:

Secretion of the 23KDHlyA peptide encoding Ser^{996} to Pro substitution.

E. coli JM101 containing pLG617 (*pUC-23KDhlyA*) or pLG609 (pTTQ-23KD*hlyA*) +/- pLG575 (*hlyB,D*) were grown in nutrient broth (with appropriate antibiotics) to an A^{450} of 0.4 before inducing expression with IPTG (0.5mM final concentration). Samples were taken before and 2 hours post-induction. Cell and culture supernatant samples were prepared as described in Materials and Methods. 0.1 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were analysed by mini SDS-PAGE (15%) and visualised by Coomassie blue.

Cellular samples:

Lane 1,2 time = 0, and 2 hours, respectively from the pLG609 culture **minus** pLG575 (*hlyB,D*)

Lane 3,4 time = 0, and 2 hours, respectively from the pLG609 culture **plus** pLG575 (*hlyB,D*)

Lane 5,6 time = 0, and 2 hours, respectively from the pLG617 culture **minus** pLG575 (*hlyB,D*)

Lane 7,8 time = 0, and 2 hours, respectively from the pLG617 culture **plus** pLG575 (*hlyB,D*)

Supernatant samples:

Lane 9,10 time = 0, and 2 hours, respectively from the pLG609 culture **minus** pLG575 (*hlyB,D*)

Lane 11,12 time = 0, and 2 hours, respectively from the pLG609 culture **plus** pLG575 (*hlyB,D*)

Lane 13,14 time = 0, and 2 hours, respectively from the pLG617 culture **minus** pLG575 (*hlyB,D*)

Lane 15,16 time = 0, and 2 hours, respectively from the pLG617 culture **plus** pLG575 (*hlyB,D*)

Arrows marks the position of the 23KD HlyA peptide, with molecular weight markers also indicated.

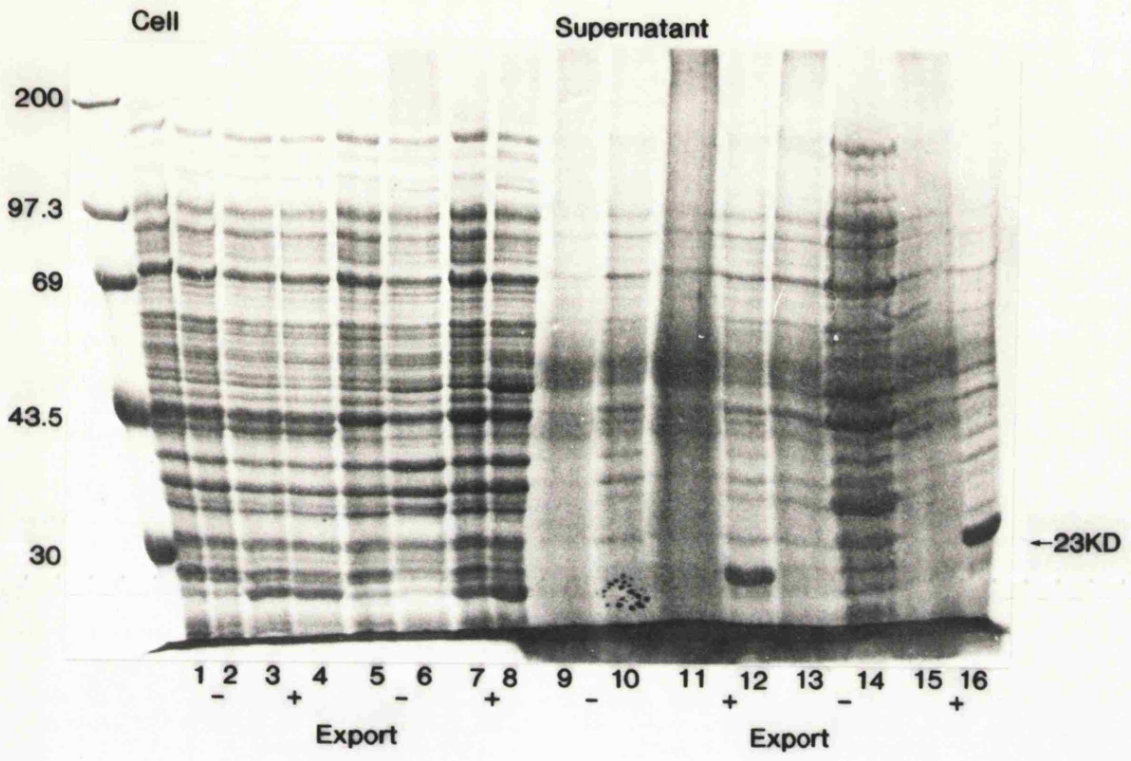

Cell
Supernatant
200
97.3
69
43.5
30
←23KD
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
- + - + - + - +
Export
Export

Figure 5:

Haemolytic assay of cultures secreting wild-type and Ser^{996} substituted (Pro) HlyA molecules.

The *E. coli* strain NM522 carrying pLG339B,D (*hlyB,D*) was transformed with either pACYC-CA,ClaI (*hlyC,A*) or pACYC-CA,Pro (*hlyC, hlyA*Ser^{996} to Pro). Cultures were started in nutrient broth (supplemented with 10mM $CaCl_2$ and appropriate antibiotics) at an A^{450} of 0.1 and samples removed at 40-60 minute intervals for optical density and haemolytic assay reading (see Materials and Methods). The release of haemoglobin by the secreted HlyA toxin was monitored at A^{543} with the haemolytic activity given in arbitrary haemolytic units (H.U). H.U. indicates the A^{543} reading obtained per A^{450} unit of cells.

Dashed lines depict the optical density (A^{450}) while the solid lines show the supernatant haemolytic activity of the cultures.

Stars represent the wild-type HlyA and the open boxes the Ser^{996} to Pro mutant HlyA cultures respectively.

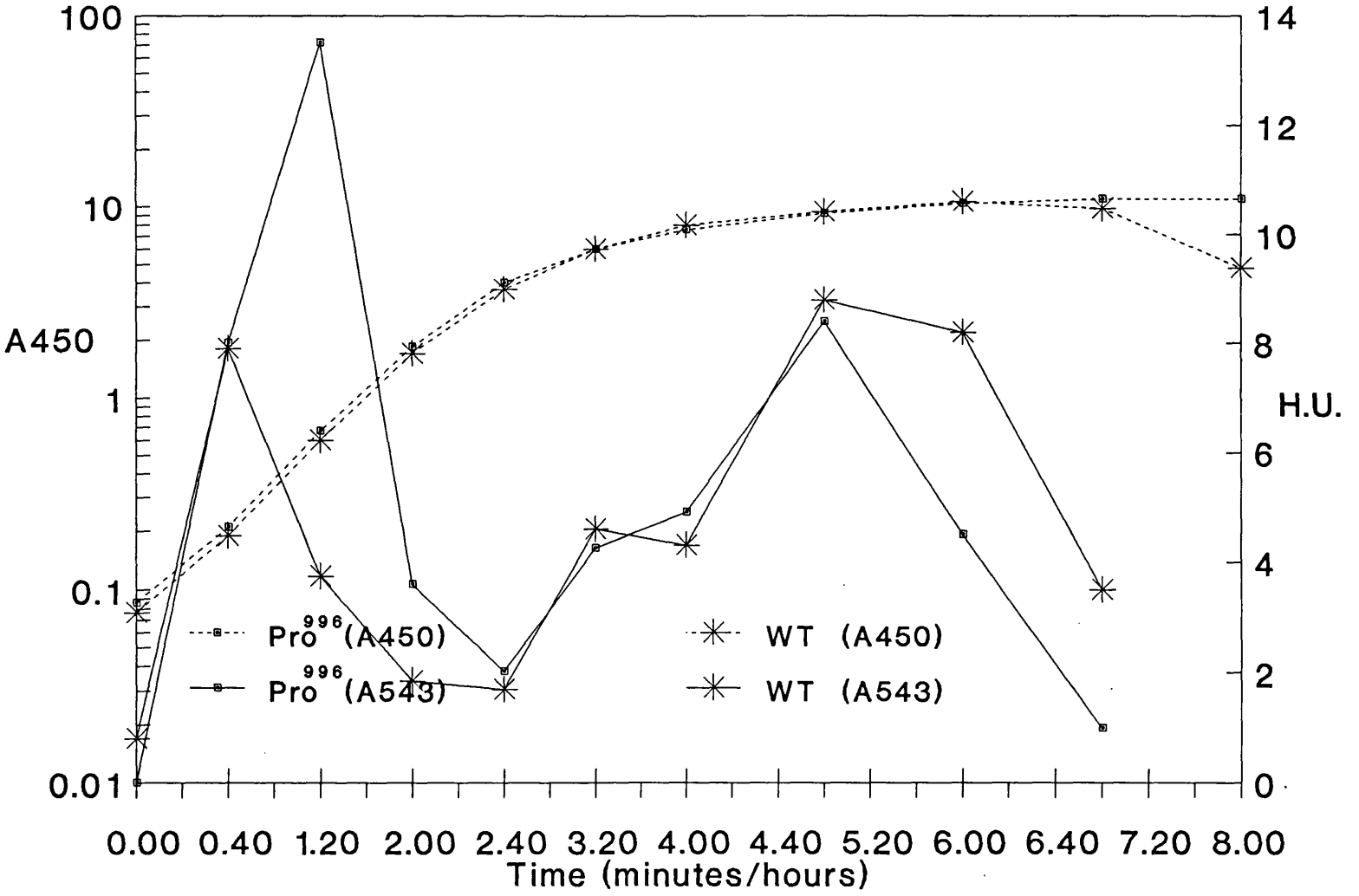
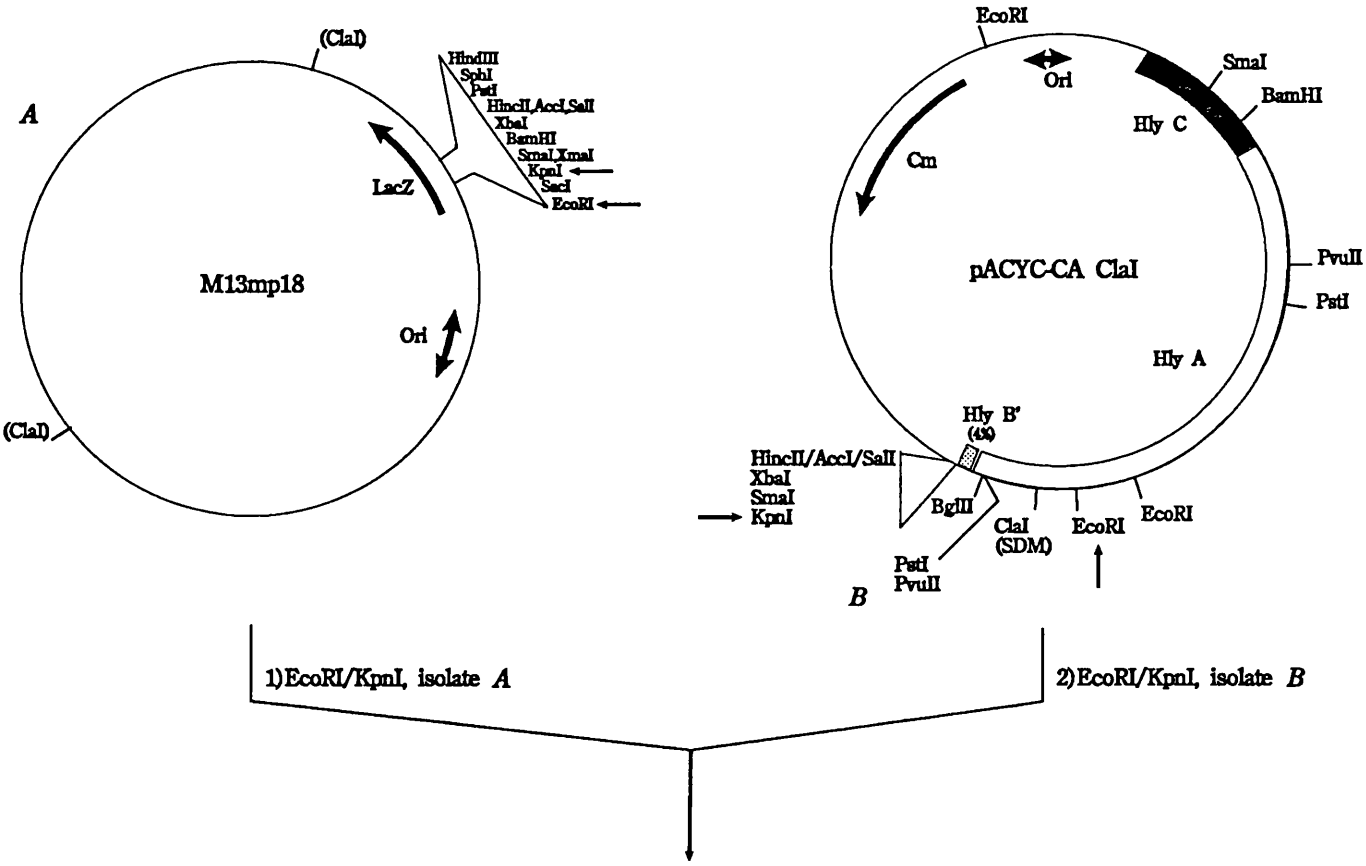

Haemolysin Assay of cultures secreting
WT and Pro996 Mutant HlyA
100
10
1
0.1
0.01
A450
14
12
10
8
6
4
2
0
H.U.
Pro996 (A450)
Pro996 (A543)
WT (A450)
WT (A543)
0.00 0.40 1.20 2.00 2.40 3.20 4.00 4.40 5.20 6.00 6.40 7.20 8.00
Time (minutes/hours)

Figure 6:
Construction of M13mp18-23KDHlyA,Cla
See text for details.
M13mp18: Yanish-Perron *et al.*, 1985
pACYC-CA,ClaI: See Chapter 7: Section 7.5.1


A
(ClaI)
HindIII
SphI
PstI
HincII,AccI,SalI
XbaI
BamHI
SmaI,XmaI
KpnI
SacI
EcoRI
LacZ
M13mp18
Ori
(ClaI)
EcoRI
Ori
SmaI
BamHI
Hly C
Cm
pACYC-CA ClaI
PvuII
PstI
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
BglII
ClaI
(SDM)
EcoRI
EcoRI
PstI
PvuII
B
1)EcoRI/KpnI, isolate A
2)EcoRI/KpnI, isolate B

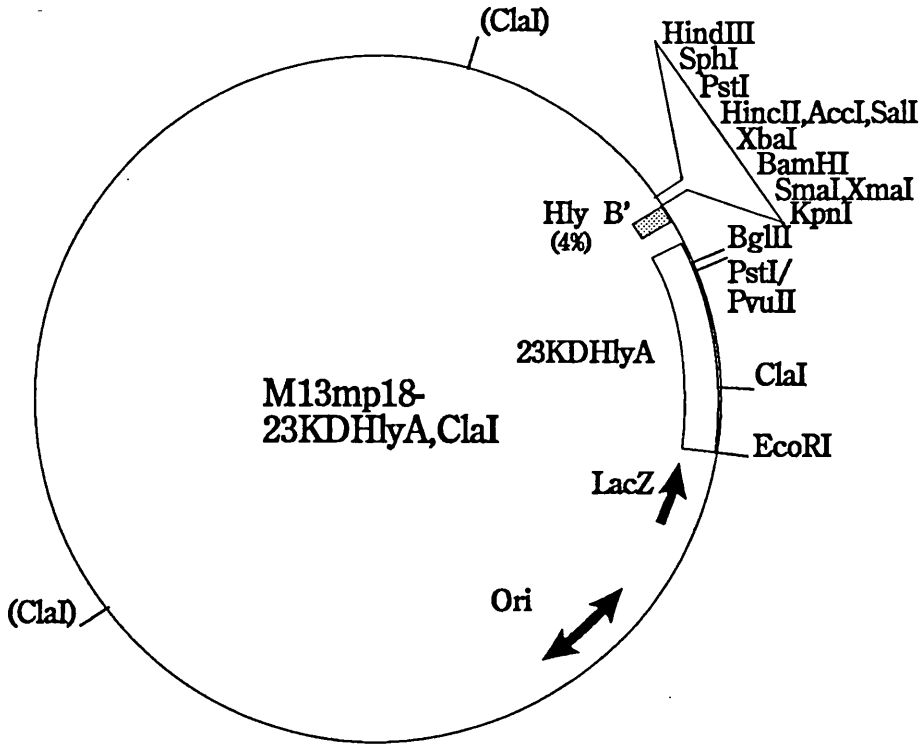

(ClaI)
HindIII
SphI
PstI
HincII,AccI,SalI
XbaI
BamHI
SmaI,XmaI
KpnI
Hly B'
(4%)
BglII
PstI/
PvuII
23KDHlyA
ClaI
M13mp18-
23KDHlyA,ClaI
EcoRI
LacZ
Ori
(ClaI)

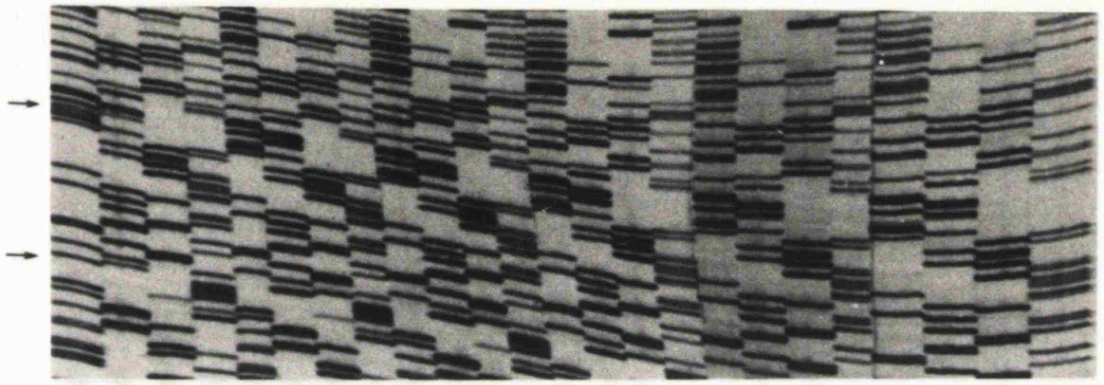
Figure 7:

Identification of "Aspartate box" site directed mutants

Three M13mp18-23KDHlyA,ClaI transformants from both site directed mutagenesis reactions (see text for details) were used to generate single stranded DNA which was sequenced using the (USB) Sequenase kit. An appropriate oligonucleotide was used to allow sequencing of the region targeted in the mutagenesis reaction.

Lanes 1-3 represent the sequence results of the three transformants from the first mutagenesis reaction with 4-6 from the second reaction.

Arrows indicate the specific site directed base changes.



T C G A T C G A T C G A T C G A T C G A T C G A

1 2 3 4 5 6

Figure 8:

Transfering mutations introduced into the 3' end of *hlyA* into the intact *hlyA* gene.

See text for details.

M13mp18-23KDHlyA,ClaI SDM mutant: Same as M13mp18-23KDHlyA,ClaI (see Figure 6) except that specific, mutations have been introduced into the 3' end of *hlyA* by a site directed mutagenesis reaction (See Materials and Methods).

pACYC-CA,ClaI: See Chapter 7: Section 7.5.1

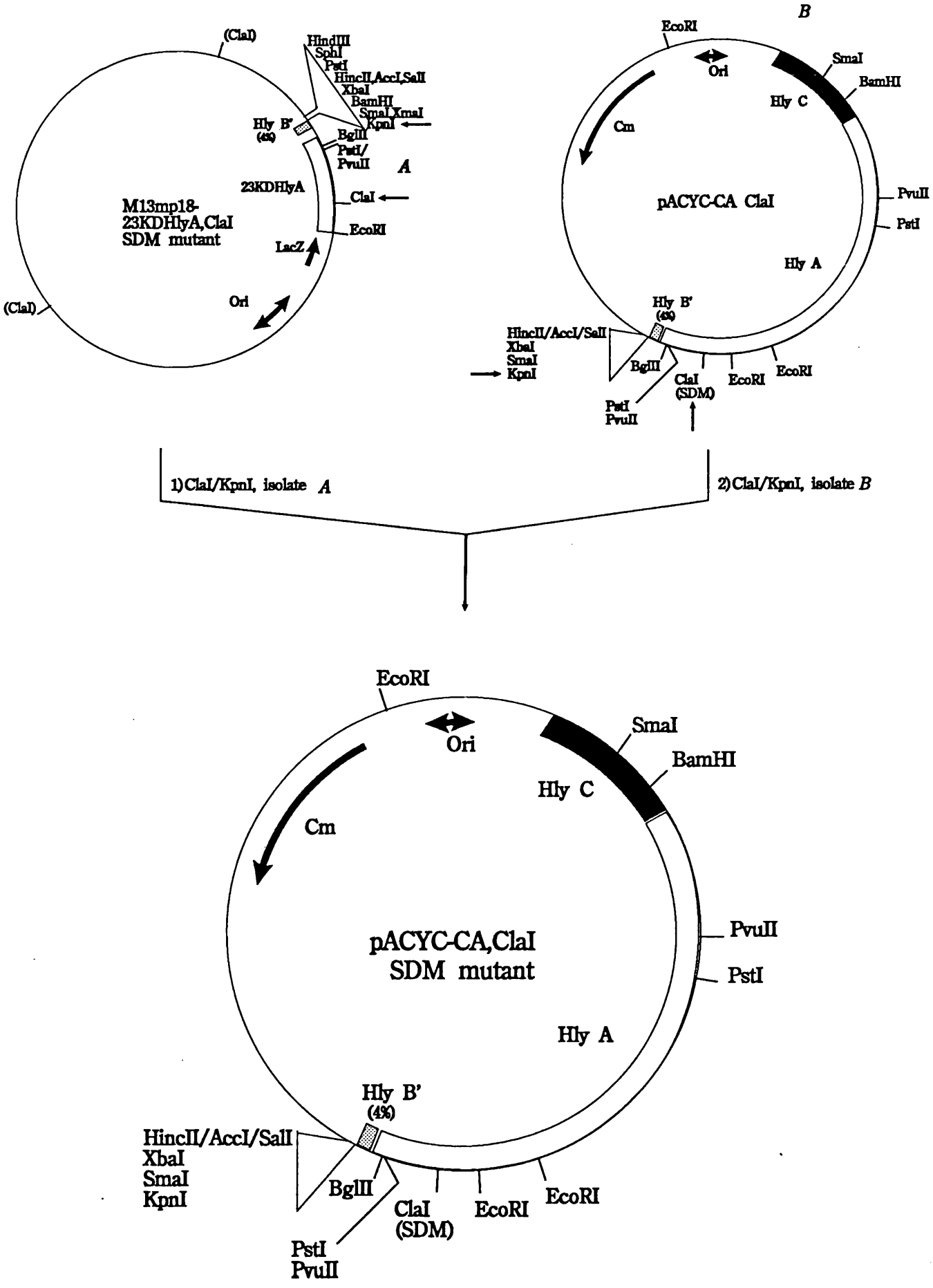

(ClaI)
HindIII
SphI
PstI
HincII,AccI,SalI
XbaI
BamHI
SmaI,XmaI
KpnI
Hly B'
(4%)
BglII
PstI/
PvuII
A
23KDHlyA
ClaI
EcoRI
M13mp18-
23KDHlyA,ClaI
SDM mutant
LacZ
Ori
(ClaI)
B
EcoRI
Ori
SmaI
BamHI
Hly C
Cm
pACYC-CA ClaI
PvuII
PstI
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
BglII
ClaI
(SDM)
EcoRI
EcoRI
PstI
PvuII
1)ClaI/KpnI, isolate A
2)ClaI/KpnI, isolate B
EcoRI
Ori
SmaI
BamHI
Hly C
Cm
pACYC-CA,ClaI
SDM mutant
PvuII
PstI
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
BglII
ClaI
(SDM)
EcoRI
EcoRI
PstI
PvuII

Figure 9:

Haemolytic assay of cultures secreting wild-type and Aspartate box mutant HlyA molecules

The *E. coli* strain NM522 carrying pLG339B,D (*hlyB,D*) was transformed with either pACYC-CA,ClaI (*hlyC,A*), pACYC-CA,Glu^{994} (*hlyC, hlyA*Glu^{994} to Lys) or pACYC-CA,Asp^{1009} (*hlyC, hlyA*Asp^{1009} to Arg). Cultures were started in nutrient broth (supplemented with 10mM $CaCl_2$ and appropriate antibiotics) at an A^{450} of 0.1 and samples removed at 40-60 minute intervals for optical density and haemolytic assay reading (see Materials and Methods). The secreted HlyA activity was determined by monitoring (at A^{543}) the release of haemoglobin from red blood cells (see Materials and Methods). The haemolytic activity is given in arbitrary haemolytic units (H.U). H.U. indicates the A^{543} reading obtained per A^{450} unit of cells.

Dashed lines depict the optical density (A^{450}) while the solid lines show the supernatant haemolytic activity of the cultures.

The stars represent the wild-type HlyA, open boxes the Glu^{994} to Lys and the diamonds, Asp^{1009} to Arg mutant cultures respectively.

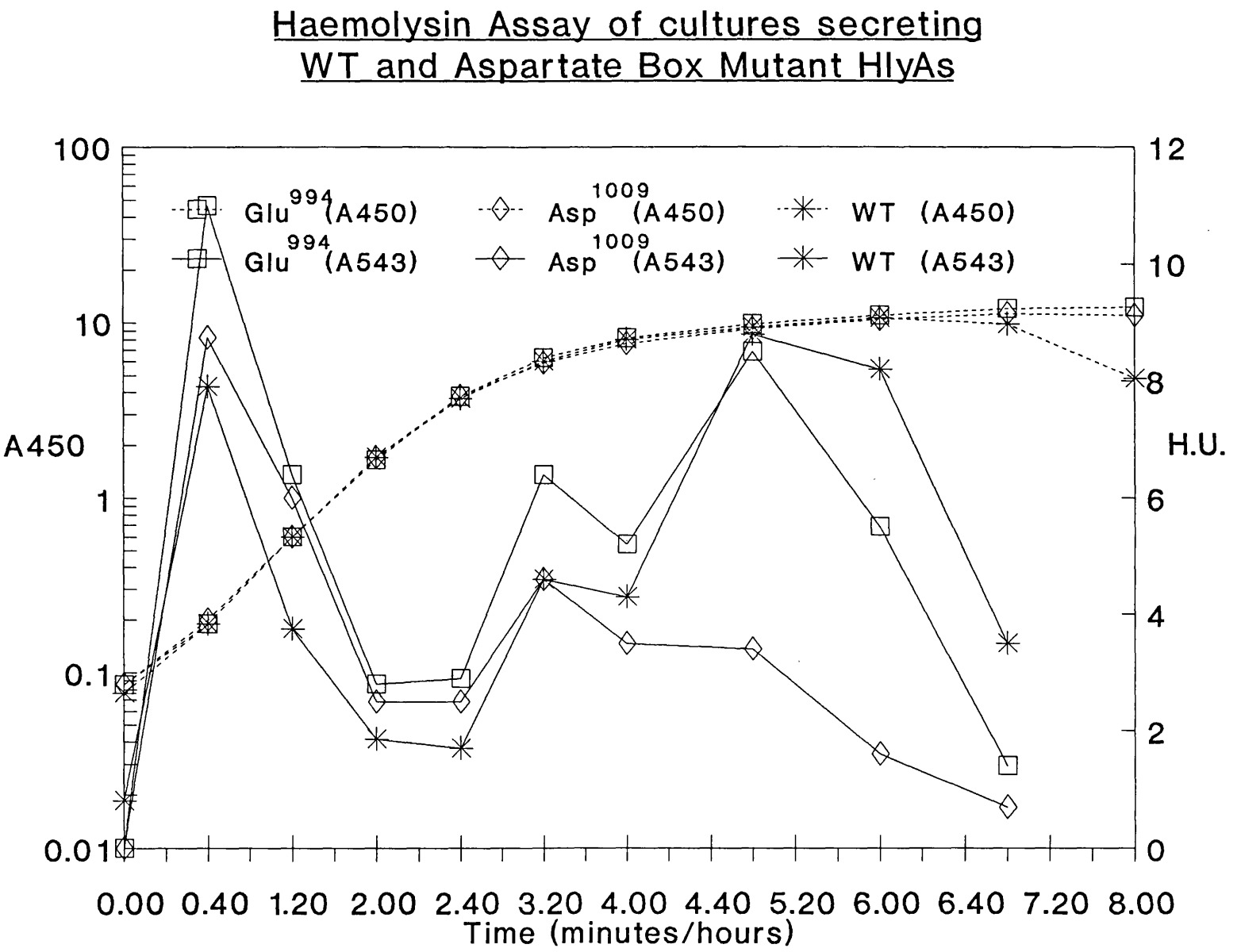

Haemolysin Assay of cultures secreting
WT and Aspartate Box Mutant HlyAs
Glu 994 (A450)
Asp 1009 (A450)
WT (A450)
Glu 994 (A543)
Asp 1009 (A543)
WT (A543)
A450
H.U.
100
10
1
0.1
0.01
12
10
8
6
4
2
0
0.00 0.40 1.20 2.00 2.40 3.20 4.00 4.40 5.20 6.00 6.40 7.20 8.00
Time (minutes/hours)

Figure 10:

Activity Half life of wild-type and the Asp^{1009} to Arg mutant of HlyA.

The *E. coli* strain NM522 carrying pLG339B,D (*hlyB,D*) pACYC-CA,ClaI (*hlyC,A*) and pACYC-CA,Asp^{1009} (*hlyC, hlyA*Asp^{1009} to Arg) were grown in nutrient broth (supplemented with 10mM $CaCl_2$ and appropriate antibiotics) to an A^{450} of 4.5 before removing a 3ml sample from each. The cells were spun out and 1.5mls of supernatant transferred to a new eppendorf tube. Sodium azide (5mM final concentration) was added to prevent the growth of any remaining cells. The supernatant samples were placed at 37°C and aliquots taken at the point of incubation and every subsequent 15 minute interval for 105 minutes and haemolytic activity determined, as before, by monitoring (at A^{543}) the release of haemoglobin from red blood cells (see Materials and Methods). The haemolytic activity is given in arbitrary haemolytic units (H.U). H.U. indicates the A^{543} reading obtained per A^{450} unit of cells.

The stars represent the wild-type HlyA and the diamonds the Asp^{1009} to Arg mutant haemolytic activity, respectively.

Half-Life of WT versus Asp^{1009} Mutant HlyA

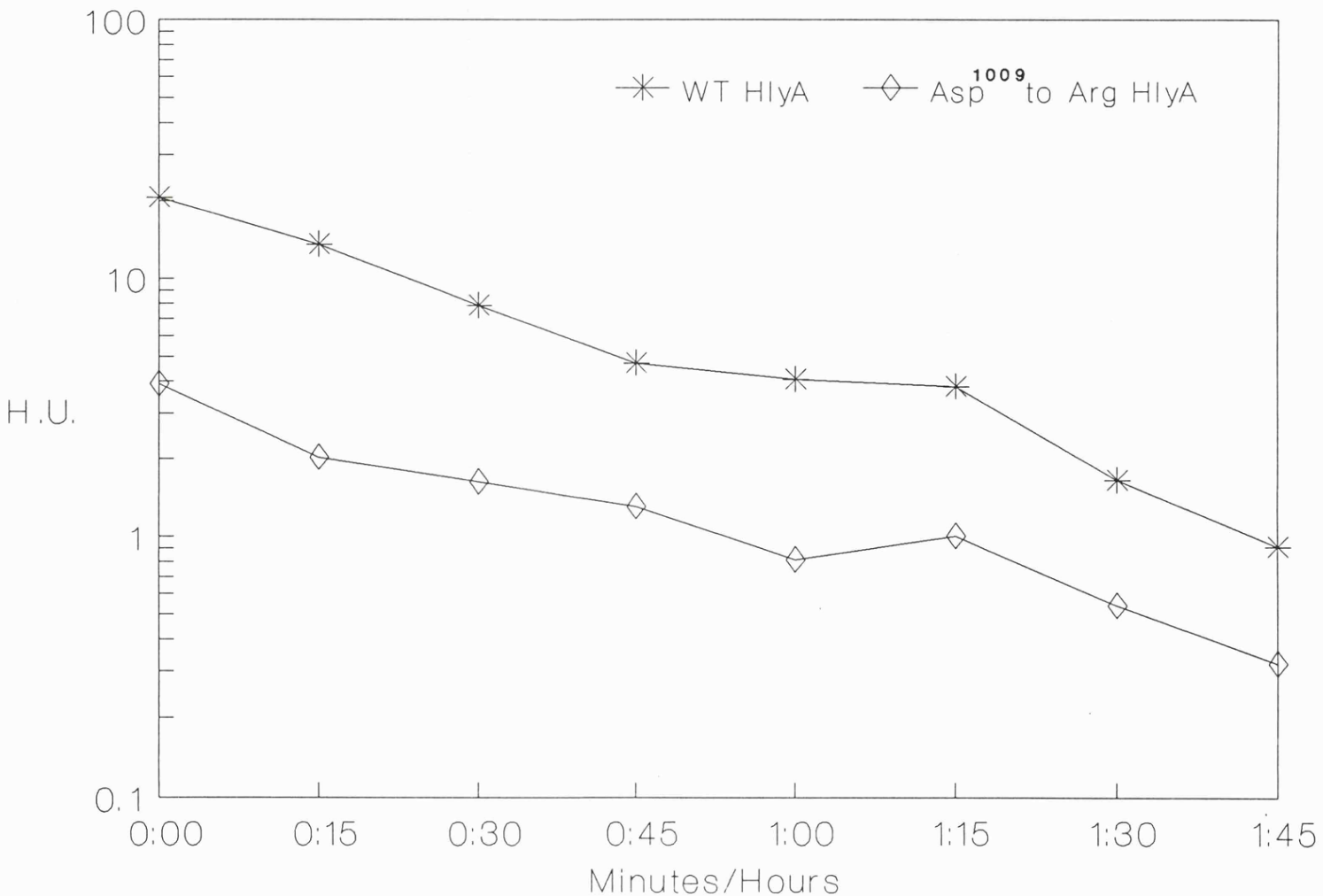


Figure 11:

a) Haemolytic assay of secreted wild-type and C-terminal truncated HlyA molecules.

The *E. coli* strain NM522 carrying pLG339B,D (*hlyB,D*) was transformed with pACYC-CA,ClaI (*hlyC,A*) or pACYC-CA,Trun (*hlyC,hlyA*deleted for final ~447bp). Cultures were started in nutrient broth (supplemented with 2.5mM $CaCl_2$ and appropriate antibiotics) at an A^{450} of 0.2 and samples removed over a 6 hour period at 60 minute intervals for both optical density and haemolytic assay reading (see Materials and Methods). The release of haemoglobin due to the secreted HlyA toxin was monitored at A^{543} with the haemolytic activity given in arbitrary haemolytic units (H.U). H.U. indicates the A^{543} reading obtained per A^{450} unit of cells.

Solid lines depict the optical density (A^{450}) while the dashed lines show the supernatant haemolytic activity of the cultures.

The stars represent the wild-type HlyA and the diamonds the HlyA truncated mutant cultures respectively.

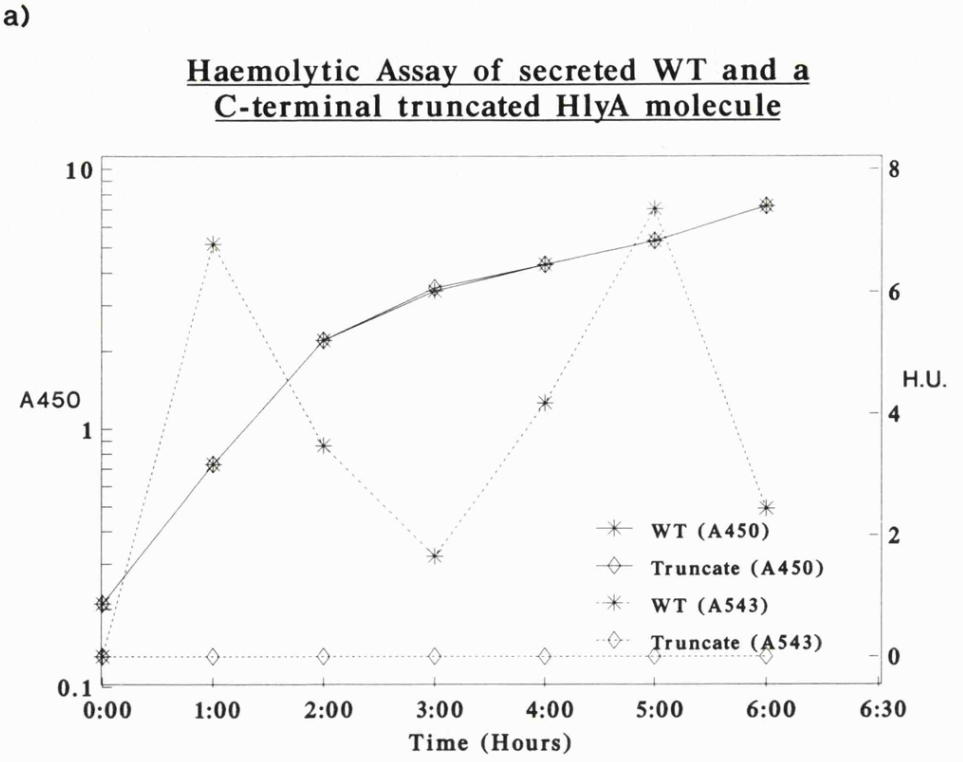
b) Cell and culture supernatant samples were also taken at the same intervals as above and prepared as described in Materials and Methods. 0.1 units of cells (where 1 unit is defined as 1ml at A^{450} = 1) and 1 equivalent units of supernatant sample were loaded. Proteins were separated by mini SDS-PAGE (11%), transferred to nitrocellulose and probed with anti-23KDHlyA antibodies.

Supernatant samples: The top panel shows the supernatant samples taken at 1, 2, 3, 4, 5 and 6h intervals. Each hourly sample shows the profile for the strain expressing the wild-type and truncated HlyA molecules, respectively.

Cellular samples: The lower panel shows the cellular samples taken at 1, 2, 3, 4, 5 and 6h intervals. Each hourly sample shows the profile for the strain expressing the wild-type and truncated HlyA molecules, respectively.

Arrows indicate the position of the wild-type and truncated HlyA molecules, with the molecular weight markers also indicated.

a)


Haemolytic Assay of secreted WT and a C-terminal truncated HlyA molecule
10
1
0.1
A450
8
6
4
2
0
H.U.
WT (A450)
Truncate (A450)
WT (A543)
Truncate (A543)
0:00
1:00
2:00
3:00
4:00
5:00
6:00
6:30
Time (Hours)

b)

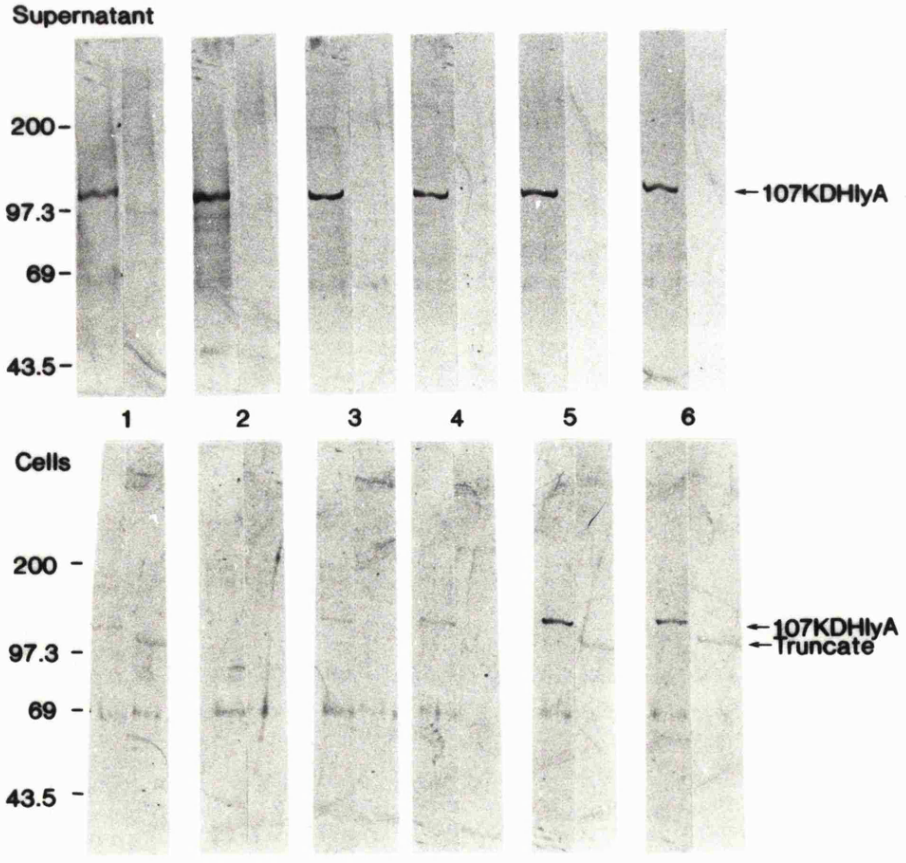

Supernatant
200
97.3
69
43.5
1
2
3
4
5
6
107KDHlyA
Cells
200
97.3
69
43.5
107KDHlyA
Truncate

Figure 12:

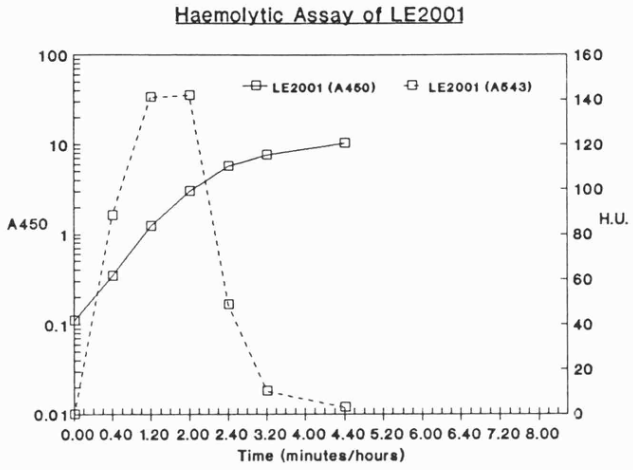
Growth curves and haemolytic activity profiles of a) the wild-type *hly* strain LE2001 versus b) 2 and c) 3 plasmid subcloned *hly* complementation systems.

H.U. indicates the haemolytic activity A^{543} reading obtained per A^{450} unit of cells.

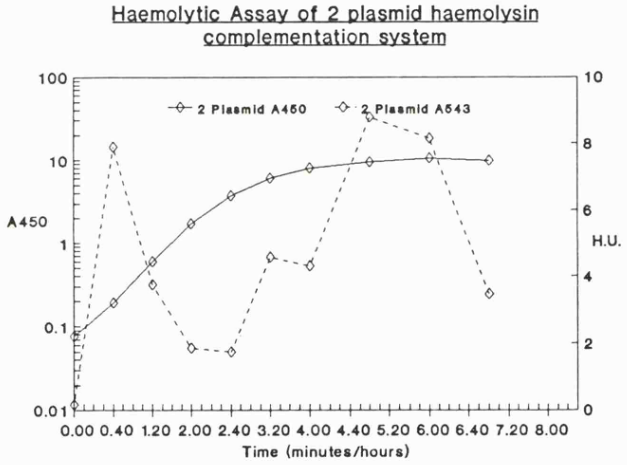
Solid lines represent the growth curve (A^{450}).

Broken lines represent the haemolytic activity (A^{543}).

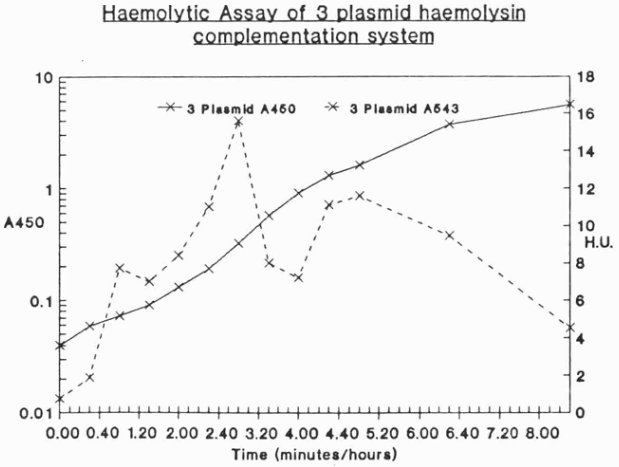
a)


Haemolytic Assay of LE2001
LE2001 (A450)
LE2001 (A543)
A450
H.U.
100
10
1
0.1
0.01
160
140
120
100
80
60
40
20
0
0.00 0.40 1.20 2.00 2.40 3.20 4.00 4.40 5.20 6.00 6.40 7.20 8.00
Time (minutes/hours)

b)


Haemolytic Assay of 2 plasmid haemolysin complementation system
2 Plasmid A450
2 Plasmid A543
A450
H.U.
100
10
1
0.1
0.01
10
8
6
4
2
0
0.00 0.40 1.20 2.00 2.40 3.20 4.00 4.40 5.20 6.00 6.40 7.20 8.00
Time (minutes/hours)

c)


Haemolytic Assay of 3 plasmid haemolysin complementation system
3 Plasmid A450
3 Plasmid A543
A450
H.U.
10
1
0.1
0.01
18
16
14
12
10
8
6
4
2
0
0.00 0.40 1.20 2.00 2.40 3.20 4.00 4.40 5.20 6.00 6.40 7.20 8.00
Time (minutes/hours)

CHAPTER 7

Investigation into the nature of the HlyA C-terminal targeting motif by random hydroxylamine mutagenesis.

7.1 Introduction:

In an attempt to identify specific residues which constitute the HlyA recognition signal, I decided to subject the DNA fragment encoding the signal sequence to random hydroxylamine mutagenesis. The screening for mutants which can no longer lyse red blood cells when complemented with the other *hly* gene products, would hopefully identify those residues essential, either directly or indirectly, for recognition by the *hly* export machinery. The nature and relative position of such residues should also indicate something of the secondary or tertiary structure requirement of the C-terminus for recognition. The experimental approach relies on the observation that some chemical agents interact with DNA, modifying it in such a way as to impair their proper base pairing in a double stranded DNA duplex. Hydroxylamine is one such chemical mutagen and has been used extensively to generate random mutants. Hydroxylamine interacts preferentially with cytosine modifying it so that it now pairs with adenine rather than guanosine. This results in a C:G to T:A transition. The conditions for generating a high rate of single point mutations on plasmid DNA using hydroxylamine has been described by Humphreys *et al.*, (1976). The objective of this experiment was to introduce random mutations, specifically within the 3' end of the *hlyA* gene which encodes the secretion targeting signal (to avoid other classes of haemolytically defective mutants, such as, those altering transcription or toxin function). Therefore, it was necessary to isolate a specific 3' *hlyA* fragment, treat with the mutagen, reconstitute into the intact *hlyA* gene, and screen for *sec*$^-$ mutations in cells expressing the products of the other *hly* genes (HlyC,B,D).

7.2 Generation of a unique restriction site to enable isolation of the 3' end of *hlyA*:

From the restriction map of the *hly* determinant of LE2001 (see Figure 1), which was previously cloned into the low copy number plasmid pOU71 (pLG570; Mackman and Holland 1984b), it is evident that there are few convenient or unique restriction sites available for the manipulation of the *hly* determinant. Indeed, there are no convenient restriction sites permiting the specific removal, and re-insertion, of the 3' portion of *hlyA* which encodes the signal sequence region. A number of unique sites were therefore introduced downstream of the *hlyA* gene by insertion of the *hlyA* gene into the multiple cloning polylinker region of the pUC expression vector (see below). In addition, a unique site was introduced upstream of the signal domain by site directed mutagenesis (as described below).

These new sites allowed the DNA fragment encoding the signal region to be readily isolated for *in vitro* mutagenesis.

7.2.1 Cloning *hlyA* into pUC18/19 expression vectors:

The *hlyA* gene, lacking the major upstream promoter region, was inserted (see below) into the high copy number vectors pUC18 and 19 (i.e in both orientations with respect to the *lac* promoter). This maximised the chances of isolating a stable *hlyA* clone, since there were doubts of the stability of high copy number vectors expressing HlyA. In this case also, due to the lack of unique restriction sites, *hlyA* had to be cloned into pUC in two steps (see Figure 2). First, the 0.8Kb BamHI/HpaI fragment from pLG609-4 (see Chapter 3: Figure 2a/b) encoding the C-terminal 23KD portion of HlyA and the first 72 bp of *hlyB* was ligated into the BamHI/SmaI sites of pUC18 and pUC19 (Norrander *et al.*, 1983). The resulting recombinant plasmids were then digested with BamHI/ BglII and the remainder of the *hlyA* gene introduced on a 3Kb BamHI/BglII fragment (containing the 3' end of *hlyC* and the majority of *hlyA*) isolated from pLG570 (Mackman and Holland, 1984b). The resulting plasmids pLG585 (pUC*hlyA*,incor) and pLG586 (pUC*hlyA*,cor) carry the *hlyA* gene in the incorrect and correct orientation, respectively, with respect to the pUC *lac* promoter (Figure 2). Indeed, difficulties were encountered in isolating the construct with the *hlyA* gene under the control of the *lac* promoter (pLG586), presumably due to lethal effects from basal level expression. These constructs allow the retrieval of the whole *hlyA* gene on a single 3.3Kb BamHI/KpnI fragment. CsCl purified plasmid DNA was prepared and the pLG585 plasmid analysed with a variety of commonly used restriction enzymes in an attempt to determine any further difference between the reported *E. coli hlyA* sequences (Felmlee *et al.*, 1985a; Hess *et al.*, 1986) and the LE2001 *hlyA* determinant of which only the 3' ~700bp have been sequenced (Gray *et al.*, 1986). The results of the restriction digests are shown in Figure 3a. The analysis of the restriction enzyme digest patterns of *hly* 2001 compared with the Felmlee *et al.*, (1985a) sequence identified a) an extra unique HindIII site at the 5' end of the *hlyA* 2001 gene, the absence of b) a HincII/HpaI site and c) a PstI site (see Figure 3b). The results also confirmed the previously observed difference in restriction pattern between LE2001 and pSF4000 for the restriction enzymes EcoRI and BglII (Mackman and Holland, 1984b).

7.2.2 Generation of a unique ClaI site near the 3' end of the *hlyA* gene:

It was decided to use the site directed mutagenesis kit (Amersham Ltd) to generate an appropriate unique restriction site within *hlyA* to allow the removal of the 3' end for *in vitro* manipulation. In order to do this the kit requires the DNA template to be in single stranded form. This required the subcloning of the 3.3Kb BamHI/KpnI *hlyA* fragment from pLG585 (see above) into the bacteriophage M13 expression vector which also permits the isolation of single stranded template. Initial attempts to subclone the *hlyA* fragment in M13mp19 (in the same orientation as the *lac* promoter) resulted in some transformants carrying deletions (data not shown). In contrast, cloning into M13mp18 (in the opposite orientation to the *lac* promoter) showed no signs of instability generating M13mp18-*hlyA* (see Figure 4). This probably reflects low level "leaky" expression of the *hlyA* gene from the *lac* promoter in the M13mp19 clone, the gene product interfering with cell viability, leading to the selection of transformants deleted for the intact *hlyA* gene.

7.2.3 Choice and location of a unique restriction site within *hlyA*:

As stated earlier (see Chapter 6; Section 6.1) all the information constituting the HlyA targeting signal is encoded, at least, within the last 12KD of HlyA, while other upstream sequences (within the last 23KD; 627bp) may be required solely for increasing the efficiency of secretion. Therefore, in order to isolate the *hlyA* region encoding the signal recognition sequence it was necessary to generate a unique restriction site, ideally, between 360 and 627bp from the end of the *hlyA* gene (encoding the final 12 and 23KD of HlyA, respectively). Utilisation of the Winsconsin DNA manipulation package and the programme Map\mismatch, identified a potential unique ClaI site within the above mentioned *hlyA* region, which could be generated by altering only one base. This change, more importantly, does not affect the coding residue or its usage preference and thus should not affect the function/activity of the molecule. This potential site was located 447bp from the end of *hlyA* and therefore would allow, together with the use of the unique KpnI site downstream of the *hlyA* gene in pLG585 and pLG586 (see Figure 2), the isolation of a fragment encoding the final 149 residues of HlyA. The appropriate 21mer oligo was synthesised and used in a site directed mutagenesis reaction together with single stranded DNA from M13mp18-

HlyA (see Figure 4) as described by the manufacturers (see Materials and Methods). Analysis of 11 resulting M13mp18-*hlyA* transformants for the presence of an extra ClaI site, revealed that 3 carried this extra site. The 3.3Kb BamHI(ClaI)/KpnI fragment was isolated from one of these transformants and recloned into pUC18 and pUC19 to generate pLG585-1 (pUC*hlyA*,incor,ClaI) and pLG586-1 (pUC*hlyA*,cor,ClaI) - the *hlyA* gene again in the incorrect or correct orientation with respect to the pUC *lac* promoter, respectively. Unfortunately, pLG586-1 (*hlyA*,cor,Cla) seemed to be somewhat unstable, as some of the transformants carried deletions, again presumably as a result of low level basal expression of HlyA directed from the *lac* promoter. pLG585-1 and pLG586-1 are identical to pLG585 and pLG586 (see Figure 2) except that they carry a unique ClaI site within the *hlyA* gene at the position shown in Figure 6.

7.3 Complementation tests to confirm the nature of *hlyA* constructs:

Before using these HlyA encoding plasmids in the hydroxylamine mutagenesis experiment, it was necessary to check that they still expressed a functional toxin that is secreted and active in the presence of the other Hly proteins (HlyC,B and D). It was not in fact expected that the pLG585-1 (pUC*hlyA*,incor,ClaI) plasmid would express the toxin due to the assumed absence of a known promoter to drive its transcription. For the complementation test appropriate plasmids encoding HlyC,B and D were also required. The pLG585-1 and 586-1 plasmids were both ampicillin resistant while the plasmid encoding HlyB and HlyD, pLG575 (*hlyB,D*; Mackman *et al.*, 1985a), was chloramphenicol resistant. Since *hlyC* was not available on a plasmid with an alternative selectable antibiotic resistance marker, I inserted the tetracycline gene from pBR322 (on a SspI/BalI blunt fragment; Bolivar *et al.*, 1977) into the ScaI site within the ampicillin gene of pLG591 (pUC*hlyC*; Nicaud *et al.*, 1985a) creating pLG591Tc (*hlyC*; construction not shown).

pLG585-1 (pUC*hlyA*,incor,ClaI) and pLG586-1 (pUC*hlyA*,cor,ClaI) were transformed into *E. coli* SE5000 cells (lacks the *lacI*^{q} repressor gene, allowing constitutive expression from the *lac* promoter), which also expressed HlyC,B and D and plated out on blood agar plates. Surprisingly, both plasmids expressed HlyA which was functional and secreted as evidenced by lysis of red blood cells on blood agar plates. Since *hlyA* was expressed in both orientations, relative to the *lac* promoter the results indicated the presence of either a promoter like sequence immediately upstream of the normal *hlyA* gene or within the

sequence immediately upstream of the normal *hlyA* gene or within the vector DNA itself.

7.4 Hydroxylamine mutagenesis reaction:

Since pLG585-1 (pUC*hlyA*,incor,ClaI) expressed the HlyA toxin and showed no signs of instability I decided to use this construct to generate and screen for hydroxylamine secretion (*sec*$^-$) mutants. CsCl prepared plasmid DNA was digested with ClaI/KpnI and the small 600bp fragment, encoding the final 149 residues of HlyA, was isolated by the "death wish" method after separation on a 0.7% agarose gel (see Materials and Methods). Approximately 1µg of this 600bp fragment was exposed to hydroxylamine as described in Materials and Methods. As a control, 1µg of intact plasmid was taken and treated identically, half in the presence or absence of hydroxylamine. This control DNA was used to transform *E. coli* cells and the survival rate estimated to be approximately 20% as a result of exposure to hydroxylamine. Humphreys *et al.*, (1976) states that a survival rate of between 10 to 30% is indicative of successful mutagenesis, with a preponderance of single mutations. This implied that the reaction was successful, but the effect on the 600bp fragment could only be judged by recloning back into an intact *hlyA* and screening for *sec*$^-$ mutants.

7.4.1 Screening for hydroxylamine *sec*$^-$ mutants:

The first ligation to reconstruct the intact *hlyA* gene using the 0.6Kb ClaI/KpnI hydroxylamine treated 3' *hlyA* fragment plus controls consisted of i) the large isolated 5.3Kb ClaI/KpnI vector fragment from pLG585-1 minus the 0.6Kb ClaI/KpnI 3' *hlyA* insert, ii) large ClaI/KpnI vector fragment plus untreated 0.6Kb ClaI/KpnI 3' *hlyA* insert and, lastly, iii) the large ClaI/KpnI vector fragment plus the hydroxylamine treated 0.6Kb ClaI/KpnI 3' *hlyA* fragment. Each ligation contained approximately 100ng of the 5.3Kb vector fragment, the 0.6Kb insert (when present) in a roughly 1:1 molar ratio of ends. Transformation of one-fifth of the products of these ligations into *E.coli* SE5000 pLG575 (*hlyB,D*) / pLG591Tc (*hlyC*), in all cases, produced fewer transformants than expected, considering a transformation efficiency of 1 X 10^7 was achieved, as determined by transforming 1ng of pUC19 as a control. Transformation with the ligation mixture of the vector 5.3Kb fragment and the **untreated** 0.6Kb fragment (ii above) surprisingly generated 1 *sec*$^-$ colony out of a total of 19 transformants (~5%), whereas in iii) the hydroxylamine **treated** 0.6Kb insert only produced 3 *sec*$^-$ colonies out of 170 transformants (~1.7%). Restriction analysis of a single *sec*$^-$

transformant from each source of the reaction mixtures, indicated above, demonstrated that all carried some form of deletion of the pLG585-1 (pUC*hlyA*,incor,ClaI) plasmid (data not shown). These results indicated two problems, a low ligation efficiency and a background of *sec*$^-$ deletions higher than the expected rate due to hydroxylamine mutagenesis.

In a repeat experiment using the hydroxylamine **treated** ClaI/KpnI 0.6Kb fragment, the total quantity of DNA in the ligation mixture was increased by 8 fold, this time with an approximate 1:1.3 ratio of vector to insert. The transformation of one fifth of the resulting ligation resulted in only 108 transformants, 4 (3.7%) of which were *sec*$^-$. Analysis of these *sec*$^-$ clones by restriction digestion indicated that one transformant appeared to lack only the 0.6Kb insert (possibly an aberrant ligation product), while another two had larger deletions (resulting in the loss of both the ClaI and KpnI restriction sites; data not shown) and the fourth was a larger than expected plasmid, produced an unrecognisable pattern (see Figure 5). It was plausible that this high background level of *sec*$^-$ deletion mutants might well be due to the complementation system being used. This consisted of 3 high copy number plasmids (see Section 7.3) all of which, unfortunately, contain overlapping regions of homology covering the *hly* DNA. Therefore, if the expression of *hlyA* is deleterious under these conditions a strong selection pressure would generate *hlyA*$^-$ mutants via some form of recombination. In an attempt to avoid these potential problems, I decided to generate a 2 plasmid complementation system, on lower copy number plasmids, reducing also the degree of *hly* homology between the plasmids. It was not possible at this stage to generate a single low copy number plasmid carrying the entire *hly* determinant and also maintain the unique restriction sites such as the ClaI site generated earlier.

7.5 Generation of a lower copy number 2 plasmid Hly complementation system:

pACYC184 is a Cm^R / Tc^R, moderately low copy number plasmid (~18 copies per cell; Chang and Cohen, 1978). In order to enable the *hlyA* gene to be cloned into this vector, retaining the unique ClaI/KpnI sites to allow the exchange of the 3' *hlyA* region, I needed, first, to generate additional unique sites downstream of the *hlyA* gene. This was achieved by cloning the XmnI/KpnI fragment (encoding the *hlyA* gene) from pLG585-1 (pUC18 *hlyA*,incor,ClaI) into the same sites of pUC19 (see Figure 6). The pUC19 vector was first deleted for its unique BamHI site (by a T_4DNA polymerase

Fill in reaction, see Materials and Methods) to avoid duplication of BamHI sites at opposite ends of the *hlyA* gene once inserted. The final recombinant plasmid, designated pLG585-2, carries extra restriction sites downstream of the *hlyA* gene (see Figure 6).

7.5.1 Cloning *hlyA* and *hlyC* onto pACYC184:

The 5' region of *hlyC*, lacking the reported upstream regulatory region was isolated from pLG577 (pACYC-*hlyC*; Nicaud *et al.*, 1985a) on a 0.7Kb PvuII/BamHI fragment and cloned into the ClaI (blunted by Klenow Fillin reaction)/BamHI sites of pACYC184, generating pACYC-C" (see Figure 7), destroying the *tet* promoter and removing the 5' end of the *tet* gene. It had been necessary to destroy the plasmid encoded ClaI site, within the *tet* promoter region, to ensure that the ClaI site within the *hlyA* gene remained unique.

The remainder of the *hlyC* gene and the entire *hlyA* gene (containing the unique ClaI site), plus the first 72bp of *hlyB* was cloned from pLG585-2 (Figure 6) on a 3.3Kb BamHI/SalI fragment into the same sites of pACYC-C" to generate pACYC-CA,ClaI (see Figure 7). This plasmid encodes Cm^R, HlyC and HlyA and retains the possibility to exchange the 3' of *hlyA* on a 0.6Kb ClaI/KpnI fragment.

7.5.2 Construction of an *hlyB,D* plasmid suitable for complementation:

Since pACYC-CA,ClaI (*hlyC,A*) and pLG575 (*hlyB,D*) have the same selectable marker (Cm^R), an alternative *hlyB,D* plasmid had to be generated for subsequent complementation analysis. This was done by cloning the 5Kb BglII fragment from pLG570 (Mackman and Holland, 1984b) carrying the last 81bp of *hlyA*, plus the *hlyB* and *hlyD* genes, into the compatible BamHI site within the Cm^R gene of the low copy number vector pLG339 (~6-8 per cell; Stoker *et al.*, 1982). This generated the Km^R plasmid pLG339BD (see Figure 8). The orientation of insertion of this BglII fragment was determined by restriction enzyme digest analysis (data not shown).

7.6 Isolation of hydroxylamine mutants using the 2 plasmid complementation system:

Both plasmids (pACYC-CA,ClaI and pLG339BD) when transformed into *E. coli* NM522 complemented each other as determined by the formation of haemolytic colonies on blood agar plates. The expression of HlyC and HlyA in the absence of known promoters supports the previous

supposition that other promoter-like sequences exist upstream of *hlyC* (see Section 7.2.1 above), although it is possible that such activity is directed from vector sequences. No indication of the formation of spontaneous *sec*$^-$ colonies was observed when streaking out to single colonies during the construction and purification of this strain. However, when the hydroxylamine treated 0.6Kb fragment was exchanged for the corresponding fragment of the pACYC-CA,ClaI plasmid, *sec*$^-$ transformants carrying deletions were still obtained (see below). Thus, of the first 338 transformants screened 4 were *sec*$^-$ (1.2%), three of which carried different sized deletions (including the loss of one or both of the ClaI/KpnI restriction sites) extending into the pACYC-CA,ClaI plasmid as determined by restriction analysis (data not shown), while the pLG339B,D plasmid gave the expected restriction pattern. The fourth *sec*$^-$ clone retained the 0.6Kb insert, and might therefore carry a genuine hydroxylamine induced *sec*$^-$ mutation.

7.6.1 Identification of a hydroxylamine induced point mutation:

The 3' end of the *hlyA* gene from the appropriate pACYC-CA,ClaI *sec*$^-$ mutant was subcloned, as before, on an EcoRI/SalI 0.8 Kb fragment into the vector M13mp18 (see Chapter 6: Figure 6). Single stranded phage DNA was then isolated and the entire hydroxylamine treated fragment sequenced as described in Materials and Methods. The reason why the 0.6Kb ClaI/KpnI hydroxylamine treated fragment was not cloned into M13mp18 itself was that this would require the destruction of the ClaI site (through insertion into the compatible AccI site, within the multiple cloning linker, destroying both sites), which would effectively imprison this fragment in the M13 vector. By using the EcoRI/SalI fragment, this allowed subsequent re-isolation of the insert, if required, on the 0.6Kb ClaI/KpnI fragment. Figure 9 shows the DNA sequencing gel autoradiograph with the observed point mutation indicated. Although the sequence data shown in Figure 9 is not the best quality the substitution of a C for a T is evident, when compared to the wild type sequence. This is the expected transition for a hydroxylamine induced mutation, and results in changing the TGG codon (tryphophan^{913}, 111 residues from the C-terminus) to TAG, a stop codon. This change is considered to be upstream of the targeting domain (thought to lie within the last 53 amino acids of HlyA; see Chapter 6; Section 6.1) and therefore provides, no information about the signal itself. However, this result demonstrated that the hydroxylamine mutagenesis reaction had been successful,

encouraging the continued isolation of *sec*⁻ mutants by this method, despite the background of *sec*⁻ deletion mutants observed previously.

7.6.2 Investigation of the mechanism generating the deleted *sec*⁻ mutants:

Restriction analysis of some of the *sec*⁻ deletion mutants isolated earlier, demonstrated that some appeared to carry a deletion equivalent in size to the intended insert, while others contained larger deletions encompassing both the ClaI and KpnI sites and extending outside the expected insertion region (see Section 7.4.1). In an attempt to understand the nature of these deletion events, I decided to sequence across some of the deletion sites. Therefore, I isolated the DNA fragment from some of the deletion clones that retained the unique BamHI/KpnI restriction sites flanking the *hlyA* gene (see pLG585-1; Figure 6), but carrying different sized deletions. The BamHI/KpnI fragments from 3 such clones were inserted into the corresponding sites within M13mp19 and the single strand template sequenced, as before (see Materials and Methods), to determine the deletion points. The results were interpreted as follows:

a) The cloned fragment which appeared to carry a deletion corresponding to the size of the intended insert was indeed that. From the sequence at the junction of the deletion it seems that the KpnI/ClaI sticky ends of the vector fragment had been blunted (due to either aberrant endo or exonuclease activity present in the ligase or restriction enzymes perhaps ?) and then ligated together without the incorporation of an insert (data not shown).

b) Sequencing two fragments carrying apparently larger sized deletions revealed that the sequence immediately counterclockwise of the KpnI site (see pLG585-1; Figure 6) corresponded to sequences at the 5' end of the *hlyA* gene. These deletion mutants were deleted for the final 1852 and 1077bp respectively of the 3' end of the *hlyA* gene (out of a total 3072bp) plus the small 145bp fragment between the end of the *hlyA* gene and the KpnI site. Comparison of the sequence found at the deletion junction site, indicated that the deletion occurred at what appeared to resemble "KpnI-like" sites. KpnI recognises the sequence GGTACC, whereas the deletions originated somewhere within the sequence GGCATCC and GGTAAC respectively (data not shown) as derived from the *hlyA* sequence encoded by the pSF4000 determinant. The latter deletion junction differs by one base from that of the natural KpnI recognition site, whereas the former differs more significantly. However, it should be noted that this region of

the *hly*A LE2001 gene has not be sequenced so that it is possible that in our determinant the former deletion site might be more KpnI "like". The ligation of the KpnI overhang end with the resulting KpnI-like overhanging ends would, presumably, result in at least one mismatch which would be repaired by the cell to either recreate or delete the KpnI site. This would explain why some of the deletion mutants isolated were deleted for both the ClaI and KpnI sites while others retained the KpnI site.

7.6.3 Isolation of additional hydroxylamine *sec*⁻ mutants:

The above results indicated that the presence of the background level of deletion mutants could be attributed to the presence of minor aberrant endonuclease or exonuclease activities in the enzymes used during the DNA manipulation steps, the products of which are readily detected by our sensitive blood assay detection system. However, it was still feasible to screen for *sec*⁻ mutants, only analysing further those that are found to contain an intact 0.6Kb ClaI/KpnI insert.

Re-transformation of another one-fifth of the same ligation mixture including the hydroxylamine treated fragment used above (see Section 7.6), this time generated approximately 1500 transformants of which 36 were scored *sec*⁻ (2.4%). Restriction analysis of the plasmid DNA from 12 of these transformants revealed that all carried deletions, some losing either or both of the KpnI or ClaI sites used for the insertion of the 0.6Kb fragment, with the deletions varying in size as before. Again, the pLG339BD plasmid (*hlyB,D*) continued to show the expected restriction pattern, implying that it had not been altered (data not shown).

In order to try to minimise any potential aberrant activity of the ligase enzyme in the overnight incubation at 15°C, due perhaps to the use of too high a ratio of the large vector fragment to the 0.6Kb insert, leading to nibbling of the overhanging ends, I decided to repeat the ligation with varying molar ratios of the large plasmid fragment to the insert. In this experiment instead of the usual approximately 1:1 molar ratio of ends used, approximately 1:2 and 1:4 molar ratios of plasmid fragment to insert were employed. Transformation of the resulting products into the *E. coli* strain NM522 containing the *hlyB,D* plasmid, pLG339BD, generated 13 *sec*⁻ clones out of about 2000 (0.65%) and 2 *sec*⁻ clones out of about 1000 (0.2%) transformants respectively. Thus, increasing the concentration of insert to vector ratio in the ligation mixture seemed to

reduce the percentage of observed *sec*⁻ mutants and hopefully the background of deletion mutants.

Restriction analysis of 10 of these *sec*⁻ transformants from the 1:2 molar ligation and the 2 *sec*⁻ clones from the 1:4 molar ligation mixture showed that in both cases half of the transformants carried deletions (data not shown). 4 of these deletion mutants appeared to lack only the expected insert, and of the other two, one had a larger sized insert and the other a deletion extending into the *hlyA* gene. Nevertheless, this still left 6 clones hopefully carrying true hydroxylamine induced *sec*⁻ mutations.

7.6.4 Sequencing of 6 potential hydroxylamine induced *sec*⁻ mutants:

The 0.8Kb EcoRI/SalI fragment from each mutant, encompassing the 0.6Kb ClaI/KpnI fragment, was cloned into M13mp18 as before and single stranded DNA isolated and sequenced using a variety of available primers to read the whole 0.6Kb region. The results of the sequence analysis was disappointing in that one of the clones carried an insert unrelated to the 3' end of the *hlyA* gene, 3 apparently carried no detectable mutations, while another was unexpectedly deleted for a single cytosine (data not shown). Finally the last mutant had a single point mutation changing an Arg^{995} (29 residues from the C-terminal) to Lys (see Figure 10). This latter change, fortuitously destroyed the unique BglII site 81bp from the 3' end of *hlyA*, providing a subsequent convenient screen for this mutant.

7.6.5 Is Arg^{995} an important residue for signal sequence recognition ? :

The above results indicated that the substitution of the large positive Arg^{995} residue for Lys abolishes secretion of HlyA as detected on blood agar. However, the substituting amino acid (Lys) is also a large positive residue and it appeared surprising that this change would cause the observed *sec*⁻ null phenotype. Indeed, subsequent expression of this mutation as the 3' *hlyA* fragment, encoding the C-terminal 23KD of HlyA from the M13mp18 *lac* promoter, resulted in its apparently normal secretion by HlyB,D (see Chapter 8; Section 8.2 for details of procedure). Secretion in this case was monitored by TCA precipitation of the supernatants from a culture expressing HlyB,D and infected with M13-23KD*hlyA* carrying the Arg to Lys substitution (see Figure 11, Lanes 6). Therefore, by this test the mutation had no effect upon secretion. This discrepancy between the normally secreted 23KD C-terminal peptide and

the non-secretion of the intact toxin, raised some important questions about the validity of the secretion tests or the method of mutagenesis. In order to clarify possible discrepancies in the secretion complementation tests the 0.6Kb ClaI/KpnI fragment from the M13mp18-23KD*hlyA* Arg to Lys mutant, was re-isolated and cloned back into the intact *hlyA* gene on pACYC-CA,ClaI plasmid. This construction was then tested for its ability to promote secretion of the HlyA toxin, on blood agar plates, in the presence of pLG339BD (*hlyB,D*). In this case the transformants now displayed a *sec*+ phenotype. This result clearly indicated that a mutation, in addition to the Arg^{995} to Lys change, must have arisen outside the ClaI/KpnI 0.6Kb fragment during hydroxylamine treatment or a subsequent *in vitro* manipulation step, leading to loss of haemolytic activity. This result unfortunately indicated that the mutagenesis procedure was flawed in some way in addition to the production of the detectable deletion mutations.

7.6.6 Investigation of the nature of another 5 *hlyA sec*- mutants:

Screening of further transformants from the original ligation mixture (see Section 7.6) led to the isolation on blood agar plates of another 5 *sec*- mutants carrying an 0.6Kb ClaI/KpnI insert which had been treated with hydroxylamine. The 0.8Kb EcoRI/SalI fragments (encompassing the treated ClaI/KpnI region) were cloned into M13mp18, as before, infected into NM522 pLG339BD (*hly*B,D) cells, cultures grown and expression induced for the synthesis of the 23KD C-terminal domain (as mentioned previously and see Materials and Methods). Unfortunately, as before efficient secretion was observed compared to the wild-type control (see Figure 11, Lanes 7 to 11). Hence, these results were again in conflict with the results obtained on the original blood plate screen apparently confirming that the *sec*- phenotype might be due to additional mutation events outside the region exposed to hydroxylamine.

7.7 Discussion:

Although, the strategy of hydroxylamine mutagenesis used to generate random mutations in the 3' end of *hly*A was successful, in that specific hydroxylamine induced mutations were identified, the great majority of the *sec*- mutants isolated were shown not to be due directly to the hydroxylamine exposure but more likely as a result of the DNA manipulation involved in the isolation of a specific *hlyA* fragment for mutagenesis. Investigation into the nature of some of these pseudo *sec*- mutants revealed the existence of deletions within the *hlyA* gene

associated with the DNA manipulations steps. Many of these "deleted plasmids" appeared in fact to lack a fragment corresponding in size to that of the intended insert. Sequencing of one such example indicated that this may have arisen as a product of aberrant ligase or restriction enzyme activity, blunting the vectors sticky overhang end, followed by blunt ended ligation without the incorporation of an insert to give the *sec*- phenotype.

A second class of deletion mutant had larger sized deletions and had lost at least one of, if not both, the ClaI and KpnI restriction sites bordering the 0.6Kb *hlyA* hydroxylamine treated fragment. Sequencing two mutants deleted for only the ClaI site suggested the aberrant recognition of the KpnI restriction enzyme for "KpnI-like" sequences within the plasmid *hlyA* encoding region. Ligation of the resulting overhangs with the genuine KpnI sticky end, would then upon cellular mismatch repair, be expected to restore or delete the unique KpnI site, as observed with the deleted mutants obtained. A possible explanation for this process is the use of a single restriction buffer for ClaI/KpnI digestion. Both enzymes were obtained from BRL together with their optimum restriction digestion buffers. However, the
manufactures state that if Buffer number 1 is used Cla1 digests the DNA with a 100% efficiency while KpnI functions at 60% efficiency but without observed aberrant activity. Thus, a single buffer was used in these experiments for convenience. Therefore, the use of the KpnI enzyme in this "non-optimal" activity buffer might be the cause of the observed "star" activity. Thus, it might be advisable, in future studies, to carry out the restriction digests in their optimal buffers (when rare events are being screened for) if this problem is to be avoided.

Despite these inherent problems, which probably could be reduced if not eliminated with time, it was worthwhile continuing with the procedure since conditions were already established where 50% of the isolated *sec*- mutants still retained the intact 0.6Kb ClaI/KpnI hydroxylamine treated insert. Disappointedly, the *sec*- nature of 11 out of 12 of these mutants as screened on blood agar, was contradicted when the mutant 3' *hlyA* fragments (encoding the C-terminal 23KD HlyA peptide) were expressed from the M13 expression vector and found to be secreted normally in the presence of HlyB,D. This result therefore indicated that the original screen for *sec*- mutants on blood agar plates must at least be, in part, due to additional mutation events outside the hydroxylamine treated fragment, affecting either the transcription/ translation, activity or

stability of the toxin. Clearly, these problems must arise as a result of either the restriction digestion, isolation and/or ligation reactions carried out in the mutagenesis procedure. One possible explanation might be the generation of small *hlyA* internal deletions or rare mutational events caused by traces of ethidium bromide bound to the DNA fragments and/or exposure of the DNA fragments to damaging ultra-violet irradation during the fragment isolation procedure. However, until this problem can be resolved it was decided to attempt another approach (see subsequent Chapter) to generate a bank of random mutants in the 3' *hlyA* region.

Figure 1:

hly determinant from *E .coli* LE2001

The restriction map of the subcloned *hly* determinant is shown (adapted from Mackman, 1984) together with the position of the *hly* genes and their observed gene product sizes (see Chapter 1; Section 1.6). Shown, immediately below are some of the useful restriction sites inferred from the sequenced regions of the LE2001 determinant (Nicaud *et al.*, 1985a; Gray *et al.*, 1986). The arrowed lines indicate the direction of known promoter activity with the broken line indicating decreased transcription due to the presence of a Rho independent terminator in the intragenic region between *hlyA* and *hlyB* (see Chapter 1; Section 1.6.4)

B=BamHI; H=HindIII; E=EcoRI; Bg=BglII; S=Sma; Bl=Bal D=DraI; P=PstI and Pv=PvuII.

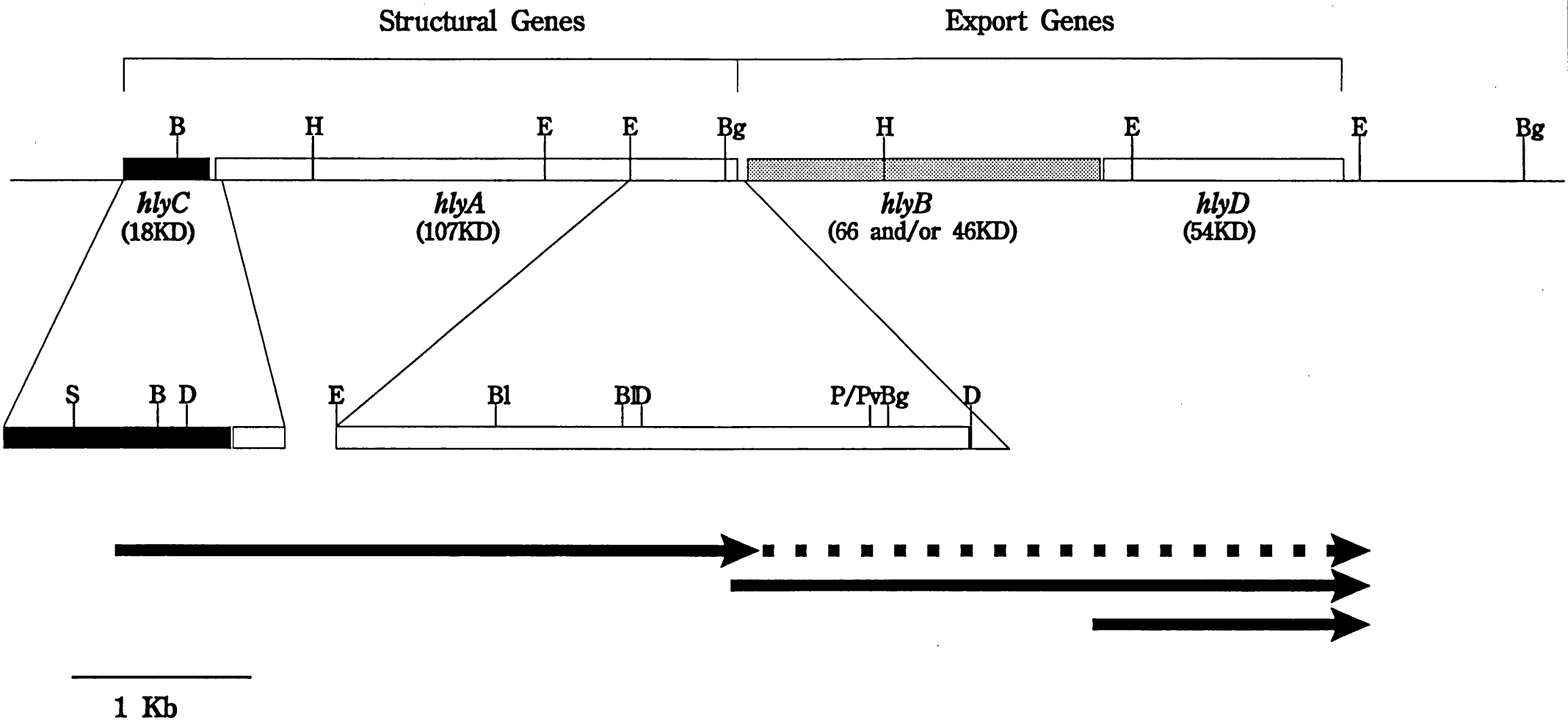
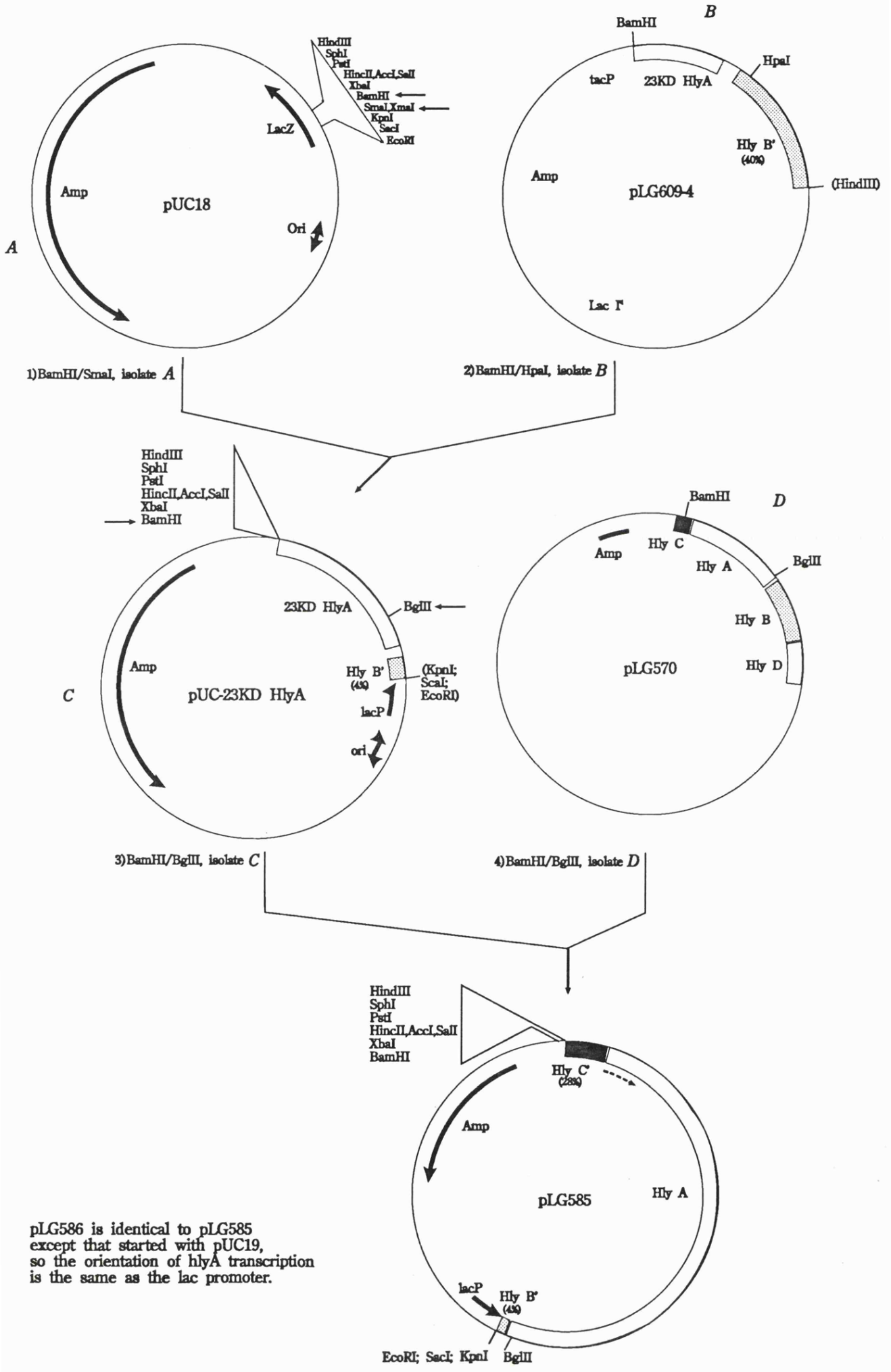

Structural Genes
Export Genes
B
H
E
E
Bg
H
E
E
Bg
hlyC
(18KD)
hlyA
(107KD)
hlyB
(66 and/or 46KD)
hlyD
(54KD)
S
B
D
E
Bl
BD
P/Pv
Bg
D
1 Kb

Figure 2:

Construction of pLG585 (pUC-*hlyA*,incor) and pLG586 (pUC-*hlyA*,cor).
pLG585 and pLG586 only differ in the orientation of insertion of the *hlyA* gene (and the 3' region of *hlyC*) relative to the pUC *lac* promoter.
See text for details. The broken arrowed line indicates the direction of transcription of the *hlyA* gene.
HlyB' indicates the presence of 4% of the 5' of the *hlyB* gene.
HlyC' indicates the presence of 28% of the 3' of the *hlyC* gene.
pUC18: Norrander *et al.*, 1983
pLG609-4: see Chapter 3; Figure 2a/b
pLG570: Mackman and Holland, 1984b



pLG586 is identical to pLG585
except that started with pUC19,
so the orientation of hlyA transcription
is the same as the lac promoter.

Figure 3:

a) Restriction enzyme digest pattern of pLG585 (pUC-*hlyA*,incor) carrying the *hlyA* gene.

Complete sequence of pUC18/19 vector portion available on EMBL database.

Approximately 0.5μg of CsCl prepared pLG585 plasmid DNA was digested in a total volume of 20ul. After 1 hour digestion time with the appropriate restriction enzyme(s), half the sample was then run out on a long 0.7% agarose gel, and ethidium bromide stained for visualisation under ultra-violet light.

Lane 1: Uncut; Lane 2: X EcoRI; Lane 3: X EcoRI/BamHI; Lane 4: EcoRI/BglII; Lane 5: X EcoRI/PstI; Lane 6: X PstI; Lane 7: X PstI/HindIII; Lane 8: X HindIII; Lane 9: X HindIII/ BglII; Lane 10: X BglII; Lane 11: X BglII/SalI; Lane 12: X SalI; Lane 13: X BamHI; Lane 14: X BamHI/KpnI; Lane 15: X KpnI; Lane 16: X PvuII; Lane 17: X PvuII/ScaI; Lane 18: X ScaI.

Molecular weight markers:

A) λ HindIII: (Kb) 23.1; 9.4; 6.6; 4.4; 2.3; 2; 0.56

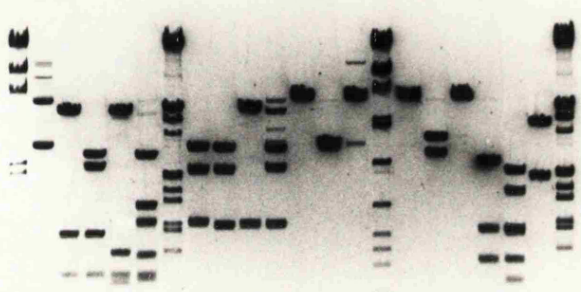
B) λ HindIII/EcoRI: (Kb) 21.2: 5.2; 5; 4.3; 3.5: 2; 1.9; 1.6; 1.3; 0.98; 0.83; 0.56

C) λ HindIII/NcoI: (Kb) 19.3; 9; 6.6; 4.3; 3.8; 2.3; 1.3; 0.77; 0.56; 0.39; 0.125.

b) Revised restriction map of *hlyA* from the LE2001 determinant compared to the pSF4000 sequenced determinant.

pSF4000 *hlyA* sequence obtained from the EMBL DNA sequence database provided by Felmlee *et al.*, (1985a).

a)



A 1 2 3 4 5 B 6 7 8 9 10 11 12 C 13 14 15 16 17 18 B

b)

Restriction map of E.coli hlyA LE2001 compared to known pSF4000 gene

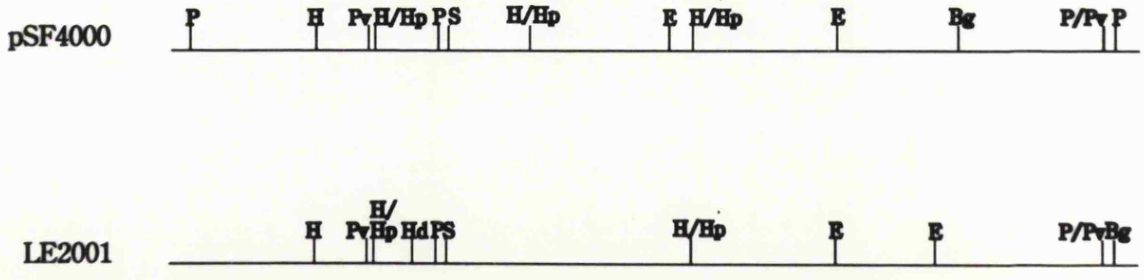
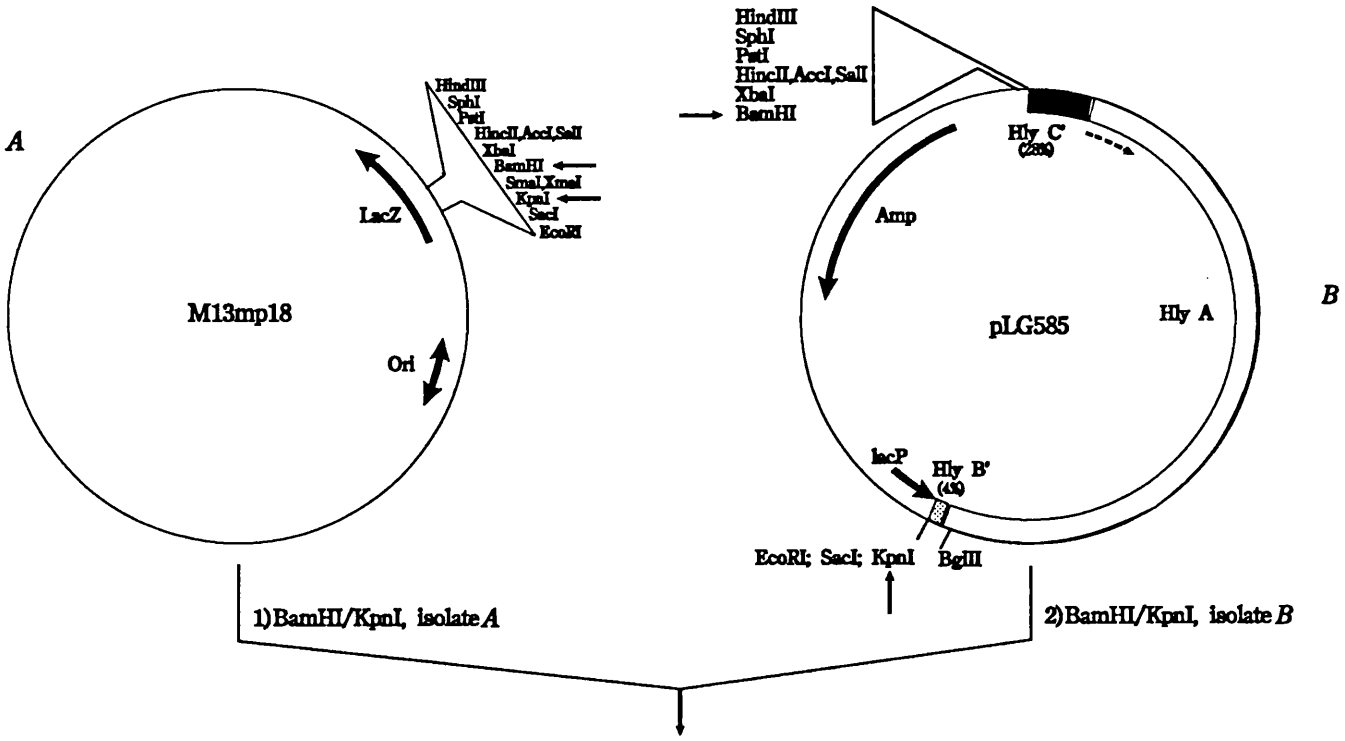
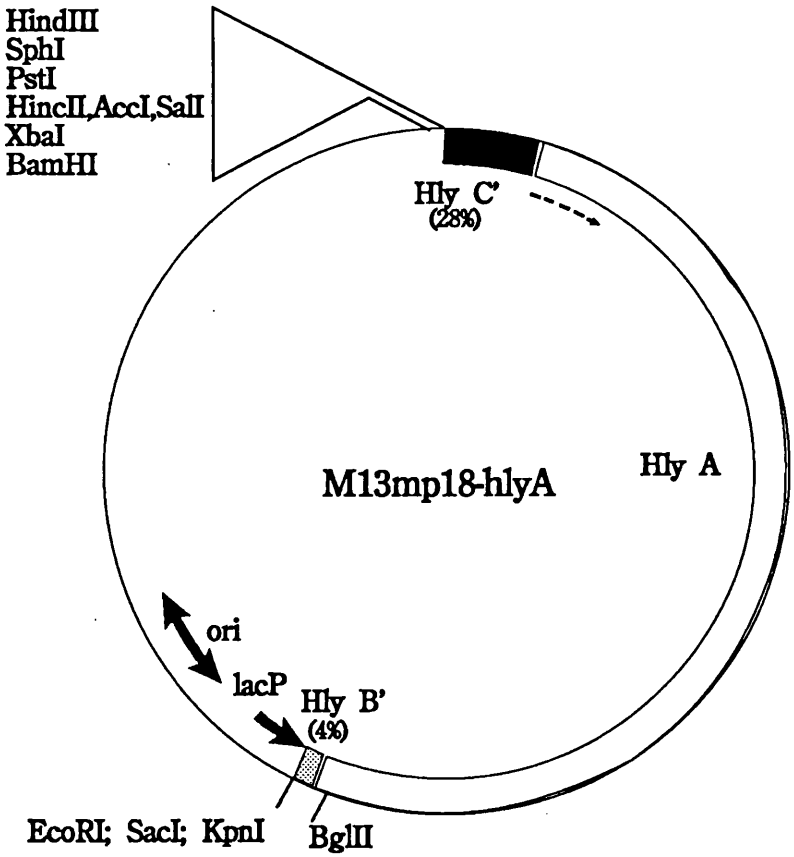


Figure 4:
Insertion of *hlyA* into the M13 single stranded vector.
For details see text.
M13mp18/19: Yanish-Perron *et al.*, 1985
HlyB' indicates the presence of 4% of the 5' of the *hlyB* gene.
HlyC' indicates the presence of 28% of the 3' of the *hlyC* gene.
Broken arrow line indicates orientation of transcription of the *hlyA* gene.





M13mp19-hlyA is identical to M13mp18-hlyA
except that started with M13mp19,
so that the orientation of hlyA transcription
is the same as the lac promoter.

Figure 5:

Restriction enzyme digest pattern of four *sec*⁻ hydroxylamine treated pLG585-1 (pUC-*hlyA*, ClaI, incor) transformants.

Four Amp^R pLG585-1 *sec*⁻ plasmids were separated from the *hlyC* (pLG591Tc) and *hlyB,D* (pLG575) plasmids by transforming isolated DNA into *E. coli* NM522 and selecting for Amp^R, Tet^S, Cm^S colonies. Isolated plasmid DNA, prepared by the mini-preparation method (see Materials and Methods), from the four transformants, along with some control pLG585-1 DNA, was digested with EcoRI and the resulting fragments separated and visualised on an agarose gel (see Materials and Methods).

Lane 1: pLG585-1 positive control

Lane 2-5: pLG585-1 *sec*⁻ transformants 1-4

The λ HindIII marker bands are indicated with the arrows showing the expected 3 band pattern (~0.3, 0.8 and 5Kb) of the control plasmid DNA, pLG585-1, digested with EcoRI. Note, that it is the 0.8Kb fragment which encodes the 3' 657bp of *hlyA* and thus the region exposed to the mutagen.

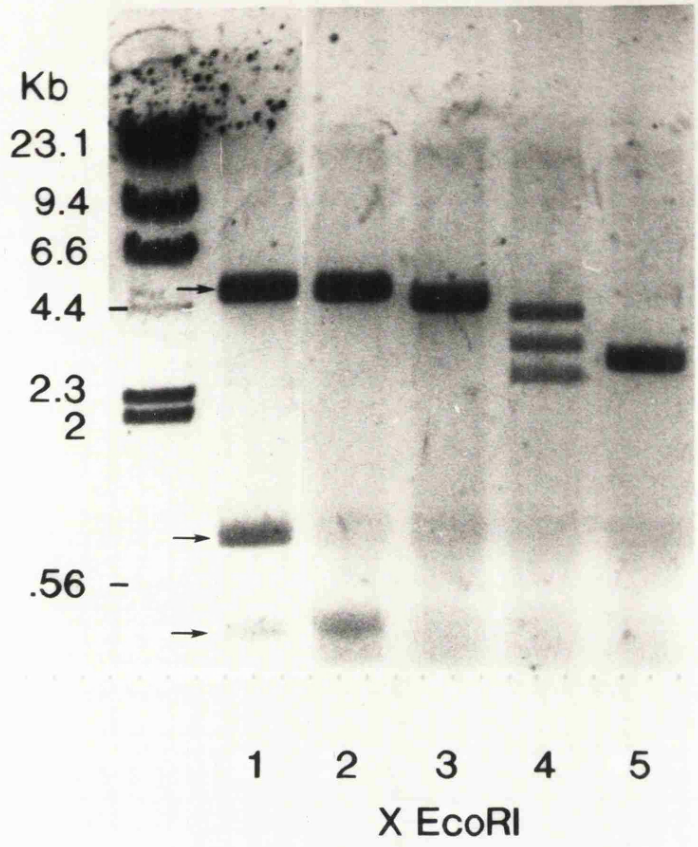

Kb
23.1
9.4
6.6
4.4
2.3
2
.56
1
2
3
4
5
X EcoRI

Figure 6:

Construction of pLG585-2

See text for details.

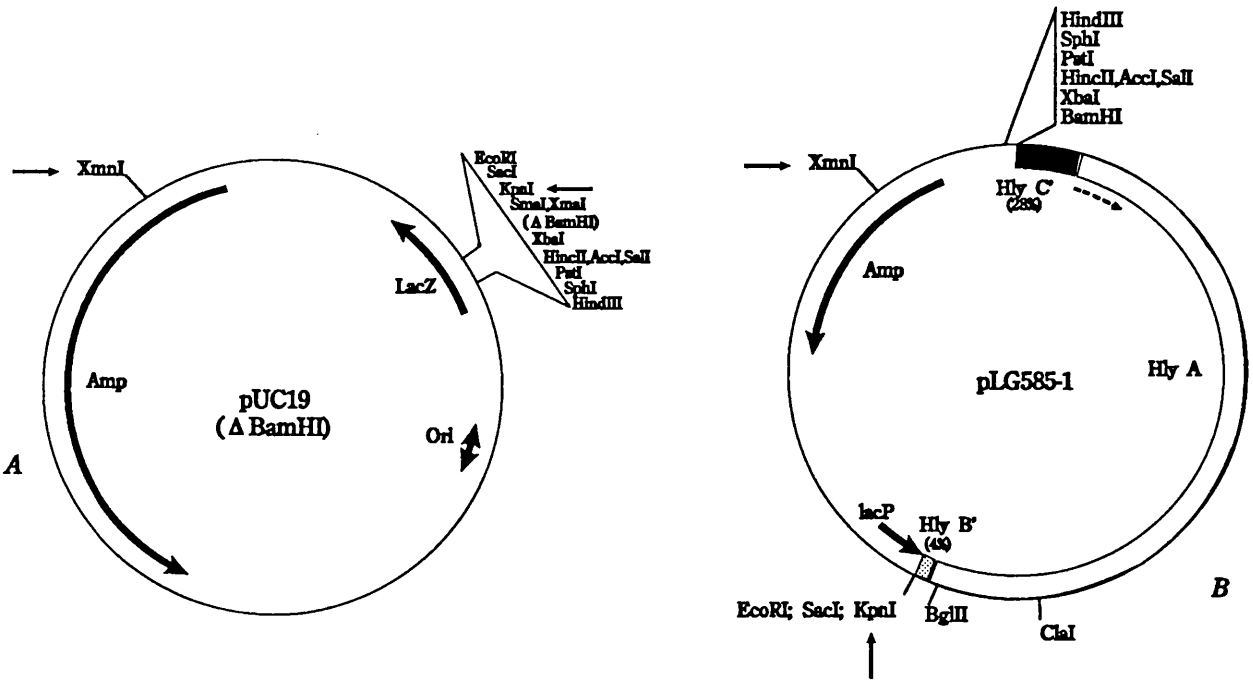
pUC19: Norrander *et al.*, 1983. The unique BamHI site was deleted, by a fill in reaction using T4 DNA polymerase (see Materials and Methods).

pLG585-1: see Section 7.2.3

HlyB' indicates the presence of 4% of the 5' of the *hlyB* gene.

HlyC' indicates the presence of 28% of the 3' of the *hlyC* gene.

Broken arrow line indicates orientation of transcription of the *hlyA* gene.



1)KpnI/XmnI, isolate *A*

2)KpnI/XmnI, isolate *B*

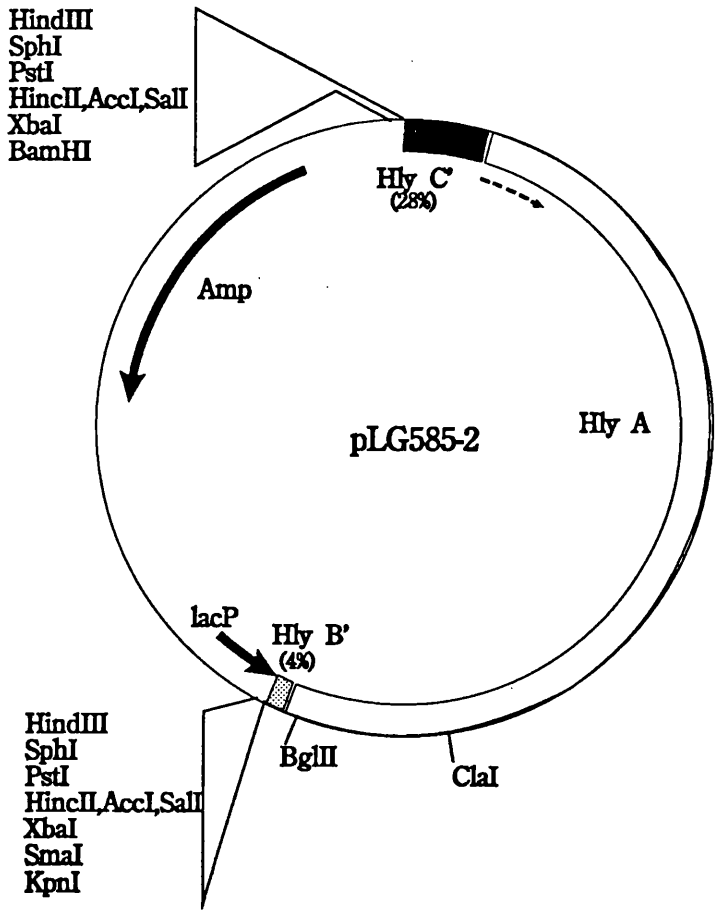


Figure 7:
Construction of pACYC-CA (*hlyC,A*)
See text (Section 7.5) for details.
pLG577: Nicaud *et al.*, 1985a
pACYC184: Chang and Cohen, 1978
pLG585-2: see Figure 6
HlyB' indicates the presence of 4% of the 5' of the *hlyB* gene.
HlyC' indicates the presence of 28% of the 3' of the *hlyC* gene.
HlyC" indicates the presence of 72% of the 3' of the *hlyC* gene.
HlyA' indicates the presence of 78% of the 5' of the *hlyA* gene.

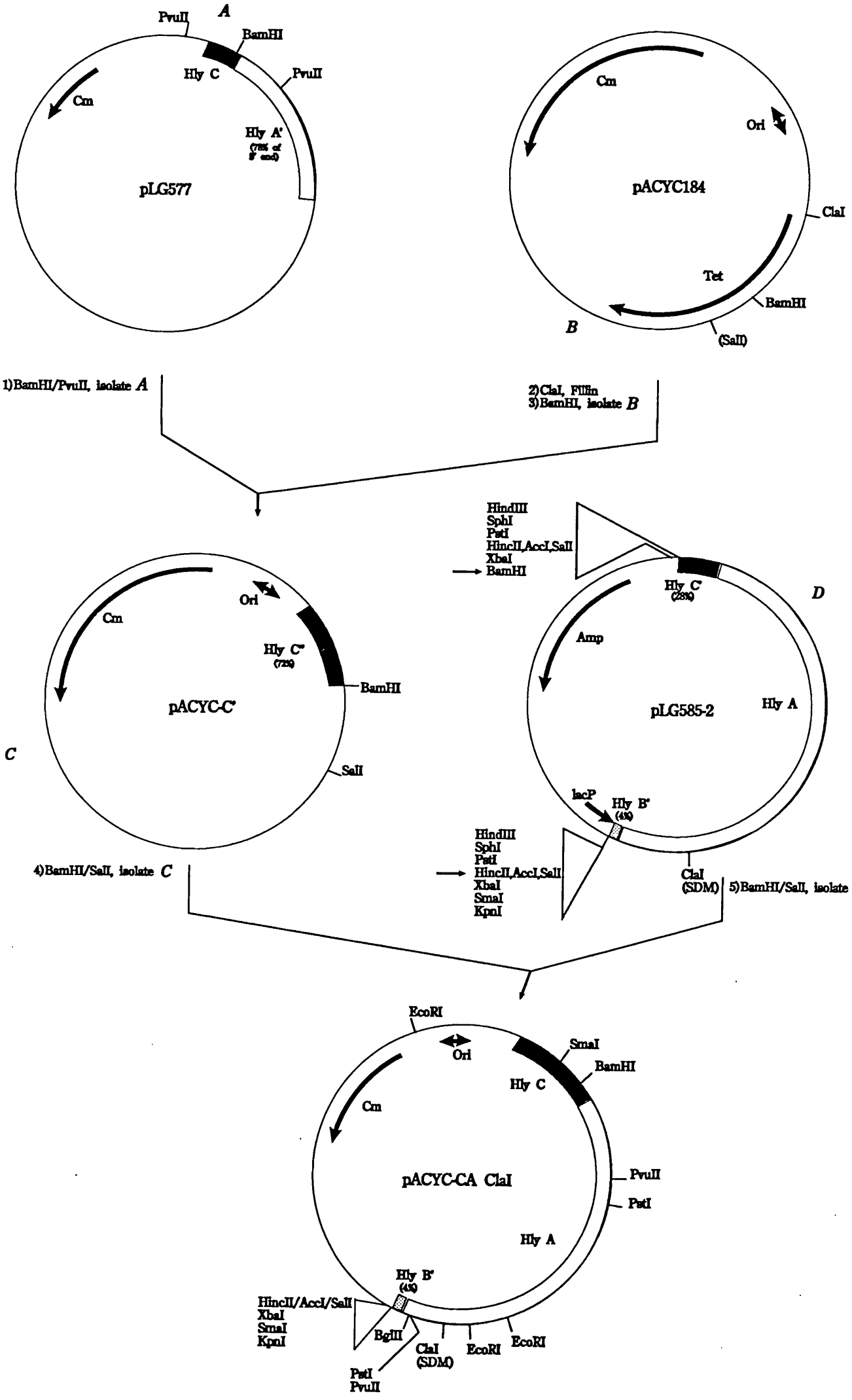

A
PvuII
BamHI
Hly C
PvuII
Cm
Hly A'
(78% of 3' end)
pLG577
Cm
Ori
pACYC184
ClaI
Tet
BamHI
B
(SalI)
1) BamHI/PvuII, isolate A
2) ClaI, Fillin
3) BamHI, isolate B
HindIII
SphI
PstI
HincII,AccI,SalI
XbaI
BamHI
Hly C
(28%)
D
Amp
pLG585-2
Hly A
lacP
Hly B'
(4%)
ClaI
(SDM)
5) BamHI/SalI, isolate
Ori
Cm
Hly C'
(72%)
BamHI
pACYC-C'
SalI
C
4) BamHI/SalI, isolate C
HindIII
SphI
PstI
HincII,AccI,SalI
XbaI
SmaI
KpnI
EcoRI
Ori
SmaI
BamHI
Hly C
Cm
pACYC-CA ClaI
PvuII
PstI
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
BglII
ClaI
(SDM)
EcoRI
EcoRI
PstI
PvuII

Figure 8:
Construction of pLG339B,D (*hlyB,D*)
See text (Section 7.5.2) for details.
pLG339: Stoker *et al.*, 1982
pLG570: Mackman and Holland, 1984b
3'HlyA indicates the presence of 4% of the 3' of the *hlyA* gene.

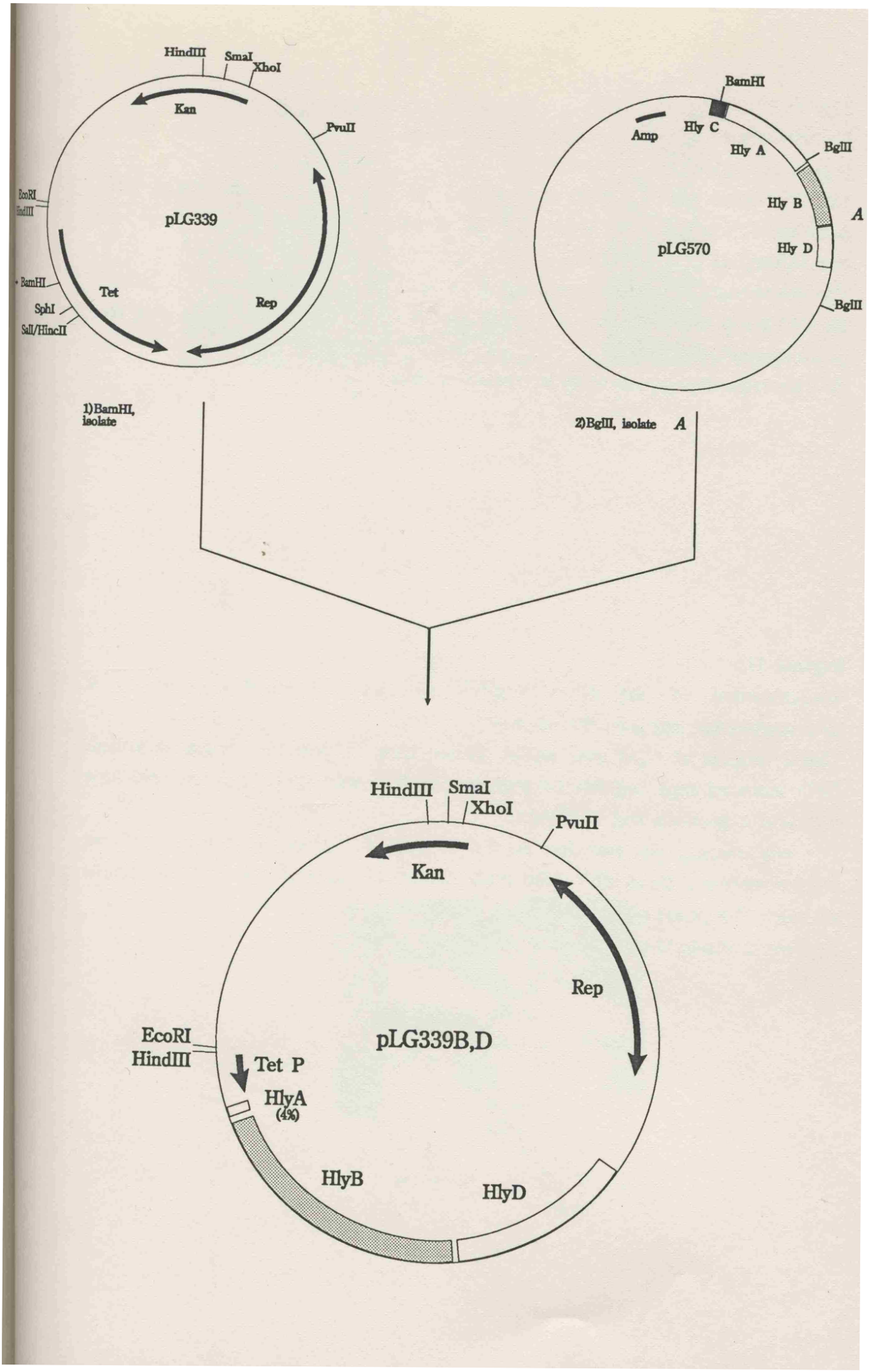
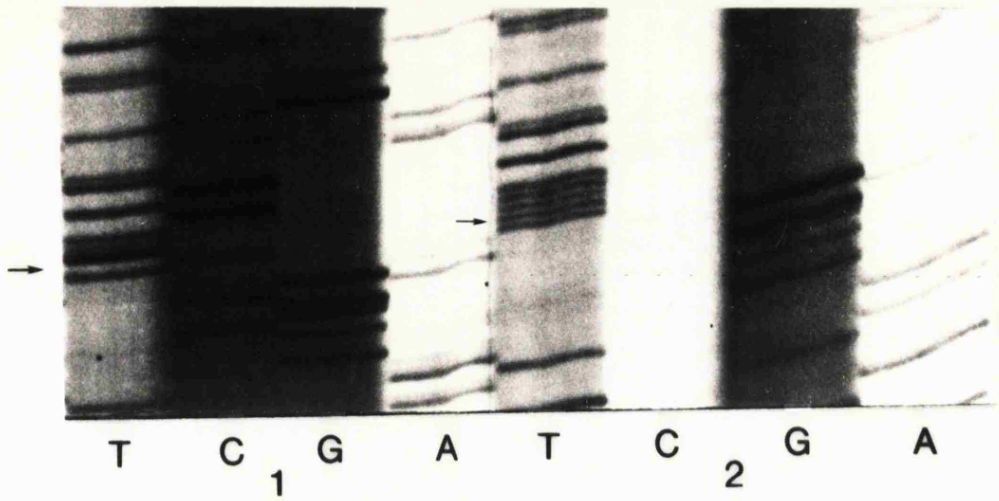

HindIII
SmaI
XhoI
Kan
PvuII
EcoRI
HindIII
pLG339
BamHI
Tet
Rep
SphI
SalI/HincII
BamHI
Hly C
Amp
Hly A
BglII
Hly B
A
pLG570
Hly D
BglII
1) BamHI, isolate
2) BglII, isolate A
HindIII
SmaI
XhoI
PvuII
Kan
Rep
EcoRI
HindIII
pLG339B,D
Tet P
HlyA
(4%)
HlyB
HlyD

Figure 9:
Identification of a hydroxylamine induced C to T transition in the 3' region of *hlyA* (Trp^{913} to stop codon).
The 3' region of *hlyA* was cloned into M13mp18, single stranded DNA isolated and sequenced together with a wild-type *hlyA* control (see text and Materials and Methods).
The sequencing gel was loaded T,C,G and A. 1 indicates those tracks derived from the wild-type sequence and 2 the transformant carrying the point mutation.
Arrows indicate the nucleotide transition from C to T.

Figure 10:
Identification of the HlyA Arg^{995} to Lys substitution due to a hydroxylamine induced transition.
The 3' region of *hlyA* was again cloned into M13mp18, single stranded DNA isolated and sequenced together with a wild-type *hlyA* control (see text and Materials and Methods).
The sequencing gel was loaded T,C,G and A tracks. 1 indicates those tracks derived from the wild-type sequence and 2 the transformant carrying the point mutation.
Arrows indicate the nucleotide transition from C to T.


T C G A T C G A
1
2

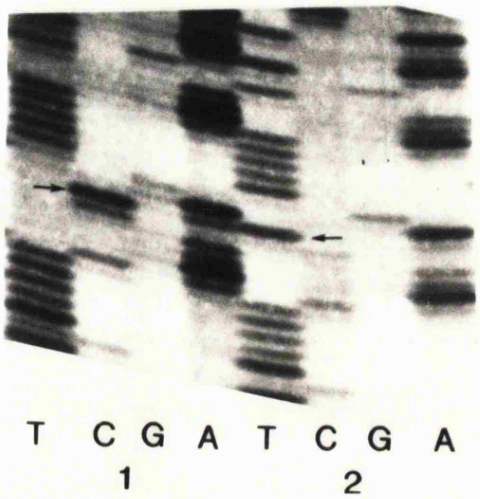

T C G A T C G A
1
2

Figure 11:

Expression and secretion of HlyA C-terminal 23KD peptides encoded by *sec*⁻ mutants isolated after hydroxylamine mutagenesis and identified on blood agar screening.

The hydroxylamine treated *hlyA* fragment (encoding the C-terminal 23KD peptide) from some of the *sec*⁻ *hlyA* transformants was transferred into M13 for expression from the *lac* promoter. The M13-23KD transformants (plus a wild-type 23KD HlyA control) were infected into NM522 cell carrying pLG339B,D (*hlyB,D*), cultures grown to an A^{450} of 0.4 before inducing 23KD HlyA expression with IPTG (0.5mM final concentration). Growth was maintained at 37°C for 2 hours before isolating and concentrating supernatant samples with TCA (10%v/v final concentration; see Materials and Methods)). 1 A^{450} equivalent unit of each supernatant sample was separated on a 15% mini- SDS-acrylamide gel and stained with Coomassie blue.

Supernatant protein profiles:

Lane 1: wild-type 23KD HlyA

Lane 2: "*sec*⁻ mutant" carrying single cytosine base deletion (see text).

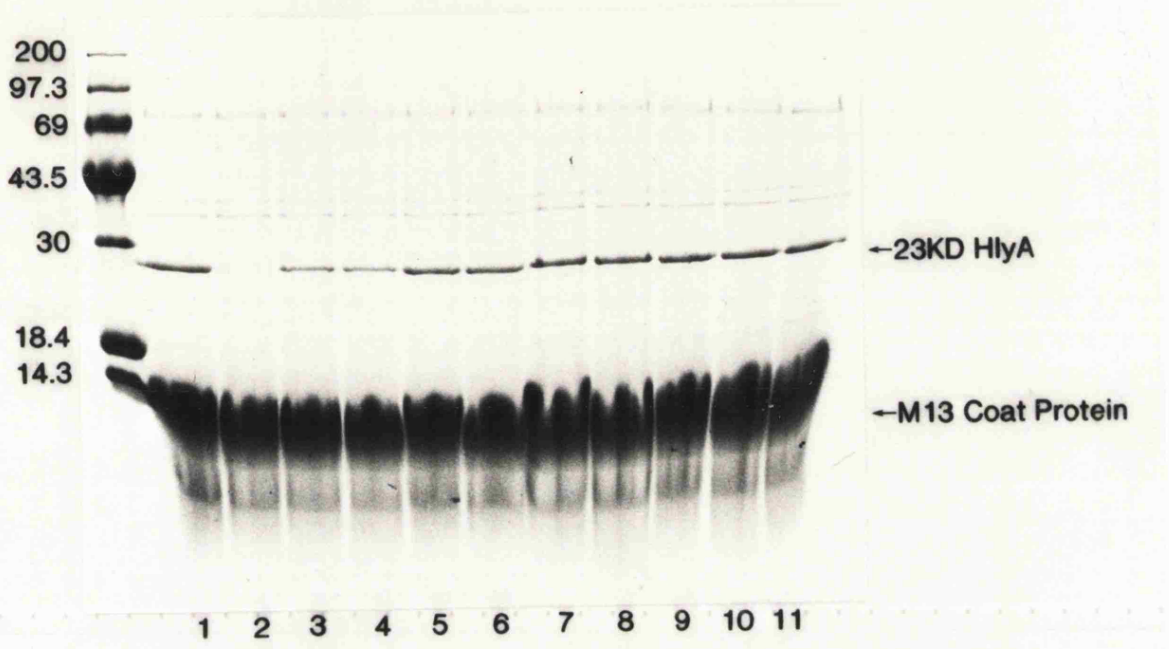
Lanes 3-5: "*sec*⁻ mutants" secreted at apparently wild-type levels.

Lane 6: "*sec*⁻ mutant" carrying Arg^{995} to Lys substitution.

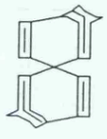
Lanes 7-11: Another 5 unsequenced "*sec*⁻ mutants" as originally screened on blood agar plates.

Molecular weight size markers are indicated together with the position of the 23KD HlyA peptides and M13 phage coat protein bands.

Note the subset of cellular band apparently released as a consequence of bacteriophage budding from the cell.


200
97.3
69
43.5
30
18.4
14.3
←23KD HlyA
←M13 Coat Protein
1 2 3 4 5 6 7 8 9 10 11

CHAPTER



Saturation Mutagenesis Analysis of the HlyA C-terminal Signal Domain

8.1 Introduction:

As argued in chapter 6 (Section 6.1) it appears that the novel HlyA C-terminal signal sequence motif is encoded within the last 53 residues (~160bp) of HlyA, with the presence of the last 27 being essential for secretion (Mackman *et al.*,1987). Colin Hughes' group in Cambridge (Koronakis *et al.*, 1989), using a series of C-terminal deletions, by the introduction of stop codons at the 3' end of the *hlyA* gene, also demonstrated that the removal of the final 8 residues reduced the level of HlyA secreted by approximately 50% whereas deletions of 21 or 27 residues virtually abolished HlyA secretion. Comparison of the protein sequence of the C-terminal region of closely related haemolysin toxins (*E. coli*, *Proteus* and *Morganella* species) has revealed some apparently conserved features at the primary and secondary levels (Koronakis *et al.*, 1989), although these do not appear to be conserved in more distantly related molecules carrying C-terminal targeting signals (see Chapter 6, Section 6.2). A number of these heterologous toxins have nevertheless been shown to be secreted in a HlyB,D dependent manner, when expressed in *E. coli*, although usually at significantly reduced efficiencies (see Chapter 6; Section 6.2). Investigation into a possible functional role in *E. coli* HlyA secretion for some of these putative primary or secondary features, by either random hydroxylamine or specific site directed mutagenesis, has so far not revealed the nature of the targeting signal (see Chapter 6 and 7). An alternative approach was therefore embarked upon in an attempt to elucidate the nature of the novel C-terminal targeting signal. The strategy chosen was to "saturate" the signal encoding region of *hlyA* with mutations, with the expectation that the isolation and identification of secretion null and defective mutants would pinpoint those residues essential for recognition. In addition, it was also anticipated that the identification of other substitutions with only partial or no effect upon secretion would further describe the nature of the targeting signal recognised by the export machinery.

The approach used to "saturate" the *hlyA* region, to generate, in theory a large number of substitutions at every position, was to synthesise a specific oligonucleotide complementary to a region of *hlyA*, under conditions which will incorporate at random a desired number of incorrect bases per synthesised molecule. The conditions to generate such a "randomised" or "doped" oligonucleotide was previously described by Hutchison *et al.*, (1986) where each nucleoside phosphoramidite (A,T,C and G) was cross-contaminated with a low level of the other three bases,

prior to the synthesis reaction, resulting in the generation of, on average, 1.5 mutations per synthesised molecule.

8.1.1 Saturation Mutagenesis of a segment of the 3'end of *hlyA*:

In order to test the saturation mutagenesis procedure in our hands it was decided to initially generate randomised oligonucleotides for only a portion of the signal encoding region. The segment chosen was a contiguous 42bp (encoding 14 residues) targeted by two randomised 21mer oligonucleotides - BK9 and BK10 (see Figure 1). These were sythesised by J. Keyte (Dept of Biochemistry, Leicester Univ.) with an automated sythesiser according to the conditions described by Hutchison *et al.*, (1986). The 3' *hlyA* region under study is also shown in Figure 1 together with the putative conserved features of the related haemolysin molecules possibly playing a role in targeting (see Chapter 6; Section 6.2). Thus, probing this region by "saturation" mutagenesis should highlight those residues essential for targeting and may help to evaluate the possible role of any such "conserved" features in secretion, such as the predicted strong alpha-helix, amphiphilic helix, "aspartate box" or the clustered charged region (see Figure 1).

8.1.2 Saturation Mutagenesis Procedure:

The 'mutations' incorporated into the randomised oligonucleotides were to be introduced into the *hlyA* gene using the Amersham site directed mutagenesis kit. The *hlyA* fragment to be mutagenised was first transferred into a vector capable of generating single stranded DNA (M13; Messing et al; 1983). As previous attempts to subclone the entire *hlyA* gene into the M13 vector had indicated instability problems (see Chapter 7: Section 7.2.3), it was decided to clone into M13 (see below) only the 3' portion of *hlyA* on a 0.8Kb EcoRI/KpnI fragment (encoding the C-terminal 23KD HlyA peptide). After the mutagenesis reaction the mutant *hlyA* fragments could then be subcloned back into the entire *hlyA* gene and screened, in the presence of *hlyC,B* and *D*, on blood agar plates for secretion defective mutants (see below). The identity of any defective mutants could then be readily determined by sequencing only the *hlyA* region probed by these initial mutagenic oligonucleotides.

8.1.3 Saturation Mutagenesis Reaction:

As described in Chapter 7 (Section 7.5) I had constructed an *hly* complementation system with *hlyC* and *hlyA* expressed from the pACYC184 based vector (low copy number; Chang and Cohen, 1978) and

hlyB,D expressed from pLG339 (low copy number; Stoker et al., 1982) generating pACYC-CA,ClaI and pLG339B,D respectively. The pACYC-CA,ClaI plasmid was constructed so that it possessed unique SalI and KpnI restriction sites downstream of the *hlyA* gene. In addition, a unique ClaI restriction site was also added (by site directed mutagenesis; see Chapter 7; Section 7.2.3) permitting the exchange of the final ~450bp of *hlyA* on a 0.6Kb ClaI/SalI or ClaI/KpnI fragment. Therefore, the 0.8Kb EcoRI/KpnI 3' *hlyA* fragment (preserving the ClaI site) was cloned into the M13 vector for the mutagenesis reaction, allowing the subsequent transfer of the mutant ClaI/KpnI *hlyA* fragments back into the intact *hlyA* gene for complementation analysis (see Figure 2 for diagrammatic representation of the proposed scheme).

The M13mp18-23KDHlyA,ClaI clone was used to generate single stranded DNA for use in an *in vitro* site directed mutagenesis reaction as described by the manufactures (see also Materials and Methods). Instead of using each of the "randomised" 21mer oligonucleotides in a separate reaction I decided to introduce equal molar amounts of both oligonucleotides into the same reaction. This would hopefully save both time and expense and generate a bank of mutants carrying mutations in both regions, allowing us subsequently to study the additive effects of double mutants upon secretion. A fifth of the product of the mutagenesis reaction was used to transform (by electroporation; see Materials and Methods) *E. coli* TG1 cells. Hundreds of M13 plaques were obtained and M13 phage stocks of 200 transformants were made and stored at 4°C (see Materials and Methods). In order to test the efficiency of the system in introducing random point mutations into this region, I decided to sequence 72 putative mutants for statistical analysis without prior screening for effects on secretion.

8.1.4 Sequence analysis of 3' region of putative *hlyA* mutants:

The first 72 M13mp18-23KDHlyA,ClaI "mutant" phage stocks were used to infect the *E. coli* F' carrying strain NM522 (F' required for sex pili production and thus M13 infection), single stranded DNA was isolated and the relevant 42bp region sequenced (see Materials and Methods). The results from this analysis are presented in Table 1. Of the 72 clones sequenced, 64 were readable of which 57% carried mutations, the rest being wild-type in sequence. Each of the 21mer oligonucleotides used in the reaction generated a similar number of mutational events, with 70, 25 and 5% of the mutants carrying single, double and triple mutations

respectively. However, since both oligonucleotides were used in the same reaction the overall outcome for the 42bp region was that 61% of the mutants carried single point mutations, 22% double and 6% carried either triple or quadruple mutations. One of the transformants carried sequences homologous to the 5' end of *hlyA* and it is not known by what aberrant event this became ligated into the M13mp18 vector.

Analysis of the spread of these mutations within the probed 14 residue region demonstrated that each residue (except number 8 [Ser^{996}] of the 14) suffered at least one mutational event (See Table 2). There does, however, appear to be a lower incidence of mutations corresponding to the ends of the oligonucleotides, indicating difficulties in generating mutants in these regions perhaps due to inefficient binding of the primer to the template at the ends of the molecule. The nature and distribution of the residues substituted in this study are depicted in Table 2.

8.2 Effect of residue substitutions on the secretion of the 23KD HlyA peptide:

Now that a number of base substitution mutants had been isolated, it only remained to characterise their effects on secretion. Instead of the time consuming process of subcloning each mutant, 3' *hlyA* fragment back into the entire *hlyA* gene, a simpler alternative procedure was developed for initial analysis. It was noted that the insertion of the 0.8Kb EcoRI/KpnI, 3' *hlyA* fragment, carrying the unique ClaI site and encoding the C-terminal 23KD of HlyA, was inframe with the *lac* initiation Met codon on the M13 vector. Thus, it was thought feasible to estimate directly the level of secretion of the various HlyA 23KD mutant peptides from cells infected with recombinant M13 clones, in the presence of HlyB,D. Any effects of the substitutions on secretion could then be detected by running supernatant samples (derived from equivalent optical densities of culture) in SDS-acrylamide gels together with a wild-type control. The levels of secreted 23KD peptide could then be quantified by analysing the gel profiles using a scanning laser densitometer and the effect of the substitutions evaluated by comparing the levels of mutant 23KD HlyA peptide to the wild-type 23KD HlyA control.

8.2.1 Initial experiments to quantify the level of secretion of 23KD HlyA mutants:

Before embarking on quantification procedures it was first essential to demonstrate that cells infected with M13mp18-23KDHlyA,ClaI phage and expressing HlyB,D would secrete the 23KD peptide after induction of

expression. Thus, nutrient broth cultures of the F' carrying strain NM522 +/- pLG339B,D (*hlyB,D*), were infected overnight with a standard amount of the wild-type M13mp18-23KDHlyA,ClaI phage stock. Next day, the infected cultures were used to inoculate fresh nutrient broth samples to an A^{450} of 0.1. 23KD HlyA expression from the M13 *lac* promoter was induced at an A^{450} of 0.44 with IPTG (0.5mM final concentration). Supernatant samples were taken for SDS-PAGE analysis 2 and 4 hours post induction as described before (also see Materials and Methods). A single uninduced total cellular sample was taken together with cellular samples from both cultures 4 hours after induction, again for SDS-PAGE analysis.

The resulting gel is shown in Figure 3 and it is evident that the 23KD peptide is efficiently secreted from this system but only in the presence of HlyB,D (Lanes 6,7 versus 4,5). The supernatant samples in Figure 3 also show unmistakably that the cells were infected with M13 as indicated by the presence of large amounts of the M13 coat proteins running ahead of the 23KD HlyA peptide. There are also other bands present in the supernatant samples but these appear to represent a subset of cellular proteins, presumably released from the periplasm during M13 budding, since the major cytoplasmic, RNA polymerase bands (close to the top of the gel in the cellular samples) are absent. This gel also establishes that all the 23KD peptide in the supernatant samples from the strain expressing HlyB,D is of HlyA origin. The cellular samples, as usual, do not appear to accumulate the 23KD HlyA peptide, even after induction, in the presence or absence of export functions, confirming its intracellular instability.

8.2.2 Quantification of the level of secretion of 23KDHlyA mutants:

From the sequencing data all those mutants that carried substitutions were taken for quantification. These experiments were approached basically as described above. Usually, 100ul of each of 11 mutants in the form of M13mp18-23KDHlyA,ClaI and one wild-type phage stock was used to infect overnight, 1.5ml of nutrient broth seeded with NM522 carrying pLG339B,D (*hlyB,D*) (see Materials and Methods). The overnight cultures were used to inoculate 12ml of fresh broth to an A^{450} of 0.1, grown at 37°C and induced for 23KD expression at an A^{450} of ~0.7, with IPTG (0.5mM final concentration). Growth of the cultures was then continued for 2 to 3 hours before isolating supernatant samples for analysis by SDS-PAGE. Large, 15% SDS-acrylamide gels were used to

analyse the protein samples derived from 5 A^{450} absorbance units of TCA precipitated supernatant sample, and stained with Coomassie blue. The resulting gels were then scanned using an LKB ultrascan laser densitometer either directly or after drying down the gel (see Materials and Methods). The output obtained for each 23KD band scanned (given in relative absorbance units for each band detected) could then be directly compared against that obtained for the 23KD wild-type HlyA peptide to determine the relative effect of the substitution on secretion. Figure 4 shows a gel comprising samples derived from two such independent experiments (with mutant 24 analysed in both experiments to provide a basis for examining the reproducibility/accuracy of this quantification system). Table 2 shows the relative level of secretion of the 23KD mutant proteins compared to the wild-type control, by this initial quantification method.

8.2.3 Results of initial quantification procedure:

Figure 4 and Table 2 together indicates the wide range of effect that the various substitutions have on the level of 23KD HlyA secretion. Table 2 summarises the position of substitutions and their effects on secretion, from this initial experiment, indicating non-secretors (mutants 9 and 50 - as expected they encode stop codons), hyper secretors (mutants 69 and 71) with the remaining mutants secreting reduced or wild-type levels compared to wild-type 23KD peptide controls. It should be noted here that mutant 24 was quantified in two independent experiments and estimated in the first case, as 30% (average of two independent readings 28 and 33%, of the same sample quantified from two separate gels) and as 33% of the wild-type levels in the repeat experiment, attesting to the reproducibility of the system.

Again the supernatant samples contain the M13 phage coat protein and the subset of cellular bands associated with the M13 budding process (see Figure 4). A minor doubt with respect to this system of quantification, concerned the initial phage infection of the F' cells which was carried out in rich broth overnight and thus the maintenance of the F'(pro+) plasmid was not directly selected. Thus, over the infection period it is possible that some cells may lose the F' and would not be infected with the M13 mutant clone with consequent over estimation of an effect upon secretion. This is particularly important since cells not carrying M13 grow faster and out compete the infected cells. Also, the isolated phage stocks in these initial experiments were never checked for purity

although the sequencing reactions should have revealed any grossly impure stocks.

In view of these possible faults with the quantification system it was decided to go back to the original phage stocks and to re-infect NM522/ pLG339B,D (*hlyB,D*), but this time on F' selective media, and to purify for single colonies infected with a M13 phage. The infection with phage is easily confirmed by analysing a sample of overnight culture for the presence of M13 coat protein in the supernatant (usually indicated as a large usually coloured pellet when spun down after TCA precipitation). Once a single colony was confirmed to carry an M13 phage it was re-streaked on selective media and stored at 4°C, and used in all future quantification experiments. Thus, we can be confident that at the beginning of an experiment that all the cells in a culture carry M13 phage, even if some lose the F' during subsequent growth in rich broth during the quantification procedure.

8.2.4 Results of the second modified quantification procedure:

Overnight cultures of NM522/pLG339B,D (*hlyB,D*) infected with different M13mp18-23KDHlyA,ClaI mutant phage stocks were used for quantifying the effect of the substitutions on 23KD secretion as described in the previous section. The quantification results from Laser scanning for each mutant analysed relative to the wild-type control, is also shown in Table 2.

These results differ from those obtained in the initial quantification experiment in that the relative level of secretion of individual mutants was higher (except for the null and putative hypersecreting mutants). However, the overall trend was the same with the lowest secretors remaining the lowest. The reproducibility of this modified procedure was shown by quantifying the relative level of secretion of several of the mutants in two or three independent experiments (percentage error shown in brackets in Table 2) or by the quantification of 2 independently isolated mutants encoding the same single substitution (mutants 27 and 34). I can only suggest that the largely quantitative differences between the two procedures probably reflect much more efficient infection of the cells by M13 when purified, established "lysogens" were used as in the second procedure.

8.3 Effect of mutations on the secretion of the HlyA toxin - a test case:

It was decided to test whether the quantitative data for a given mutant measured by secretion of the C-terminal HlyA 23KD peptides could be duplicated when transferred into the intact HlyA molecule. The effect of substitutions on the secretion of the HlyA toxin can then be quantified by assaying for the secreted toxin activity compared to a wild-type control. This should tell us whether the quantification system used above truly reflects the effects that these substitutions have on the secretion of the intact HlyA molecule. Therefore, as test cases the 0.6Kb ClaI/KpnI fragment from mutant 19 (26+/-3% secretor) and 16 (~100% secretor) were substituted for a similar fragment from the pACYC-CAPro mutant (Ser^{996} to Pro, wild-type secretor but missing the unique *hlyA* BglII site; see Chapter 6; Section 6.3.2). The insertion of the mutant fragments was verified by screening for the re-introduction of the unique BglII site. The 0.6Kb ClaI/KpnI fragment from mutant 19 was recombined into the intact *hlyA* gene in two independent cloning procedures, generating pACYC-CA19.1 and pACYC-CA19.2. The wild-type pACYC-CA,ClaI and the three mutant plasmids (19.1, 19.2 and 16) were then transformed into NM522 carrying pLG339B,D (*hlyB,D*) and the growth curves and secreted haemolytic assay monitored as usual (see Materials and Methods). The results are shown in Figure 5. The supernatant samples were assayed in duplicate and no significant error was observed between the two samples (data not shown). From Figure 5 it can be seen that the cultures grew virtually identically while the haemolytic profile differed. As noted in Chapter 6, it was observed that the haemolytic activity of the first peak was variable from experiment to experiment, whereas the second peak could be accurately reproduced. Analysis of the differences in the level of secretion of the various toxin molecules at each time point in this second peak, taking wild-type levels as 100%, showed that the mutant 16 toxin was secreted between 70 to 100% of the wild-type levels, while both the mutant 19 HlyA molecules were secreted at significantly lower levels, between 26 and 55% of that of the wild-type over the final 2 hour period. The results of the two methods of assaying the effect upon secretion did vary therefore in these two cases although this was largely quantitative rather than qualitative. Therefore, it appears that the results obtained by densitometry are indeed generally valid and can be taken as a reproducible index of the level of secretion in a given mutant.

8.4 Preliminary interpretation of data obtained from the C-terminal HlyA mutants:

8.4.1 Features of particular regions or residues possibly required for efficient secretion of the HlyA peptide by the export machinery:

i) Null secretors:

It was expected that the majority of secretory null mutants isolated would be found to encode stop codons, and so far this is the case. Null mutants 36 and 50 also carry 3 other mutations encoding 2 further substitutions. Mutant 9 encodes the substitution, Ser^{999} to Thr, followed by an unexpected base deletion resulting in a frameshift and the generation of a stop codon 5 residues from the natural C-terminal end point. The null phenotype of this mutant is not solely due to the short truncate as Koronakis *et al.*, (1989) previously demonstrated that the final 8 residues are not required for secretion. It is worthwhile pointing out that the frameshift mutation changes the characteristic of the HlyA C-terminus from being "weakly" hydrophobic with a high incidence of both hydroxylated and small residues (Koronakis *et al.*, 1989) to that of a strongly hydrophobic C-terminus (data not shown). The percentage of small residues within this 24 residue frameshift region falls from 71 to 42%, hydroxylated residues from 30 to 10% whereas the number of non-polar residues is reduced from 70% to 40% (data not shown). Thus, it is not surprising that this mutant is not secreted as the whole nature of the C-terminus has been altered by the frameshift mutation.

Unfortunately, the mutants described above have not been very informative as we cannot access whether the observed abolition of secretion is due entirely to the deletion of a C-terminal region or to the additional alterations.

ii) Low Secretors (20 to 30%):

One of the most defective mutants so far isolated and quantified is mutant 25 which also encodes a stop codon (at Gln^{1002}) 22 residues from the C-terminal end. This result is surprising since it appears to contradict results presented by Koronakis *et al.*, (1989) who reported that removal of the final 21 residues virtually abolished secretion in their system. Quantification using the first assay system described (infecting an HlyB,D containing strain overnight with phage stocks) gave relative secretion levels of 10 and 12% in two independent experiments whereas the second method (isolating and using single colonies already infected with the mutant phage) gave a 20% level of secretion (see Table 2). Thus,

this result suggests that the C-terminal boundary of the signal motif or some essential feature of it, lies between residue 21 and 27 (as deletion of the final 27 residues abolishes secretion; Gray *et al.*, 1986; Koronakis *et al.*, 1989).

Both mutants 6 and 19 secrete at approximately 25% of the level of the wild-type 23KD HlyA peptide (28+/-3% and 26+/-3% respectively) and both carry a substitution involving Phe^{989}. In mutant 19 the Phe residue (large non-polar aromatic residue) is replaced by another large non-polar residue, Leu. In mutant 6 it is replaced by the large intermediate polarity aromatic residue, Tyr, but this mutant also encodes for a second substitution of Arg^{995} (large positively charged residue) to Lys, another large positively charged residue. As the nature of this second substitution is not altered it is probable that the defect is due entirely to the Phe to Tyr change. This is corroborated by the fact that a hydroxylamine generated C-T transition substituted the same Arg (995) for a Lys, without any obvious effect on the secretion of the 23KD HlyA peptide encoding the change (Chapter 7: Section 7.6.5). These results indicate that Phe^{989} is a critical residue required at this position since even its replacement with residues of apparently similar nature substantially reduced secretion.

It is worthwhile pointing out here that the replacement of Asp^{1009} (small negatively charged residue) by Arg (large positively charged residue) by SDM (See Chapter 6; Section 6.4.3) resulted in an approximately 70% reduction in HlyA secretion as determined by haemolytic assays. However, this residue in itself does not appear essential for recognition as deletion of the final 22 residues (mutant 25), and thus also the region carrying the Arg^{1009} residue, only reduced secretion (by about 80%). This could be interpreted to imply that the substitution may be influencing the efficiency of secretion by disrupting the overall structure of the C-terminal domain of HlyA resulting in inhibition of the proper presentation of the essential features of the signal motif.

iii) Reduced level secretors (31 to 70%):

Mutants 24, 53 and 46 were estimated to secrete 40 (+/-8%) , 53(+/-2%) and 61% of wild-type levels, respectively. All three mutants unfortunately carry multiple substitutions making it difficult to determine the contribution of each substitution to the secretion defect. However, mutant 53 has only two residue changes, one being a Glu^{993} (large polar, negatively charged residue) to Asp (small polar negatively charged residue). As, this substitution is present by itself in mutant 16 and does

not appear to alter secretion levels, the secretion defect in mutant 53 is probably due to the replacement of Asp^{990} (small polar, negatively charged residue) to Val (large non-polar residue). Asp^{990} was also replaced by a Ala (small non-polar residue) in mutant 68, but this time without any significant effect on secretion levels. These results suggest that the size of the residue at this position might be an important feature in secretion.

iv) Substitutions producing wild-type levels of secretion (71 to 120%):

The remaining 16 mutants were found to secrete in the range of 71 to 116% of wild-type 23KD HlyA levels and will be classified as wild-type in nature, allowing for errors in the quantification procedure. Nevertheless, the analysis of these mutants is informative in revealing those substitutions that can occur in certain positions without affecting secretion of the HlyA molecule.

v) Effects of single residue substitutions on secretion (Phe^{989} to Gln^{1002})

$\underline{Phe}^{989}$

See section ii) above.

$\underline{Asp}^{990}$

The small negative character of this residue has been altered by substitution to Asn (small polar, mutant 50), Val (large non-polar, mutant 53) and Ala (small non-polar, mutant 68). Whereas, the effect on secretion of the first replacement cannot be assessed due to the presence of multiple substitutions (including a stop codon), the replacement by a small non-polar (Ala) residue does not appear to significantly affect secretion. However, the change to a Val (large non-polar) reduces secretion by about 2 fold, as the possible effect of a second substitution (Glu^{993} to Asp) can be neglected as the presence of this substitution alone does not alter the level of secretion of the 23KD peptide encoded by mutant 16. Thus, it would appear that the size of the residue at this position can influence the level of secretion whereas the polarity or the charged characteristic of the residue at this position do not appear to be essential for efficient secretion.

Val^{991}

Analysis of the mutants altered for this residue revealed that this Val residue (large non-polar) can be deleted (mutant 67), substituted for either a Leu (large non-polar mutant, 69), Ala (small non-polar mutant 59) or Ile (large non-polar, mutant 32) without apparently altering secretion levels. These results suggest that size is not an essential feature of this residue, if required at all, for recognition/secretion. The presence of a second substitution in mutant 69 (Leu^{1000} to Ile) has to be noted. However, this double mutant appears to be wild-type in secretion levels and the substitutions are very similar in nature.

Lys^{992}

This Lys (large positively charged residue) was found to be replaced by Gln (large polar residue, mutant 71), Asn (small polar residue, mutant 58) or Thr (small non-polar residue, mutant 4) all without significantly altering secretion levels. Again, this indicates that there is no particular feature (size, charge or polarity) required in this position for signal recognition and indeed may not be important in the process.

Glu^{993}

The only informative substitution of this large negatively charged residue was to another negatively charged but small residue, Asp (mutant 16) again without any detectable effect on secretion suggesting that at least the size of the residue at this position is not critical for secretion.

Glu^{994}

The base triplet encoding this large negative residue was only mutated twice, so far, generating a stop codon (mutant 50), or a small positively charged residue, Asp (mutant 24). Unfortunately, both of these mutants also carry other substitutions so are not very informative. However, a previously engineered site directed mutagenesis experiment converted this Glu^{994} residue (large negatively charged) to a Lys (large positively charged residue) with apparently no effect on secretion (see Chapter 6; Section 6.4.3). Thus, the nature of the charge of this residue at this position, at least does not appear to be essential for secretion.

Arg^{995}

Not very much information was obtained on the substitutions permitted at this position, since the only replacement, obtained was to a similar residue, Lys in mutant 6 which also carries a second substitution Phe^{989} to Tyr with an overall 75% reduction in secretion levels. Indeed, this

former replacement (Arg^{995} to Lys) was fortuitously duplicated in a previous hydroxylamine mutagenesis reaction without any apparent effect on 23KD HlyA peptide secretion (Chapter 7; Section 7.6.5). Obviously, the isolation and analysis of more mutants at this position is required to define any characteristics of this residue required, if at all, for secretion.

Ser^{996}

This is the only codon in the saturation mutagenesis reaction that, so far, has not being mutated, and this could be due to its position at the very end of the 21mer oligonucleotide making the binding of the "mismatch" tail difficult. However, another engineered site specific mutagenesis reaction was employed previously to test whether this small polar residue could be substituted for Pro (small non-polar residue with alpha-helix disrupting characteristics; see Chapter 6; Section 6.3). Again, analysis of the effect of this substitution on HlyA toxin secretion using haemolytic assays did not reveal any significant alteration in the secretion of the mutant compared to the wild-type control. Thus, neither the polar characteristic of this residue nor its potential to be incorporated into a predicted alpha helical structure seems to be critical for recognition and thus secretion.

Ala^{997}

This small non-polar residue can apparently be replaced by either a small polar Ser residue (Mutants 22 or 28) or the small negatively charged residue, Asp (mutant 27 or 34) without altering secretion levels significantly. This suggests that the non-polar nature of this residue is not important for targeting. Mutants 36 and 24 also encode the replacement of Ala^{997} for the small non-polar residue, Pro, but as they both carry multiple substitutions the observed reduction in secretion cannot be attributed to any one substitution. The requirement for small residues at this position has not, as yet, been challenged and this will therefore require the isolation and analysis of more mutants.

Ala^{998}

Of the mutations isolated at this codon position three are silent changes with others being substitution for other small residues. In the first case (mutant 26) the small non-polar Ala^{998} residue is converted to a small polar Ser, with no effect on secretion levels, indication a minor role, if any, for the polar character in secretion. In the second case (mutant 46) the substitution is to a small non-polar residue, Pro. However, this mutant also encoded changes of Ser^{999} to Cys and Gln^{1002} to Arg with

an overall reduction of secretion to about 40% of wild-type. No conclusions can therefore be drawn due to the nature (multiple substitutions) of this mutant.

Ser^{999}

Mutants 24, 46 and 28 encode substitutions of the small polar Ser^{999} residue for small non-polar residues Ala and Cys and Thr, respectively. Both mutants 24 and 46 are reduced secretors but also carry multiple (3) substitutions and thus the effect on secretion cannot yet be assigned to specific substitution event. However, mutant 28 only carries two substitutions, the Ala^{999} to Thr and Ala^{997} to Ser. This latter substitution by itself has no apparent effect on secretion as evidenced by mutant 22 and thus the Ala^{999} (small polar) to Thr (small non-polar) replacement can be taken to have no significant effect on secretion. Thus, it seems as if the presence of either small polar or non-polar residues are permitted at this position.

Leu^{1000}

This residue (large non-polar) has so far only being replaced by other large non-polar residues Val (mutant 50) and Ile (mutant 69). Unfortunately, the former substitution is only one of four including a stop codon so is non informative. However, mutant 69 although carrying two substitutions (Leu^{1000} to Val and Val ^{991} to Ile) is "wild-type" for secretion suggesting that the replacement of either of these residues with other similar residues has no effect on secretion.

Leu^{1001}

The encoding triplet has been mutated so far 4 separate times, two generating silent substitutions (data not shown) the third replacing the large non-polar Leu^{1001} by another large non-polar residue, Met (mutant 47) and the fourth to a large polar residue, Gln (mutant 40). The latter substitution can only be, and was, generated by a rare double mutation in this single triplet codon. Both of these single substitutions do not appear to alter secretion levels suggesting that the polarity is not important at this position although as above the effect of placing small residues at this position has not yet been tested.

Gln^{1002}

Only two substitutions have been obtained so far at this position, one the generation of a stop codon (mutant 25) and the other the replacement of the large polar Gln with a another large polar but positively charged

residue, Arg (mutant 46). Mutant 46 is a reduced secretor (~60% of wild-type levels) but unfortunately carries two other replacements as well so the effect cannot yet be assigned to one substitution.

In summary, the major single substitutions having a significant effect upon secretion were alteration of Phe^{989} the adjacent Asp^{990}, Asp^{1009} and Glu^{978} (see below). The substitutions of the Phe^{989} residue are predicted to either disrupt (mutant 19) or extend (mutant 6) the short predicted weak turn region. The substitution of Asp^{990} to Val is also predicted to shorten the predicted turn region, by introducing a beta-sheet forming potential. The Asp^{1009} to Arg (SDM mutant) is predicted to disrupt the long stretch of turn forming residues, while the Glu^{978} substitutions weakened (mutant A; see below) or converted the small alpha-helical forming region to predicted beta sheet formation (mutant C; see below). Although these single substitutions do appear to disrupt predicted secondary structures these features, in themselves, do not appear to be essential for secretion as they have also been disrupted by other substitutions without any apparent effect on secretion levels (see below) suggesting an important role for the character of these particular residues in secretion.

8.4.2 Effect of disrupting "conserved" features within the C-terminus of related HlyA molecules on secretion:

i) Strong alpha-helix Asp^{990} to Leu^{1001}:

The C-terminal 53 residues of each of the mutants was analysed by the suite of secondary structure prediction programmes available on the Leeds Biovax system for mutants altering the predicted strong alpha-helical structure (see Chapter 6; Section 6.1). An insight into a possible functional role for the strong alpha-helix in the secretion process was obtained from a number of mutants. Thus, while mutants 24 and 46 are predicted to either disrupt or shorten the predicted helix (see Figure 6) and do reduce 23KD HlyA peptide secretion by 50-60% of wild-type level, they also carry multiple substitutions which might be the source of the defect. However, mutants 22, 26, 27 ,28 and the Ser^{996} to Pro mutant obtained by SDM do not significantly affect secretion although they are predicted to either disrupt, or weaken this predicted helical structure arguing against an important role in secretion. Thus, it would appear that the formation of a strong helical structure in this region.

ii) Charged -O+--+ region (Asp^{990}-Arg^{995})

a) Conversion of -O+--+ to **O**O+--+

Alteration of this -O+--+ charged region to OO+--+ (mutant 53) reduced secretion by ~53% whereas another similar charge change (mutant 68) had no apparent effect. This suggests that it is not essential to have a negative charge at this position (Asp^{990}), though the nature of the residue influences secretion (see Section 8.4.1).

b) Conversion of -O+--+ to -O**O**--+

Mutant 4, 58 and 71 all carry single substitutions changing the character of this charged region to -OO--+, and do not alter the level of 23KD peptide secretion. Therefore, it would appear that the presence of a positively charged residue at position 992 is not essential for secretion.

c) Conversion of -O+--+ to -O+-++

By site directed mutagenesis the negative character of residue Glu^{994} was changed to a positive Lys without any significant effect on secretion. This result only implies that a negative charge is not essential at this position, and the replacement of the charged character has yet to be tested.

d) Alteration of overall charge in region 990-995

As seen above the overall -1 negative charge in this clustered charged region has been converted to 0 , -2 and +1 (see a, b and c above) without any significant effect on secretion of the 23KD HlyA peptide, indicating a non-crucial role for the overall charge of this region in the secretion process.

iii) "Aspartate Box" Glu^{994} to Asp^{1009}

Initially the "Aspartate box" was defined as a group of 12-14, mainly small uncharged residues flanked by two negatively charged residues, usually Asp (see Chapter 6). As shown in Chapter 6 the replacement of the small 5' "aspartate" (Glu^{994}) for the large positive Lys residue had no apparent effect on secretion. However, the conversion of the 3' Asp (1009) to the large positive Arg residue reduced secretion to ~70% of wild-type levels. However, the attempted introduction of large residues into the " aspartate" box was not successful (see Chapter 6; Section 6.4.1) so this remained untested. Nevertheless, the above results together with the later realisation that the related *Proteus* and *Morganii* HlyA molecules do not conserve the 3' Asp residue rule out this feature, as described above, as being important in secretion.

This initial saturation mutagenesis study only probed the region encoding the first 8 of the intervening 14 "aspartate box" residues (see Figure 1). In addition, analysis of all the predicted possible substitutions within this region arising from single base mutations, surprisingly, revealed a limitation on the number and nature of the possible substitutions (see Figure 7). Indeed, it appears that an alternative strategy will have to be pursued in order to alter the nature of the residues in this region in order to obtain more radical changes and therefore to elucidate the role of these residues in targeting.

iv) Amphiphilic Helix Leu^{972} to Phe^{989}

Although the region encompassing the predicted amphiphilic region (Koronakis *et al.*, 1989) was not investigated in the above study, preliminary analysis of this region was carried out by Sally Taylor (MSc. student, 1990; this laboratory) using a strategy identical to that described above. The five mutants (A-E) isolated and analysed by her are also shown in Table 2.

Mutant A (12% secretor) replaces Glu^{978} (large negative residue) for Lys (large positive residue) making the molecule defective for efficient secretion. However, this change does not alter the amphiphilic nature of this region (see Figure 8), since Glu^{978} was already the only polar residue in the non-polar face. This result implies rather that it is the positively charged nature of this substituting residue that is hindering secretion.

Mutant B (14% secretor) surprisingly carries multiple substitutions (5) and so the defect in secretion cannot be assigned to one event. However, these substitutions do alter the amphiphilic nature of this region by inserting polar and non-polar residues into the non-polar and polar faces, respectively (see Figure 8).

Mutant C (51% secretor) has a single substitution, altering the same residue as in A (above) but this time from Glu^{978} (large negatively charged residue) to a large non-polar residue, Val. Indeed, this alteration increases the amphiphilic nature of this region by replacing the single polar residue in the non-polar face of the amphiphilic helix with a non-polar residue, nevertheless this results in reducing secretion levels. This result is difficult to square with any crucial role for an amphiphilic helix

per se and further emphasises that the negative character of the Glu^{978} residue is crucial for secretion.

Mutant D (59%) is a double mutant Ile^{983} to Ser (large non-polar to small polar) and Ser^{984} to Ala (small polar to small non-polar). This mutant therefore results in the introduction of a polar and a non-polar residue into the non-polar and polar faces of the predicted amphiphilic helix, respectively (see Figure 8). Such changes might have been expected to have a more marked effect on function if the formation of an amphiphilic helix were a crucial factor in targeting.

Mutant E (Wild-type secretor) has Ile^{976} to Thr (large non-polar to small non-polar) suggesting that at least the size of the residue at this position is not essential for efficient secretion. This mutation does not alter the amphiphilic potential of this region (see Figure 8).

Although the amphiphilic nature of this region is not in question there are doubts whether this region is able to adopt an 18 residue long amphiphilic helix since the suite of eight secondary structure prediction programs (available a Leeds. BioVax system) used in this study predict a quite different secondary structure (see Figure 1). However, it is possible that a helical structure may be adopted if inserted into a hydrophobic membrane environment. Nevertheless, some of the amino acid substitutions described above cast serious doubts as to whether an amphiphilic helix has a role in the targeting signal. Further studies of the essential nature of this region will in any case require the isolation and characterisation of more single substitutions.

8.5 Discussion:

This chapter has demonstrated the success of the "saturation" mutagenesis procedure in introducing random point mutations into a specific region of interest. There are, however, some limitations imposed by the nature of the codon table, which in particular in the region of the "aspartate box" precludes many radical amino acid changes based on single base substitutions (see Figure 7). Nevertheless, the results obtained from the analysis, so far, of only 25 mutants which carry substitutions has been very instructive in identifying at least 3 individual residues which are important for function and in providing strong arguments against any essential role in secretion for several secondary structural or other features which had been predicted to be present in the C-terminal region of HlyA. The analysis has also been invaluable in

further defining, in contrast, to more ambiguous deletion/ fusion analysis, the extent or borders of the targeting motif. Thus, the ability of a C-terminal 23KD HlyA mutant truncated for the final 22 residues (mutant 25) to be secreted at significant levels suggests that the final 22 residues do not encode any major "critical" features, necessary for secretion. This, contrasts with earlier studies that the deletion of the final 27 residues virtually abolishes secretion (Gray *et al.*, 1986; Koronakis *et al.*, 1989). Together, these results indicate that the right hand border of the signal motif lies close to the Gln^{1002} residue 22 amino acids from the C-terminus. Data from Mackman *et al.*, (1987) and Koronakis *et al.*, (1989) had placed the left hand border of the targeting signal motif within the final 27 or 53 residues of HlyA. From the mutagenesis data described above it now appears that the signal motif extends at least to Glu^{978}, 46 residues from the C-terminus.

Data obtained from this initial use of saturation and site directed mutagenesis, has revealed a wide spectrum of secretion defective mutants ranging from null secretors to wild-type in character, over the entire signal encoding region. Unfortunately, all the null secretion mutants isolated, so far, carry stop codons, deletion or frameshift mutations, suggesting that the targeting motif is very tolerant of substitutions. The isolation of mutants throughout a stretch of 46 C-terminal residues gave rise to a number of strongly defective mutants (Glu^{978} to Lys; Phe^{989} to Leu or Tyr and Asp^{1001} to Arg) reducing secretion by 70 to 90% of wild-type levels. These results suggest that these residues are critical for interaction with the membrane translocator and hence efficient secretion. The spread of other substitutions reducing secretion to lesser degrees, or not at all, indicated that the role of these residues may not be an essential one and are therefore presumably more concerned with the overall "folding" of the C-terminus, permitting the correct presentation of the "critical" residues for interaction with the Hly export protein(s).

The proposition that the recognition signal is composed of a subset of "crucial" residues within the overall "extended" structure of the C-terminal region is supported by the inability to identify any conserved "targeting" features between the wide range of molecules now shown to carry C-terminal targeting signals, which in several cases have been found to use the HlyB,D pathway, albeit with varying efficiencies (see Chapter 6; Section 6.2). Indeed, the apparently "promiscuous" nature of the Hly export proteins for different C-terminally targeted molecules, with apparently no conserved features, also suggests the existence of a subset

of critical residues maintained for interaction/ recognition with the export proteins. Thus, it would appear that the overall C-terminal structure is flexible, having diverged with time in other species, but nevertheless has maintained some residues/features crucial for targeting, which can still be recognised by more distantly related export proteins. From the results of the mutagenesis studies described here these elements are more likely perhaps to be individual or clusters of key residues rather than a specific secondary or tertiary structure. However, one should perhaps exercise some caution in the use and interpretation of secondary structures prediction programmes as the methods presently available are only considered to be 60-70% accurate but with an apparently reduced ability to correctly predict secondary structures at the C-terminal region of proteins (E. Eliopoulos; Dept. of Biophysics, Leeds Univ. Personal communications)

Amino acid substitutions which reduce secretion levels can conceivably be acting at different levels, by either a) reducing the intracellular stability of the HlyA molecule b) affecting the recognition of HlyA by the translocator or c) interfering with the actual translocation process itself subsequent to recognition. Since the intracellular 23KD HlyA peptide is intrinsically extremely unstable, whilst the normal secretion process is thought to occur almost instantaneously with synthesis (Felmlee and Welch, 1988), if transport is blocked it is therefore unlikely that "stability" mutants would affect the secretion rate. However, the peptides once secreted if rendered more unstable by substitutions may breakdown extracellularily although one would expect to see the tell-tale breakdown products, which have not been observed. At the moment it is not possible to differentiate between substitutions affecting recognition and those affecting the translocation process although it should be possible to distinguish between these two classes by designing "competitive" experiments. Thus, mutants inhibited at the translocation process would be expected to also inhibit the secretion of, say, a hybrid molecule carrying a secretable passenger molecule fused to the HlyA wild-type signal domain. In contrast, recognition defective mutants should not inhibit the secretion of other hybrid molecules carrying the wild-type signal. However, at this stage it is not important to differentiate between the two classes as both are vital for the identification of all the requirements needed for efficient secretion.

The inability to obtain mutants that increase the secretion of the molecule significantly above wild-type levels argues that the secretion process is

already tuned for working at maximal rates with the number of export channels probably being the limiting factor as suggested previously (Holland *et al.*, 1990a; 1990b). The screening of further mutants of the signal domain may nevertheless still yield mutations of this class.

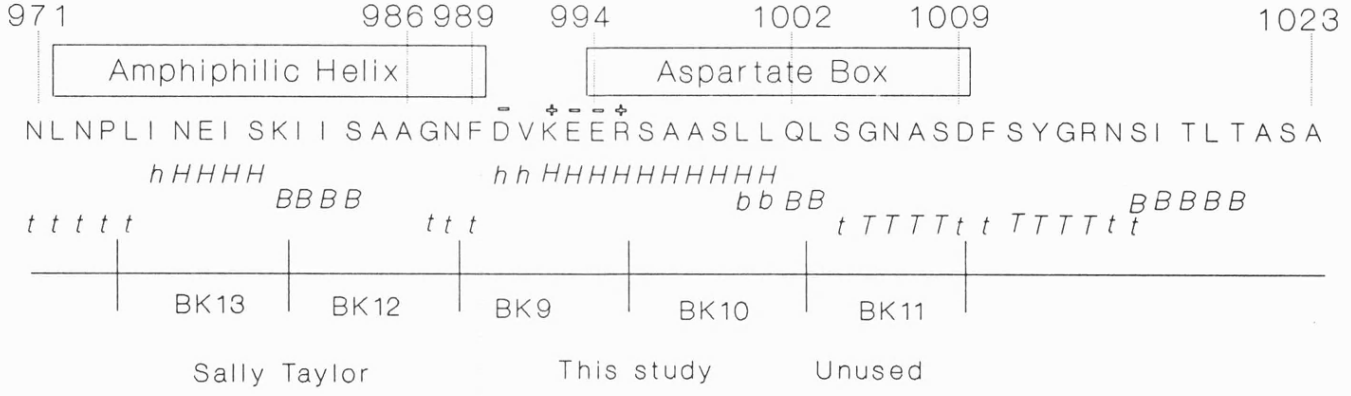
The "saturation" mutagenesis study, so far, has not identified any structural features essential for the secretion process. On the contrary, these studies appear to have ruled out any vital role for the predicted strong alpha-helix or the cluster of charges for recognition or efficient secretion. Other data is also incompatible with a putative role for the predicted amphiphilic helical region although the isolation and evaluation of more mutants (encoding single substitutions) in this region must be done before a final assessment can be made. On the other hand this does not appear to be a conserved motif in the more distantly related molecules and the majority of analytical programmes do not predict an extended helical structure even in HlyA. Amino acid substitutions obtained by mutagenesis in this study have also ruled out a specific role for the 5' negatively charged residue of the so called aspartate box whilst isolation of a charged Asp residue within the box also had no effect upon secretion. In addition, it is now evident that the 3' Asp residue is not always conserved in closely related haemolysins. These data all question the importance of this motif in promoting secretion although the effect of altering the size characteristic of residues within this region remains to be tested.

Although the isolation and identification of substitutions that do not affect secretion have been and are very informative in themselves, the screening process to identify such mutations is very slow. In the future it might be better to concentrate only on the isolation of defective, preferably null mutants for sequence analysis, since the analysis so far suggests that these might be quite rare. In this way it may be possible to establish the distrubution of hot spots or key residues involved in secretion which in turn will lead to the elucidation of the structure and function of the C-terminal signal. For the future also is the need if possible to divide the *sec*$^-$ mutations into those altering features required for recognition as opposed to those required for the translocation process, perhaps by the competitive experiments mentioned earlier ?.

Figure 1:

C-terminal region of HlyA to be probed by "saturation" mutagenesis strategy.

The C-terminal 53 residue region of the LE2001 HlyA protein is shown together with "conserved" features in related haemolysin toxins together with predicted secondary structures. The segmented line below the sequence indicates the regions to be probed by the various "randomised" oligonucleotides (see text).



Regions covered by synthesised "random" oligonuceotides

Table 1:

Sequencing data obtained from 72 transformants from site directed mutagenesis reaction using oligonucleotides BK9 and BK10.

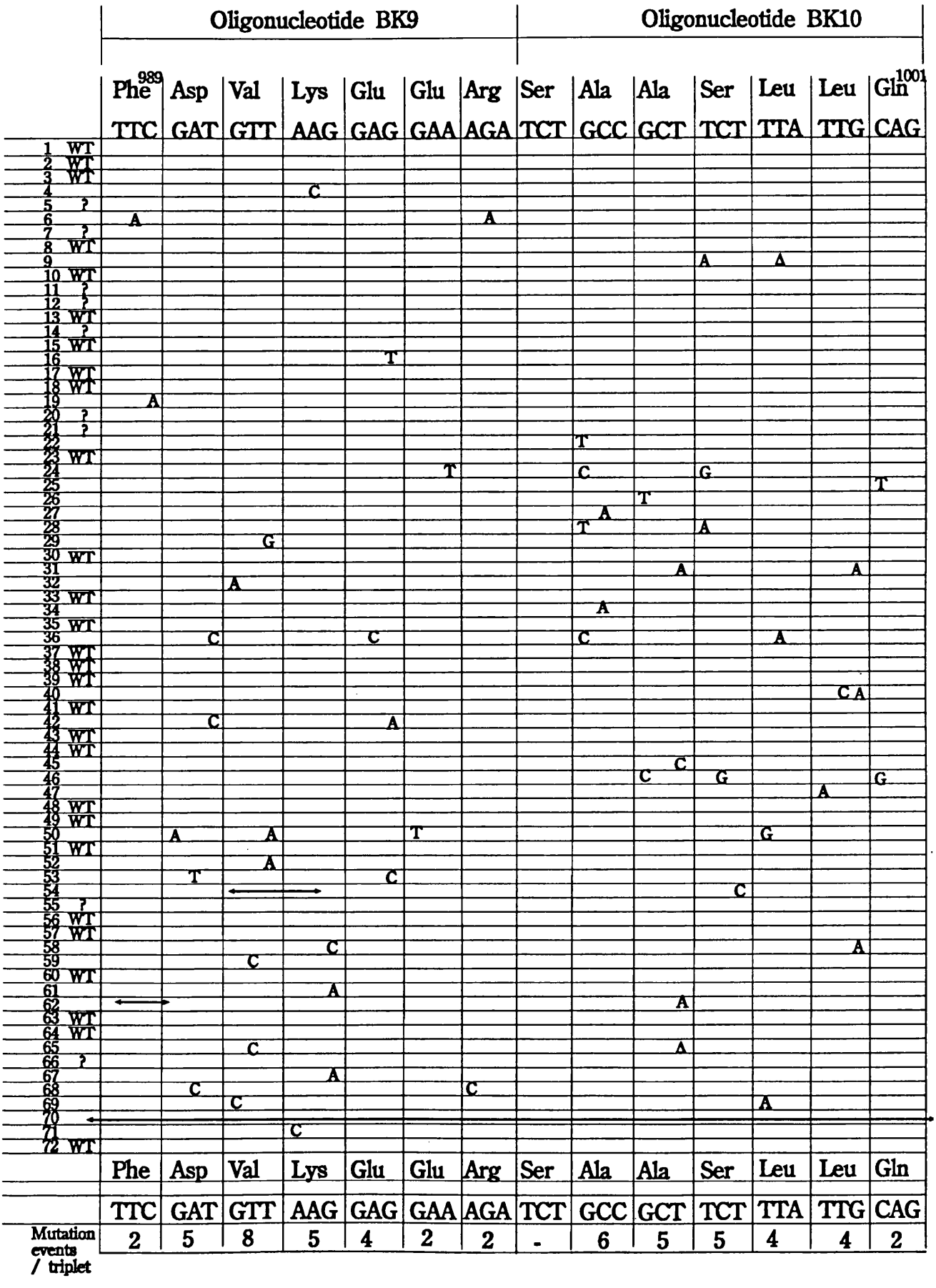

Oligonucleotide BK9
Oligonucleotide BK10
989
1001
Phe Asp Val Lys Glu Glu Arg Ser Ala Ala Ser Leu Leu Gln
TTC GAT GTT AAG GAG GAA AGA TCT GCC GCT TCT TTA TTG CAG
WT
?
Mutation events / triplet
2 5 8 5 4 2 2 - 6 5 5 4 4 2
Extent of deletion
Deleted base
Wild-type sequence
Sequence unclear

Table 2:

Distribution of substitutions and quantification of secretion levels.

The number of the mutants quantified in this study, carrying one or more substitutions, is given on the left hand side together with those generated by Sally Taylor (designated A to E). The level of secretion, compared to wild-type levels, for a series of C-terminal deletion mutants, generated by Colin Hughes (Koronakis et al., 1989) and specific site directed mutants generated earlier (Chapter 6; this study) are represented by the coloured numerals on the right hand side. The other two columns on the right hand side indicate the secretion levels of the mutants generated in this study estimated by the initial (left most column) and the revised quantification strategy, respectively (see text for details). The error margin (percentage) is also given (in brackets) for those mutants whose effect on secretion was re-quantified in other independent quantification analyses.

Preliminary Data on Level of Secretion of C-terminal Mutants

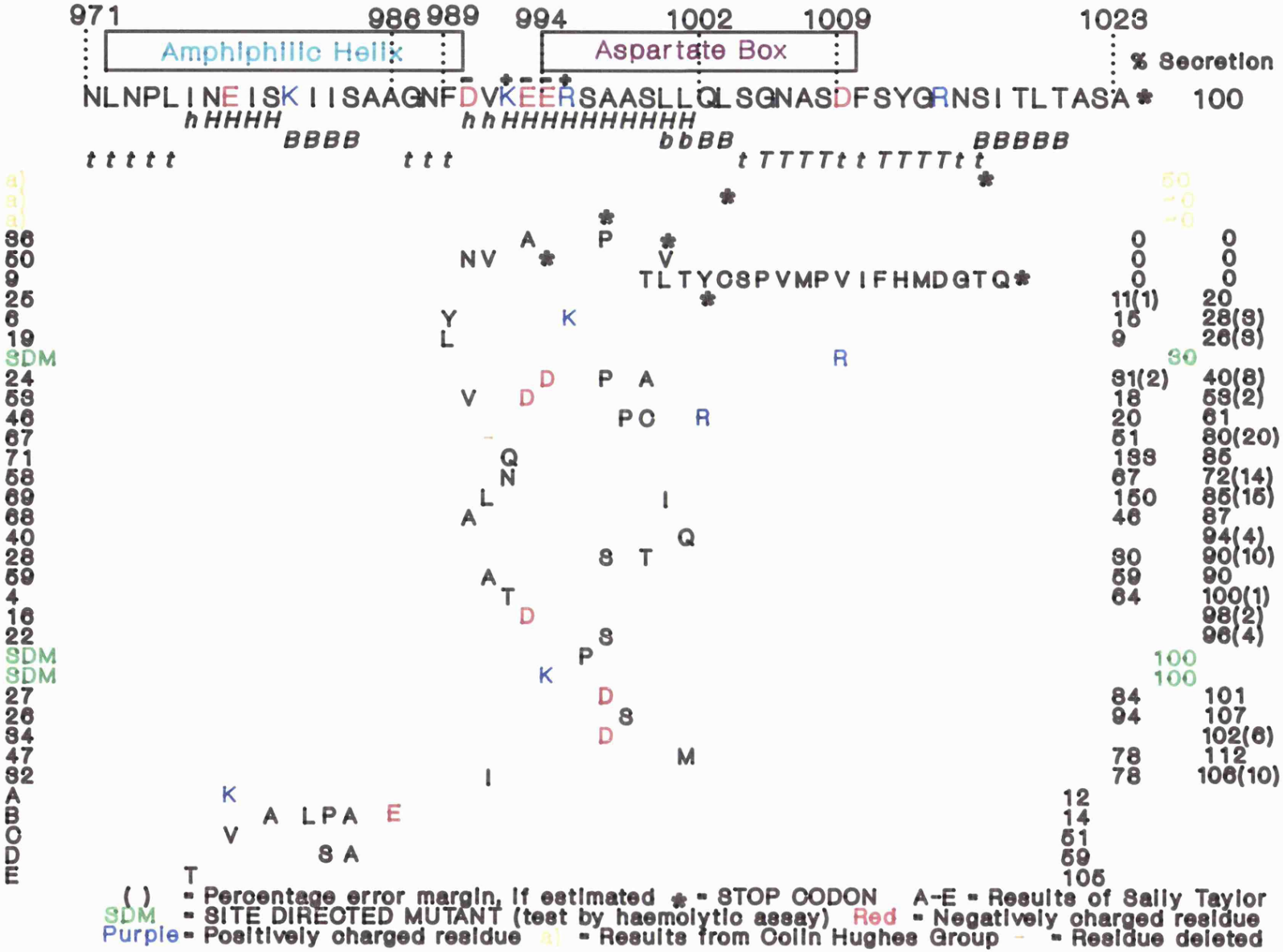


Figure 3:

Secretion of the 23KD HlyA peptide expressed from M13 in a HlyB,D dependent manner.

Cultures of NM522 +/. pLG339B,D (*hlyB,D*) were infected overnight with M13mp18-23KDHlyA,ClaI and used to inoculate fresh nutrient broth to an A^{450} of 0.1. Growth was maintained at 37oC and 23KD HlyA expression induced from the lac promoter at an A450 of 0.44 by the addition of IPTG (0.5mM final concentration). Supernatant samples were taken for SDS-PAGE analysis 2 and 4 hours post-induction. Only a single cellular sample was taken before induction from the NM522/M13mp18-23KDHlyA,ClaI culture together with cellular samples extracted 4 hours post-induction. 0.1 unit of cells (where 1 unit is defined as 1ml at A450 = 1) and 1 equivalent units of supernatant sample were analysed by mini-SDS-PAGE (15%) and visualised by Coomassie blue.

Cellular samples: lane 1,2, time = 0 and 4 hours post induction, respectively from the culture **minus** pLG339B,D (*hlyB,D*) and lane 3, time = 4 hours post induction, from the culture **plus** pLG339B,D (*hlyB,D*).

Supernatant samples: lane 4,5, time = 2 and 4 hours post induction, respectively from the cultures **minus** pLG339B,D (*hlyB,D*) with lanes 6,7, time = 2 and 4 hours post induction, respectively from the culture **plus** pLG339B,D (*hlyB,D*).

Arrows marks the position of the 23KD HlyA peptide and M13 coat protein, with molecular weight markers also indicated.

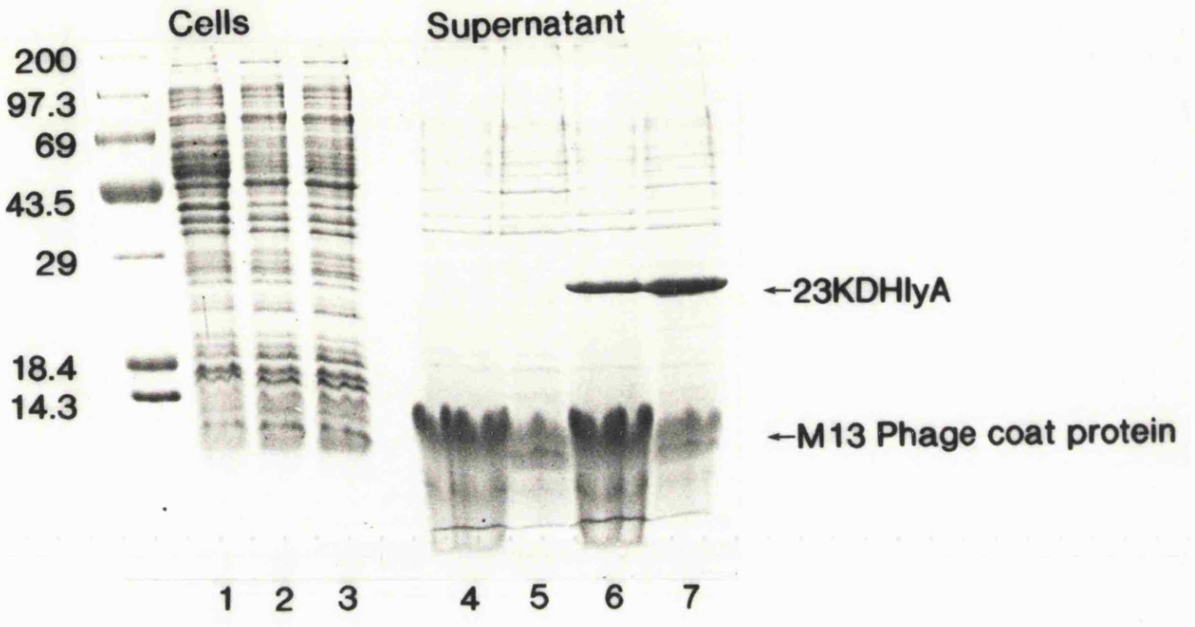

Cells
Supernatant
200
97.3
69
43.5
29
18.4
14.3
←23KDHlyA
←M13 Phage coat protein
1 2 3 4 5 6 7

Figure 4:

Supernatant profiles from cultures expressing HlyB,D and different M13mp18-23KDHlyA,ClaI mutants.

NM522 carrying pLG339B,D (*hlyB,D*) was infected overnight with different M13mp18-23KDHlyA,ClaI mutants and an M13mp18-23KDHlyA,ClaI wild-type as a control, and used to inoculate fresh nutrient broth to an A450 of 0.1. Growth was maintained at 37oC and 23KD HlyA expression induced from the lac promoter at an A450 of ~0.7 by the addition of IPTG (0.5mM final concentration). Supernatant samples were taken for SDS-PAGE analysis 3 hours post-induction and concentrated as usual. 5 equivalent units of supernatant sample were loaded (where 1 unit is defined as 1ml at A450 = 1). Proteins were analysed by SDS-PAGE (15%) and visualised by Coomassie blue.

The results from two independent experiments are shown with the WT (wild-type derived sample) being the first sample from both set of experiments, with the number of the mutant analysed also indicated.

Again arrows marks the position of the 23KD HlyA peptide and M13 coat protein, with molecular weight markers also indicated.

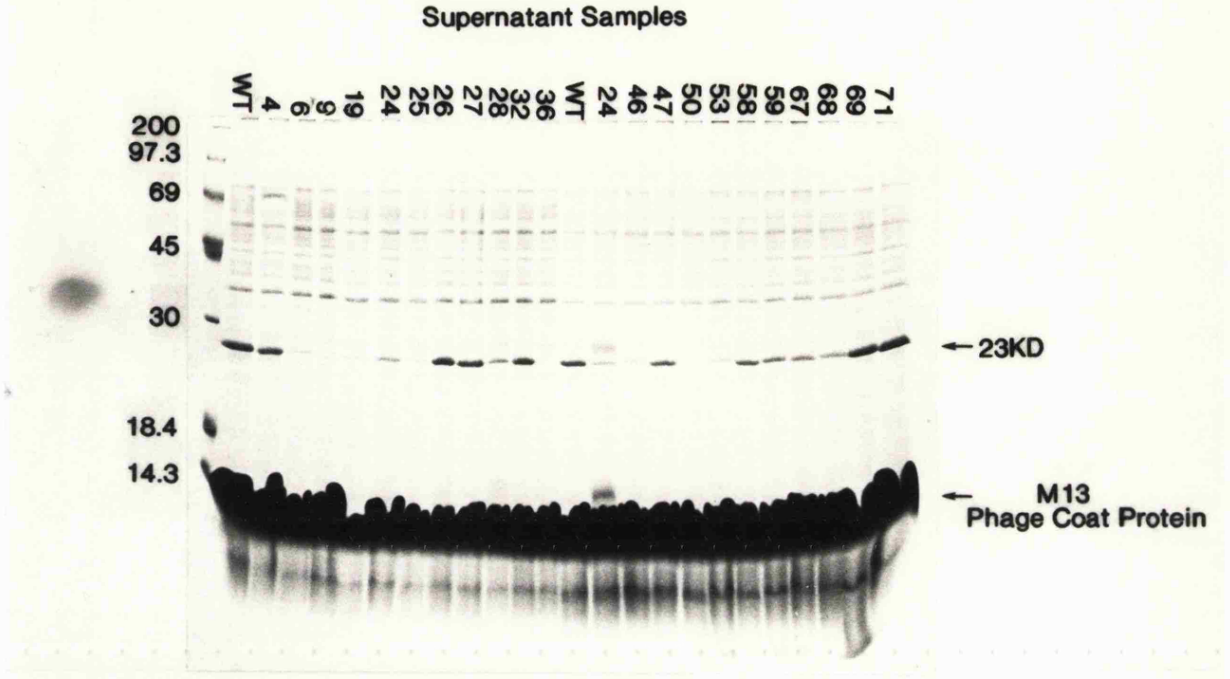

Supernatant Samples
WT
4
6
9
19
24
25
26
27
28
32
36
WT
24
46
47
50
53
58
59
67
68
69
71
200
97.3
69
45
30
18.4
14.3
23KD
M13
Phage Coat Protein

Figure 5:

Effect of substitutions on the secretion of the intact active HlyA toxin.

The mutations present in 3' portion of *hlyA* in mutants 19 and 16 (see Table 2) were transferred into the intact *hlyA* gene in pACYC-CA,Pro (see text). The mutant 19 *hlyA* fragment was exchanged for the pACYC-CA,Pro fragment in two independent cloning reactions generating pACYC-CA19.1 and 19.2 respectively. NM522 cultures carrying pLG339B,D (*hlyB,D*) and the wild-type or mutant pACYC-CA,ClaI plasmids were initiated in Luria broth (supplemented with 10mM CaCl2 and appropriate antibiotics) at an A450 of 0.2 and samples removed at 40 to 60 minute intervals for optical density and haemolytic assay readings (see Materials and Methods). The release of haemoglobin from red blood cells by the secreted HlyA toxin was monitored at A^{543} with the haemolytic activity given in arbitrary haemolytic units (H.U). H.U. indicates the A^{543} reading obtained per A^{450} unit of cells.

Solid lines depict the optical density (A^{450}) while the dashed lines show the supernatant haemolytic activity of the cultures.

Haemolytic assay of Mutant versus Wild-type HlyA molecules

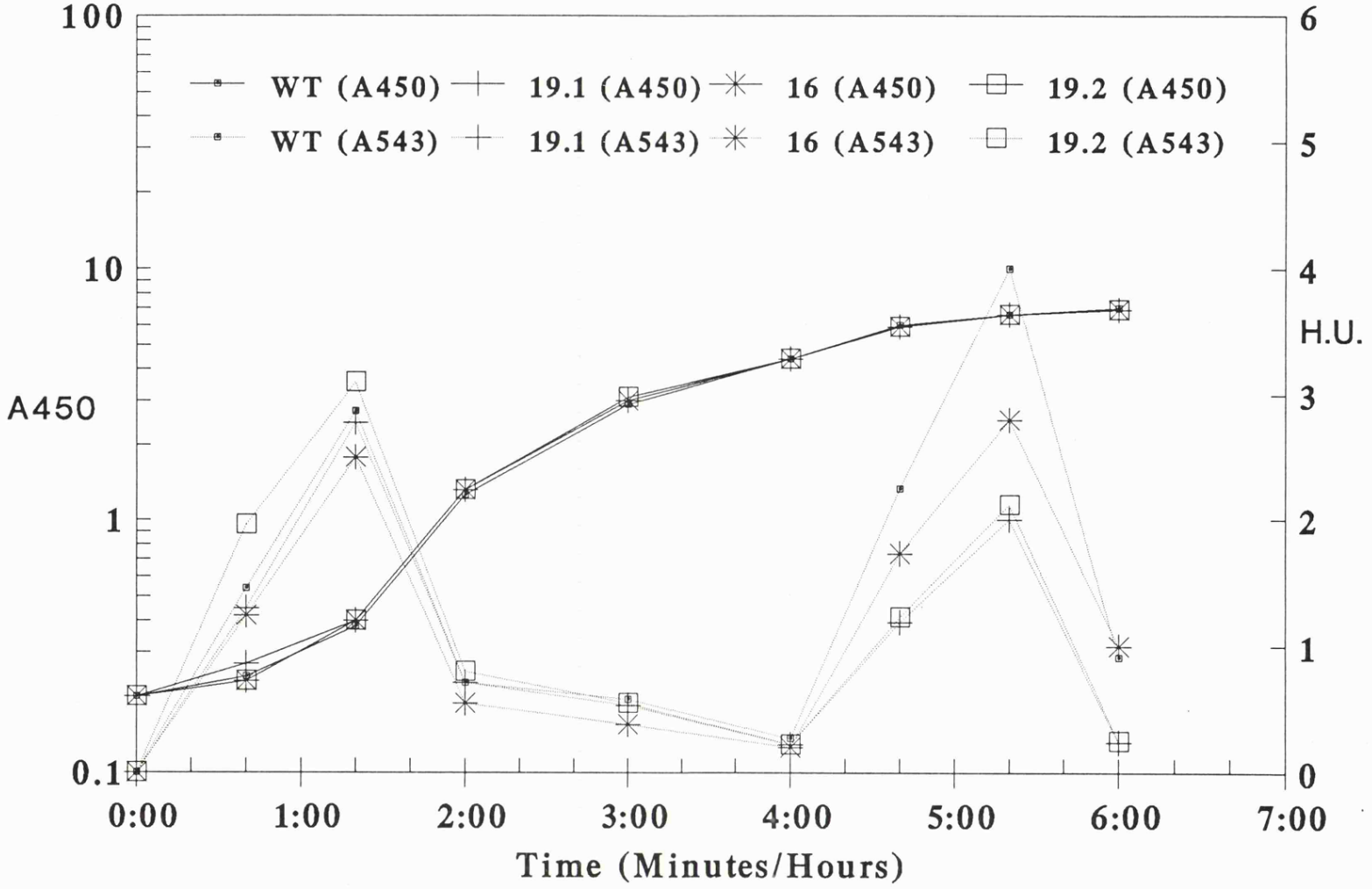


Figure 6:

Substitutions predicted to disrupt/weaken the putative strong alpha-helix (Asp^{990} to Leu^{1001})

The number of the mutant is given together with the estimated level of secretion compared to a 23KD HlyA wild-type control. The substitutions encoded by the mutants altering the predicted helical region are highlighted in bold together with changes in the prediction secondary structure in this region. The secondary structure predictions were again obtained from a suite of 8 secondary structure prediction programmes, available on the Leeds Biovax system (see Materials and Methods). h,b and t represent predicted helical, beta sheet and turn forming residues. Capital letters indicate predicted by at least 5 out of the 8 programmes and small letters less that 5 out of eight.

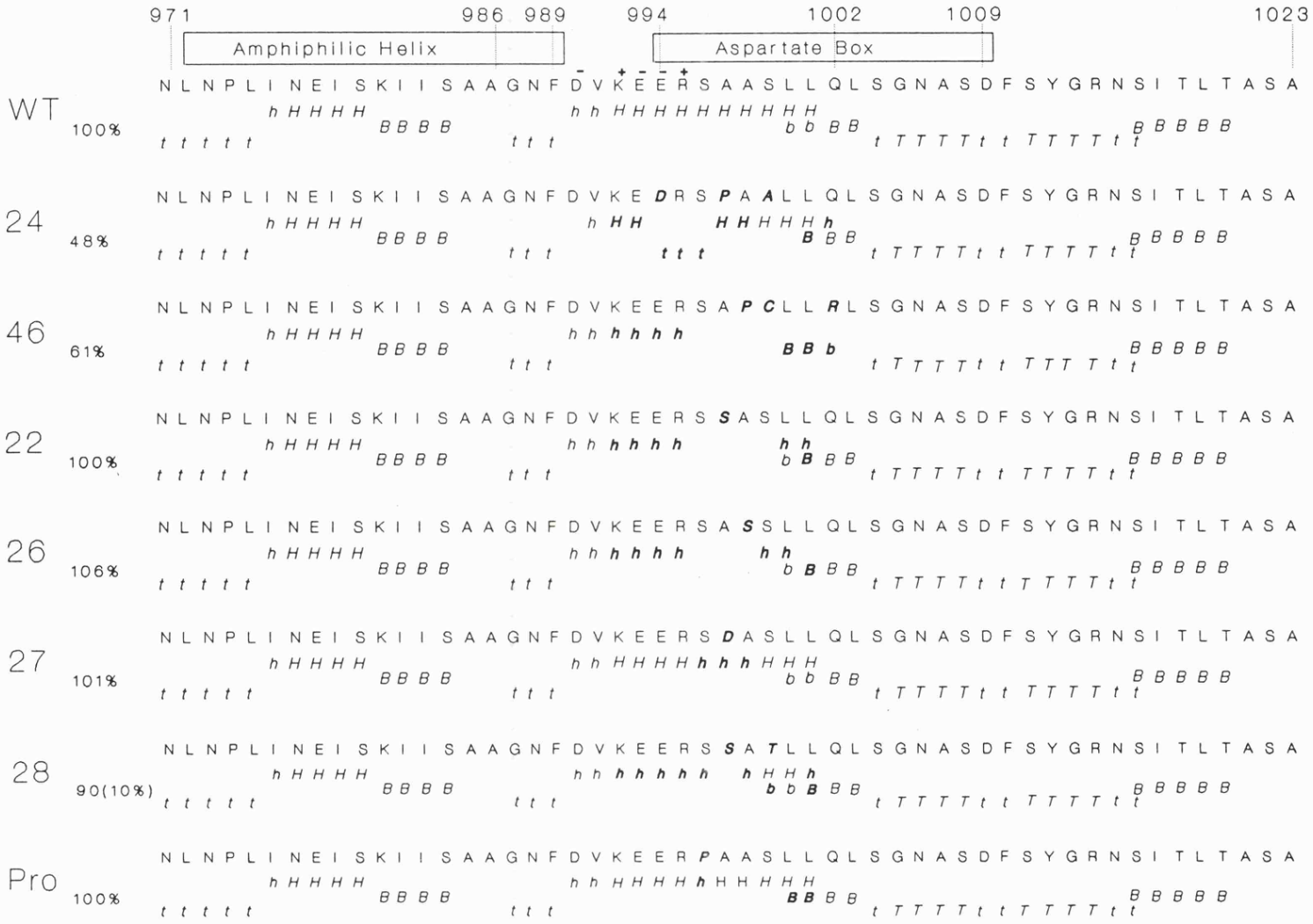

971
986
989
994
1002
1009
1023
Amphiphilic Helix
Aspartate Box
WT
100%
24
48%
46
61%
22
100%
26
106%
27
101%
28
90(10%)
Pro
100%

Figure 7:

Theoretical substitutions possible, by single point mutations, over the probed region (Phe^{989} to Gln^{1002}).

The residues probed in this study are given in the one letter code from Phe^{989} (F) to Gln^{1002} (Q). The theoretically possible substitutions for each codon are given with the number of times a particular substitution is predicted to occur (if greater that 1) given on the left hand side with the actual number of each substitution obtained so far shown on the left hand side of the columns.

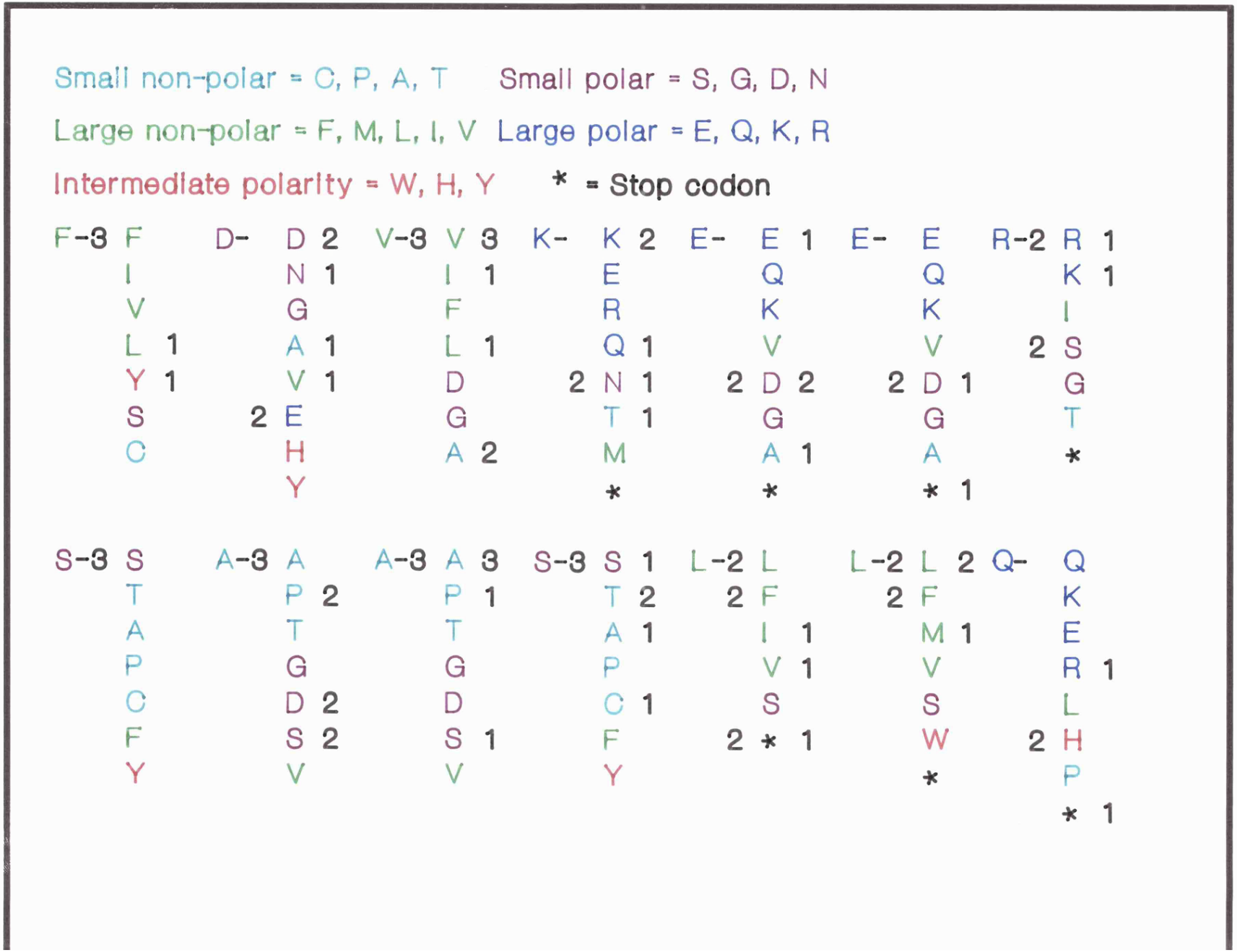
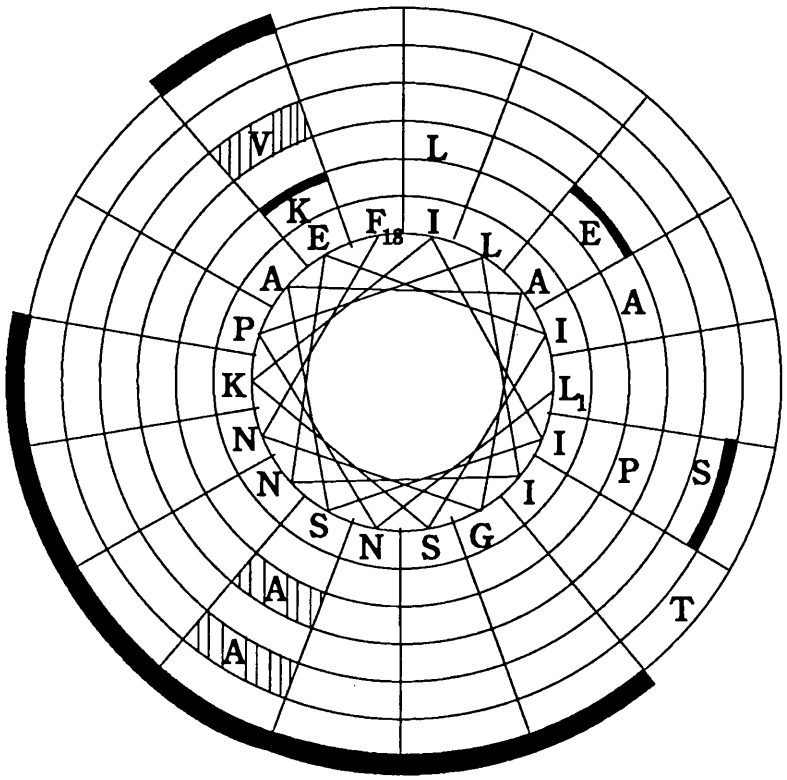

Small non-polar = C, P, A, T
Small polar = S, G, D, N
Large non-polar = F, M, L, I, V
Large polar = E, Q, K, R
Intermediate polarity = W, H, Y
* = Stop codon
F-3 F I V L 1 Y 1 S C
D- D 2 N 1 G A 1 V 1 2 E H Y
V-3 V 3 I 1 F L 1 D G A 2
K- K 2 E R Q 1 2 N 1 T 1 M * 2
E- E 1 Q K V 2 D 2 G A 1 * 2
E- E Q K V 2 D 1 G A * 1
R-2 R 1 K 1 I 2 S G T *
S-3 S T A P C F Y
A-3 A P 2 T G D 2 S 2 V
A-3 A 3 P 1 T G D S 1 V
S-3 S 1 T 2 A 1 P C 1 F Y 2
L-2 L 2 F I 1 V 1 S 2 * 1
L-2 L 2 F M 1 V S W * 2
Q- Q K E R 1 L H P * 1

Figure 8:

Substitution in the predicted Amphiphilic helix.

The potential for an amphiphilic helical forming region within HlyA suggested by Koronakis et al., (1989) is displayed on a Schiffer and Edmunson (1967) helical wheel. The polar face is indicated by the surrounding thick filled lines and the non-polar face is unmarked. The substitutions obtained to date in this region by Sally Taylor (Masters student, 1990; this laboratory: see Table 2), for mutants A to E are shown from the centre outwards, with the wild-type sequence being the innermost. Alterations to the amphiphatic nature of this region are highlighted as follows:- non-polar residues into polar face by hatched boxes and polar residues into the non-polar face by thick black lines.


V
L
K
E
F18
I
L
E
A
A
A
P
I
K
L1
N
I
P
S
N
I
S
N
S
G
T
A
A

CHAPTER 9

General Discussion:

One of the common features in most living organisms, whether uni or multi-cellular, is the basic composition of the unit cell. What makes a cell unique, is the possession of an enclosing membrane which secures it independence but not its dependency on the surrounding environment. Thus, the cell still needs to communicate with the surrounding environment in order to obtain information regarding the status, nutrient and other requirements to substain life. To solve this problem the cell evolved appropriate protein receptors and channels in the enclosing membrane in order to assimilate all the cellular requirements from its environment. This, basic cellular unit has been supplemented in the unicellular Gram negative bacteria by the addition of a second enclosing protective membrane, with the intervening space constituting the periplasm (see Chapter 1: Section 1.1.2). Each of these compartments contain specific sets of proteins performing quite specific functions. Although lacking such an outer membrane the eukaryotic cellular unit is more highly evolved and specialised, with the assignment of several specialised functions grouped together in separate internal compartments, such as mitochondria, chloroplasts and peroxisomes.

As the average cell synthesises 10^3 to 10^4 different polypeptides, and requires the accurate and swift delivery of each protein to its correct compartment for proper and efficient cellular functioning, this necessitated the adoption of mechanisms to target proteins efficiently to their respective compartments. Indeed, the nature of a general mechanism for this purpose has been elucidated over the last few decades, which directs proteins to the enclosing membranes of unicellular bacteria and the endoplasmic reticulum of eukaroytic cells. Those proteins to be transported out of the cytoplasmic compartment usually possess an N-terminal extension which in most cases is cleaved off during the membrane translocation process. Extensive studies in bacteria have shown that the N-terminal extension apparently only contains the information directing transfer across the cytoplasmic membrane, as the signal peptides from proteins targeted to different envelope compartments can be swapped without altering the final location of the mature protein. Surprisingly, analysis of the N-terminal signal sequences of many exported proteins have not revealed any conserved targeting features at the primary sequence level. However, isolation and analysis of signal defective mutants (see below) have led to the elucidation of a general targeting motif but how this actually interacts with the translocation machinery is still unknown.

The transient nature of these N-terminal targeting signals very conveniently demarcated the boundaries of the signal motif. Surprisingly, systematic searches for export defective mutants involving the signal peptide, for example such as that carried out with B-lactamase by Koshland and Botstein (1980), proved unsuccessful in that out of hundreds of mutants characterised only two (both nonsense mutations) prevented export. However, a powerful selection procedure for isolating export defective mutants was developed by Bassford *et al.*, (1979) which involved fusing the N-terminus of export proteins to the export incompatible cytoplasmic protein LacZ. These hybrids were generally found to jam the export channel leading to cell death at high levels of expression. Therefore, signal defective mutants were then isolated by selecting for viable cells with higher LacZ activity (or in the case of MBP-LacZ fusions, maltose resistance) indicating the now cytoplasmic location of the hybrid due to its inability to be targeted to the export channel. This procedure was used by many groups to isolate signal sequence mutations in different export proteins such as β-lactamase, MalE, PhoA (all periplasmic) and LamB (outer membrane) (Bassford and Beckwith, 1979: Bedouelle *et al.*, 1979; Emr *et al.*, 1980a; Michaelis *et al.*, 1983). Analysis of these mutants identified substitutions, insertions and deletions in the N-terminal targeting region of the hybrid protein. The single residue substitutions fell into three general classes when introduced back into an otherwise wild-type export protein a) those that cause a severe export block (>95%) and confer an essentially null phenotype b) those that are "leaky" in export and often not phenotypically distinguishable from the wild-type, blocking export and/or processing to varying degrees, while class c) reduced translation of the respective proteins indicating a possible coupling of translation with translocation. This powerful selection procedure allowed the rapid detection and isolation of the class b) and c) mutants, even though they are phenotypically indistinguishable from the wild-type protein when recombined into an otherwise wild-type export protein. Thus, it was possible to identify those features, that although not essential for export, were required for an efficient translocation process. The class a) mutants were generally found to involve mutation events confined to a specific subset of amino acids, for example introducing a charge into the hydrophobic region, but the pattern was not consistent between different signal sequences.

From these sorts of analysis those features important for targeting were elucidated. The severely defective mutants highlighted the importance of a hydrophobic region, the length of which was quite variable (10 to 15

residues) between export proteins. The isolation and analysis of extragenic suppressors to some of these first site mutations suggested that this region was also required to adopt an alpha-helical structure. However, in the case of the OmpA signal sequence the results indicated an equal probability for the formation of either an alpha-helical or β-sheet structure. Thus a number of suppressor mutants which resulted in an increased rate of translocation were correlated with an increasing probability of forming a β-sheet structure (Goldstein *et al.*, 1990). There is some evidence from physical studies on the LamB and PhoE signal peptides that they adopt β-sheet conformations in an aqueous environment or on lipid surfaces whilst α-helical structure is formed when inserted into phospholipid monolayers (Briggs *et al.*, 1986; Batenburg *et al.*, 1988b)

The study of the otherwise phenotypically wild-type export defective mutants (Class b above) have indicated the importance of the net positive charge (although not essential) at the extreme N-terminus of the signal peptide. This is required for optimal processing and translocation, consistent with the proposed Loop model (Inouye *et al.*, 1982), the charged residues catalysing interaction with the negative phospholipid bilayer (see Section 1.3.2: iv). Other mutants from this class affected the cleavage process leading to the accumulation of the precursor form (see Gennity *et al.*, 1990 for review). Analysis of selected mutants together with those generated by mutagenesis led also to a general consensus sequence for the cleavage site of AXB (von Heijne, 1983) with A generally being Ala or another aliphatic residue B usually a Gly or Ala residue, with low specificity at position X. There is some evidence that the presence of a β-turn may also be required for recognition by the signal peptidase at the cleavage site.

The question has been asked as to whether signal sequences must be N-terminally located in order to function and many studies in fact indicate when inserted into the mature protein they tend to act as stop transfer signals (Coleman *et al.*, 1985), indicating the importance of position or context. Nevertheless, some proteins do apparently employ internal signal sequences, such as ovalbumin (Tabe *et al.*, 1984) and the Yersina secreted YopS proteins (G. Cornelis, personal communication) presumably the particular context again is important in this case in avoiding any effect upon subsequent translocation.

Targeting of proteins to eukaryotic intracellular compartments such as the mitochondria, usually, employs transient N-terminal signal sequences. However, these differ in composition from the type considered above, in being generally larger (up to 70 residues long), being rich in positively charged and hydroxylated residues, and frequently lacking acidic residues. It is now widely accepted that these targeting signals lead to the interaction of these proteins with proteinaceous components in the outer membrane of the mitochondria which facilitate their translocation, across the membrane and entry into the appropriate compartment (see Chapter 1: Section 1.4.2).

Another intriguing question is how do Gram negative bacteria distinguish between those proteins destined to be exported to the cellular envelope compartments from those to be secreted from the cell ?. As stated before most proteins destined for the periplasmic or outer membrane compartments of bacteria carry classical N-terminal signal sequences which function to engage the protein with the inner membrane translocation machinery, the signal usually being cleaved during the translocation process. The final sorting of the protein in most cases is apparently dependent on the conformation and composition of the individual polypeptide, which is thought to determine the optimum location for function (see Chapter 1: Section 1.5.1). Thus, one could conceive that the cellular mechanism for secreting proteins from Gram negative organisms would use the same general translocation mechanism to gain entry to the periplasm or outer membrane compartments, with other components required for release from or to translocate the protein across the outer membrane. Indeed, this does seem to be the case for the majority of secreted proteins investigated to date, such as the secretion of pullulanase from *Klebsiella pneumoniae* (Pugsley *et al.*, 1990), cholera toxin from *Vibrio cholerae* (Hirst *et al.*, 1987a), elastase, alkaline phosphatase, lipase and phospholipase C from *Pseudomonas aeruginosa* (Lazdunski *et al.*, 1990), and pectate lyase and cellulase from *Erwinia chrysanthemi* (Andro *et al.*, 1984). The secretion of all these proteins also appears to be dependent on the presence of additional accessory proteins. Another class of Gram negative, secreted proteins exist which also utilise the classical N-terminal signal, translocation machinery, but do not appear to enter the periplasmic space en route to the external medium. One such protein is exotoxin A from *Pseudomonas aeruginosa* (Lory *et al.*, 1983) which is thought to gain direct access from the cytoplasm to the external media via zones of contact between the inner and outer membranes. Indeed, the secretion of exotoxin A has been shown to be

dependent upon the same gene products required for the secretion of other *Pseudomonas aeruginosa* proteins carrying N-terminal signals and involving detectable periplasmic intermediates (Lazdunski *et al.*, 1990). A third class of secreted proteins are secreted to the periplasm via the classical N-terminal signal sequence route but do not appear to require any accessory proteins to ensure their specific release from the periplasm. One such proteins is the IgA protease from *Neisseria gonorrhoeae* (Polhner *et al.*, 1987). Perhaps these proteins, while not requiring any unique accessory components, require the presence of ubiquitous outer membrane proteins to aid their transfer across the membrane

Recently, a fourth class of secreted proteins has emerged which appear to utilise a novel C-terminal targeting signal. Indeed, detailed studies on the original member of this family, the *E. coli* haemolysin protein HlyA, demonstrated that the protein was translocated independently of the general export pathway (Gray *et al.*, 1986; Gray *et al.*, 1989) and is also secreted directly to the external medium without a detectable periplasmic intermediate (Gray *et al.*, 1986; Felmlee and Welch, 1988; Koronakis *et al.*, 1989). Unlike the N-terminal signal sequences, whose boundaries are conveniently defined by the removal of the peptide during export this is not the case for C-terminal targeting signals. Nevertheless, C-terminal deletion studies have shown that the presence of the final 27 residues is essential for secretion (Gray *et al.*, 1986; Koronakis *et al.*, 1989) while the C-terminal 23, 15 or 12.5KD of the HlyA molecule encodes all the information required to ensure their own specific secretion (Nicaud *et al.*, 1986; Mackman *et al.*, 1987; this study) by the Hly, membrane bound, export proteins (HlyB,D). The discovery that these domains encode all the information required for both recognition and secretion was confirmed from the ability of these HlyA domains to promote the secretion of heterologous "passenger" domains when fused C-terminally (Mackman *et al.*, 1987; Chapter 3 and 4 this study). Thus, the features required for efficient targeting must reside at least within the final 120 residues (12.5KD) of HlyA. The inability to detect efficient secretion of the C-terminal 4KD (39 residues) of HlyA when expressed independently or fused to 40KD of the prochymosin molecule suggests that this region is insufficient to encode all the information/features required for recognition and efficient secretion (see Chapter 4). Nevertheless, the ability on one occasion to detect minor secretion levels of the afore-mentioned hybrid peptide might suggest that the molecule carries a weakly recognised targeting signal but that other sequences are required for either efficient

recognition or to allow the proper presentation of the targeting motif unhindered by the downstream passenger domain, which may otherwise occlude the recognition signal. If this is the case then it would be expected that the final 4KD HlyA peptide when expressed independently would be secreted in the presence of HlyB,D, albeit, possibly at low levels. This was not observed (see Chapter 4: Section 4.6). However, this could be due to a number of plausible reasons such as the possible instability of such a small peptide or the inability to detect, stain or resolve such small peptides by SDS-PAGE (see Chapter 4: Section 4.8). Indeed, it has been reported recently (C. Wandersman, personal communications) that the final 39 residues (fused to 12 residues of LacZ in the construct in question) of an Erwinia protease (also carries a C-terminal targeting signal), is secreted from *E. coli* by the protease specific export proteins. However, this ~6KD peptide was not detectable on stained gels and was only visualised on Western blots using specific antibodies directed against the C-terminus. The inability to detect the HlyA 4KD peptide even by Western analysis could be due to the very weak cross-reactivity of the HlyA antibodies to regions within even the final 12.5KD region of HlyA (unpublished observations; this laboratory). Thus, it remains a possibility that the final 39 residues of HlyA do encode sufficient information for at least weak recognition, but still lacks sequences required for efficient secretion.

More recent studies have suggested that the HlyA C-terminal targeting signal residues entirely within the final 53 residues (Gray *et al.*, 1986; Mackman *et al.*, 1987; Koronakis *et al.*, 1989) with evidence from single substituting mutants (Chapter 8 this study) indicating that residues within the last 46 residues are essential at least for efficient secretion. It remains to be determined whether these residues are involved in the recognition, translocation or stability of the mutant molecules. The isolation of mutants carrying severely defective (10 to 30% secretors), single missense substitutions along the entire length of the C-terminal 53 residue region, suggests that a number of dispersed individual residues are critical for efficient secretion. The essentially "wild-type" phenotype of mutants substituted for intervening residues suggests that the nature of these residues is not essential for the secretion process. The tolerance of the C-terminal region to missense substitutions suggests that the overall tertiary structure is flexible, with the recognition not involving a tightly folded structure but probably a loose overall structure. The possible role of predicted secondary structures in the secretion of HlyA has been probed by mutagenesis and the results argue against important roles for

the predicted strong alpha-helix, "aspartate box" or clustered charged regions in targeting or efficient secretion (see Chapter 8). Such a role for the predicted amphiphilic helix has not yet been investigated rigorously but data already obtained is incompatible with a crucial role. However, the isolation and characterisation of more substitutions in this region will be required to determine what features/residues contribute towards efficient secretion.

The approach of "random" mutagenesis to isolate secretion defective mutants although very powerful has not so far generated any of the potentially most informative mutants, null secretion mutants, as was the case in the initial screen for N-terminal targeting mutants. However, in the latter case it was possible to isolate export null mutants using powerful selection techniques, involving LacZ hybrid technology as mentioned above. Thus, it might be worthwhile to design a similar strategy (also employing LacZ hybrid technology) to rapidly screen for all classes of secretory defective mutants. In this respect it may be feasible to exploit the apparent inability of the Hly export machinery to secrete the hybrid 200KD LacZ-HlyA protein (see Chapter 3: Section 3.5.4) under conditions of high expression which leads to cellular death. It should therefore be possible to select for defective mutants that reduce/abolish recognition of the hybrid protein by the Hly translocator and at the same time allow the secretion of a co-expressed wild-type HlyA molecule, the secretion of which could be monitored on blood agar plates. This should lead to the rapid isolation of secretion signal defective mutants, the precise nature of the mutation being then studied after recombination into an otherwise wild-type HlyA molecule.

It is still worthwhile to continue to screen a larger population of transformants following saturation or random mutagenesis for rare secretory null missense mutants, as these have proved to be ideal starting points for the selection of both inter and extragenic suppressors in the study of the N-terminal secretion signal. The characterisation of such suppressor mutants should provide valuable information not only of residue interaction within the targeting domain but also interactions of HlyA with the export machinery.

Once mutants have been isolated that reduce the secretion of the HlyA toxin molecule further analysis will be needed to determine whether the defect could be due to reduced secretion by interference with either targeting, translocation or the stability of the mutant molecules. Thus,

experiments will have to be devised to classify the mutants in order to separate those features required for specific recognition/ targeting from those required for subsequent steps in the overall efficiency of secretion. Simple competitive experiments, which monitor the effect of the mutant molecule on the secretion of other co-expressed molecules carrying wild-type signals, should be able to distinguish between mutants defective in targeting from those in translocation. Thus, substitutions inhibiting recognition should not affect the secretion of a co-expressed molecule carrying a "wild-type" signal whereas those inhibiting translocation would be expected to also hinder the secretion of such "wild-type" molecules. Ideally these experiments would be most effective and more simple to carry out on secretory null mutants but it should be possible to adapt such a procedure to analyse secretory defective molecules. Substitutions affecting secretion due to increased instability could be identified by a combination of half-life studies and Western blot analyses of mutant molecules compared to the wild-type toxin. However, it is not envisaged that such mutants would seriously affect the secretion process as the HlyA molecules are intrinsically very unstable intracellularily. Moreover, secretion normally appears to be extremely rapid (Felmlee and Welch, 1988), suggesting that each molecule is sequestered and secreted immediately upon completion and synthesis of the C-terminal targeting signal. On the other hand, any extracellular instability of such mutant molecules should be readily detected by "tell-tale" breakdown products.

Recently, a whole family of related haemolysin molecules and other more distantly related secreted proteins have emerged which also possess C-terminal targeting signals. Proteins that fall into this class include the haemolysin toxins from *Escherichia coli* (Hartlein *et al.*, 1983; Mackman and Holland, 1985b; Felmlee *et al.*, 1985a), *Proteeae* and *Morganii* species (Koronakis *et al.*, 1987), leukotoxin from *Pasteurella haemolytica A1* (Lo *et al.*, 1987), adenylate cyclase from *Bordetella pertussis* (Glaser *et al.*, 1988) and proteases from *Erwinia chrysanthemi* (PrtB and C, (Delepelaire and Wandersman, 1989) *Pseudomonas* (A. Lazdunski, personal communication) and *Serratia marcesens* species and a nodulation protein from *Rhizobium leguminosarum* (Economou *et al.*, 1990). From studies on some of these systems (HlyA, LtkA, CyaA and PrtB) it appears that at least 2 membrane bound accessory proteins (Hartlein *et al.*, 1983; Mackman and Holland, 1985b; Felmlee *et al.*, 1985a; Strathdee and Lo, 1989; Glaser *et al.*, 1988; Wandersman *et al.*, 1989) are specifically required to promote secretion, apparently together with another minor outer membrane protein (Wandersman and Delepelaire, 1990). These

proteins are thought to interact together to form a specific translocation apparatus across both membranes. All the membrane bound export translocator proteins studied so far are homologous with those of the Hly export machinery which indeed have been found to cross-complement and secrete from *E. coli* some of the other C-terminally targeted proteins (LtkA: Strathdee and Lo, 1989; CyaA: Masure *et al.*, 1990; PrtB: C. Wandersman, personal communication; and more recently a *Pseudomonas protease*: A. Lazdunski, personal communication) to small but significant levels. Although it would be envisaged that the comparison of the C-terminal region from all these molecules would reveal some "conserved" targeting motif no such features have so far been identified either at the primary or secondary level (see Chapter 6: Section 6.2.1). This supports the idea, from the analysis of C-terminal mutants (see Chapter 8) that the targeting signal might indeed be composed of a subset of critical residues required for interaction with the export machinery, with the overall structure of the C-terminus being relatively pliable as long as some of these critical residues remain available for recognition, as suggested by the inability to isolate secretory null mutants.

One of the Hly export proteins required for the secretion of haemolysin from *Escherichia coli* besides having homologues in the other Gram-negative secretion systems, also appears to be a member of a large family of ATP-dependent translocators involved in the export of various molecules from chloride ions and drugs in Man, pigments in Drosophila to peptides and cyclic polysaccharides in prokaryotic species (for review see Blight and Holland, 1990). This, may indicate the evolution of this system for the translocation of a variety of molecules from different organisms, but it is difficult to imagine how one class of protein can recognise and transport peptides, ions and complex carbohydrates. In the case of large polypeptides it is possible therefore that one role for the HlyD homologue, unique so far to these systems, is to confer specificity for such transport upon the HlyB translocator. Other as yet unidentified protein factors may similarily function in the other HlyB-like systems to specify recognition and transport with different molecules. If this is so then it also has implications for the nature of the translocation process itself since it is difficult to envisage how such a wide variety of different molecules can be translocated across membranes by a similar mechanism. One such common mechanism might involve initial binding of the export molecule to the translocator, followed by its "shuttling" across the membrane via a "revolving door" type mechanism presumably through a hydrophilic environment in a single step (Holland *et al.*, 1989).

A final question for consideration is why has a protein targeting system evolved which places the signal at the C-terminus as opposed to the almost ubiquitous N-terminal position in other targeting systems ?. The latter situation presumably has considerable advantages as it permits early identification of the nascent export protein allowing the interaction of accessory proteins (see Chapter 1: Section 1.4.5), to maintain the protein in an loose conformation compatible with export ?. Thus, in the case of C-terminal targeting signals it is likely that the protein destined for secretion will have folded to some degree, by the time the signal is synthesised and becomes available for recognition. Perhaps, as suggested above the Hly export system can translocate the HlyA molecule when in a partially or fully folded state, unlike the SecA,Y pathway (for review see Randall *et al.*, 1987; Meyer, 1988; Eilers and Schatz, 1988). This hypothesis is supported by the ability to secrete (see Chapter 3) the export incompatible cytoplasmic proteins Chloramphenicol acetyl transferase and LacZ molecules (Gentz *et al.*, 1988; Lee *et al.*, 1989) when fused to the HlyA signal domain as shown in this study. Both of these molecules would be expected to have assumed a considerable degree of "folding" by the time the HlyA targeting signal is synthesised. However, these findings still do not indicate why such a translocation system has evolved. Perhaps, it is necessary for the full length of the HlyA molecule to adopt a folded conformation appropriate for recognition and activation by cytoplasmic HlyC protein prior to translocation and secretion.

In conclusion, although these studies have generated an insight into the nature and location of this novel targeting signal, as well as determining some of the limitation for its use in the secretion of heterologous proteins, we have not yet being able to define the recognition motif. However, continued isolation, classification and analysis of *hlyA* signal domain mutants should, with time, highlight all those residues essential for the recognition and or translocation processes. This should hopefully allow us to deduce the nature of the recognition motif for this family of proteins bearing C-terminal targeting signals. The isolation of secretion defective HlyA mutants should also give us an opportunity, by selecting for extragenic suppressor mutants, to investigate its interaction with the Hly export protein(s) and or TolC and generate an insight into the elusive mechanism of translocation.

APPENDIX

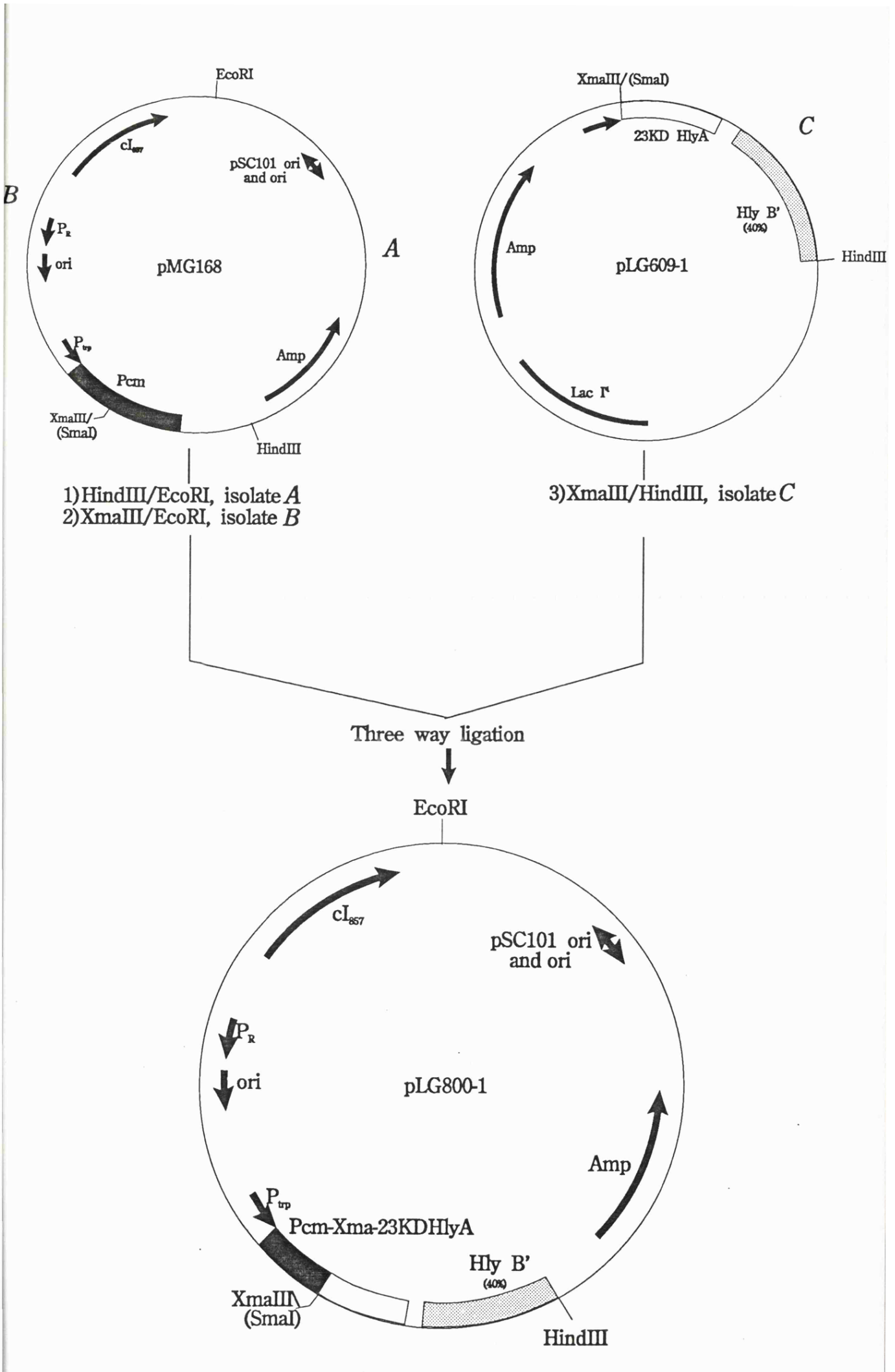
Appendix 1:

Construction of pLG800-1 (*pcm*-Xma-23KD*hlyA*)

(see text for details of construction)

pMG168: see Figure 1 for details, Wright *et al.*, 1986

pLG609-1: see Figure 2a/b for details


EcoRI
cI857
pSC101 ori
and ori
B
P_R
ori
pMG168
A
P_trp
Pcm
Amp
XmaIII/
(SmaI)
HindIII
1) HindIII/EcoRI, isolate A
2) XmaIII/EcoRI, isolate B
XmaIII/(SmaI)
23KD HlyA
C
Hly B'
(40%)
Amp
pLG609-1
HindIII
Lac I'
3) XmaIII/HindIII, isolate C
Three way ligation
EcoRI
cI857
pSC101 ori
and ori
P_R
ori
pLG800-1
Amp
P_trp
Pcm-Xma-23KDHlyA
Hly B'
(40%)
XmaIII
(SmaI)
HindIII

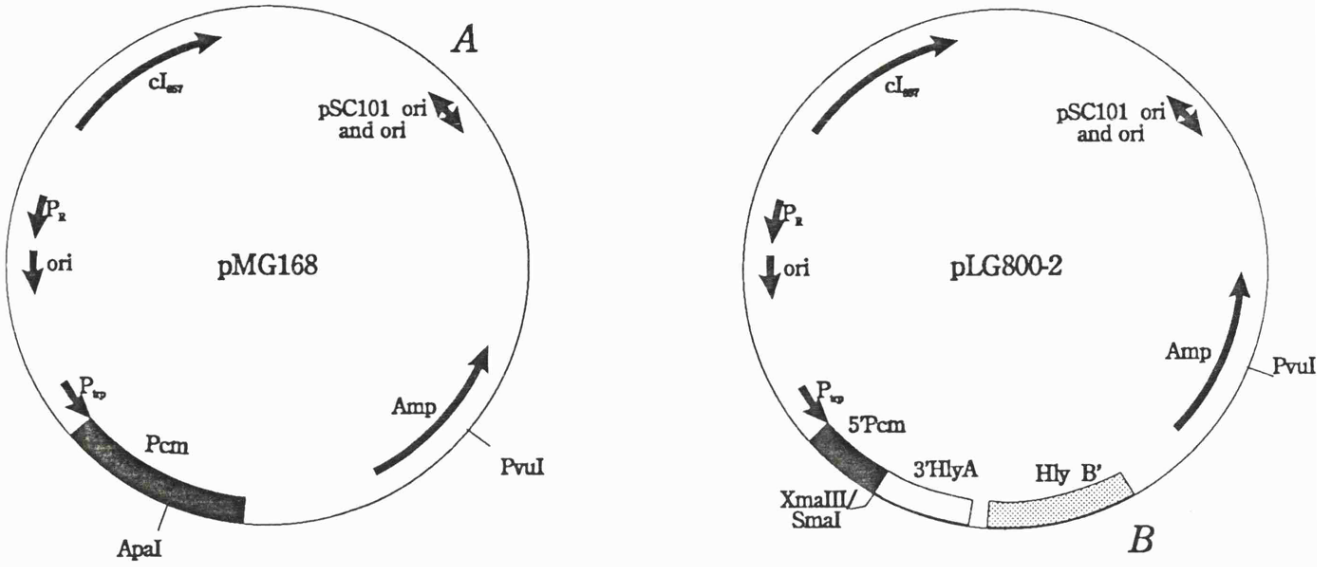
Appendix 2:

Construction of pLG801 (*pcm*-Apa-23KD*hlyA*)

(see text for details of construction)

pMG168: see Figure 1 for details, Wright *et al.*, 1986

pLG800-2: see Section 3.2 and Appendix 1 for details



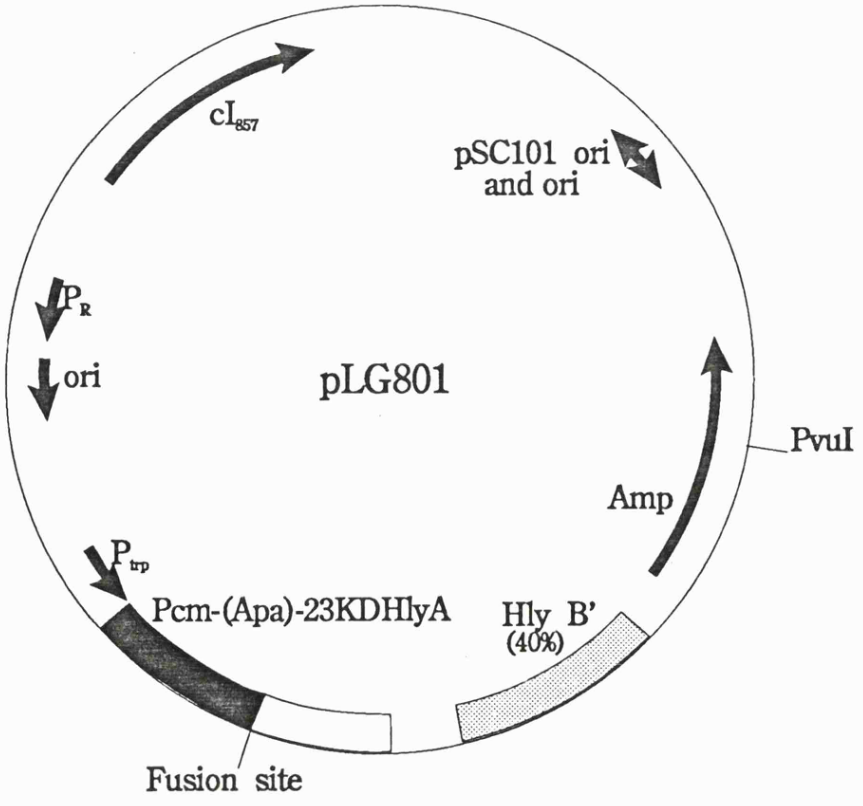
1) Cut ApaI

2) Blunt with T4 DNA polymerase

3) Cut PvuI, isolate *A*

4) Cut SmaI/PvuI, isolate *B*

Ligation



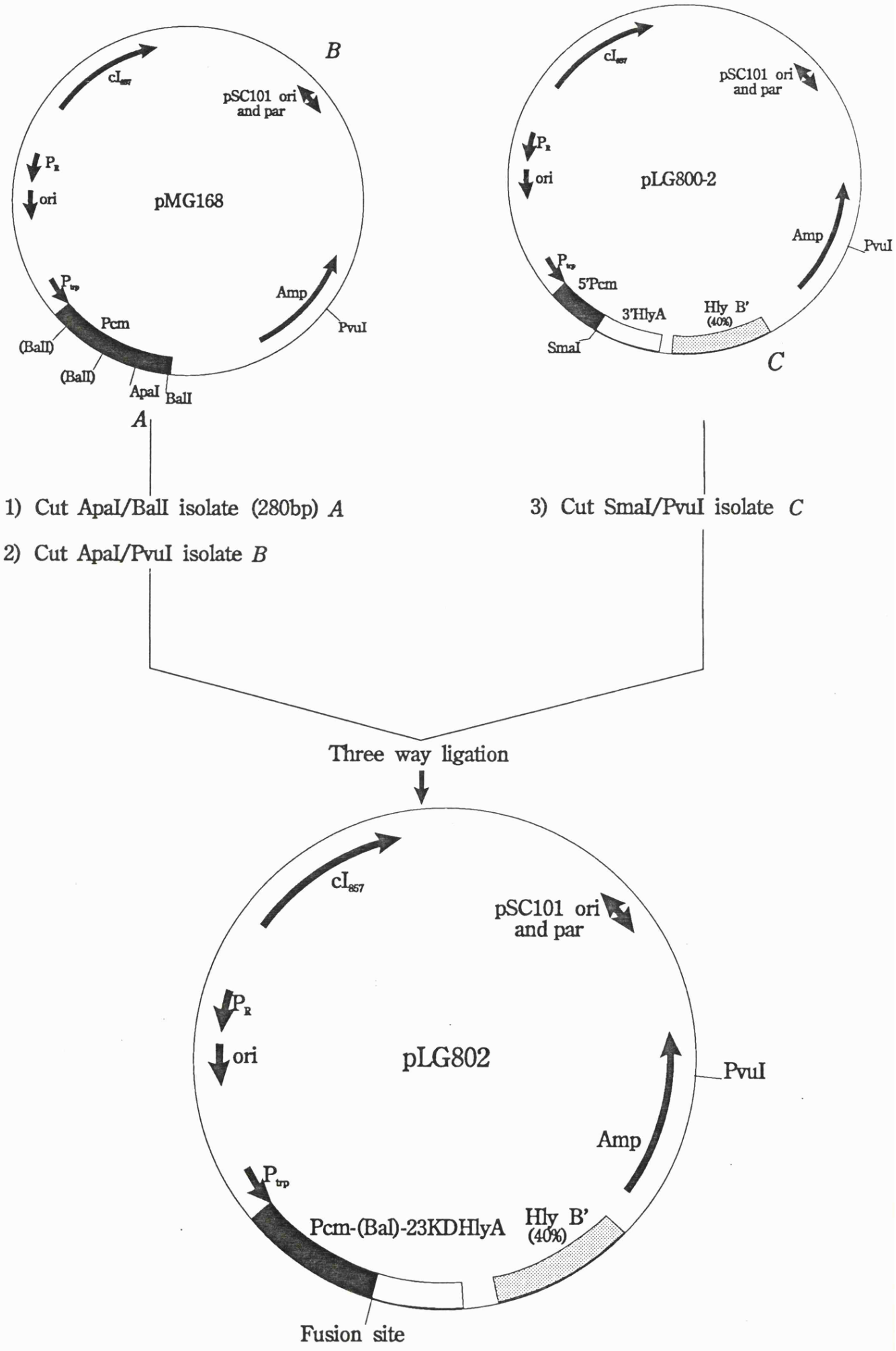
Appendix 3:

Construction of pLG802 (*pcm*-Bal-23KD*hlyA*)

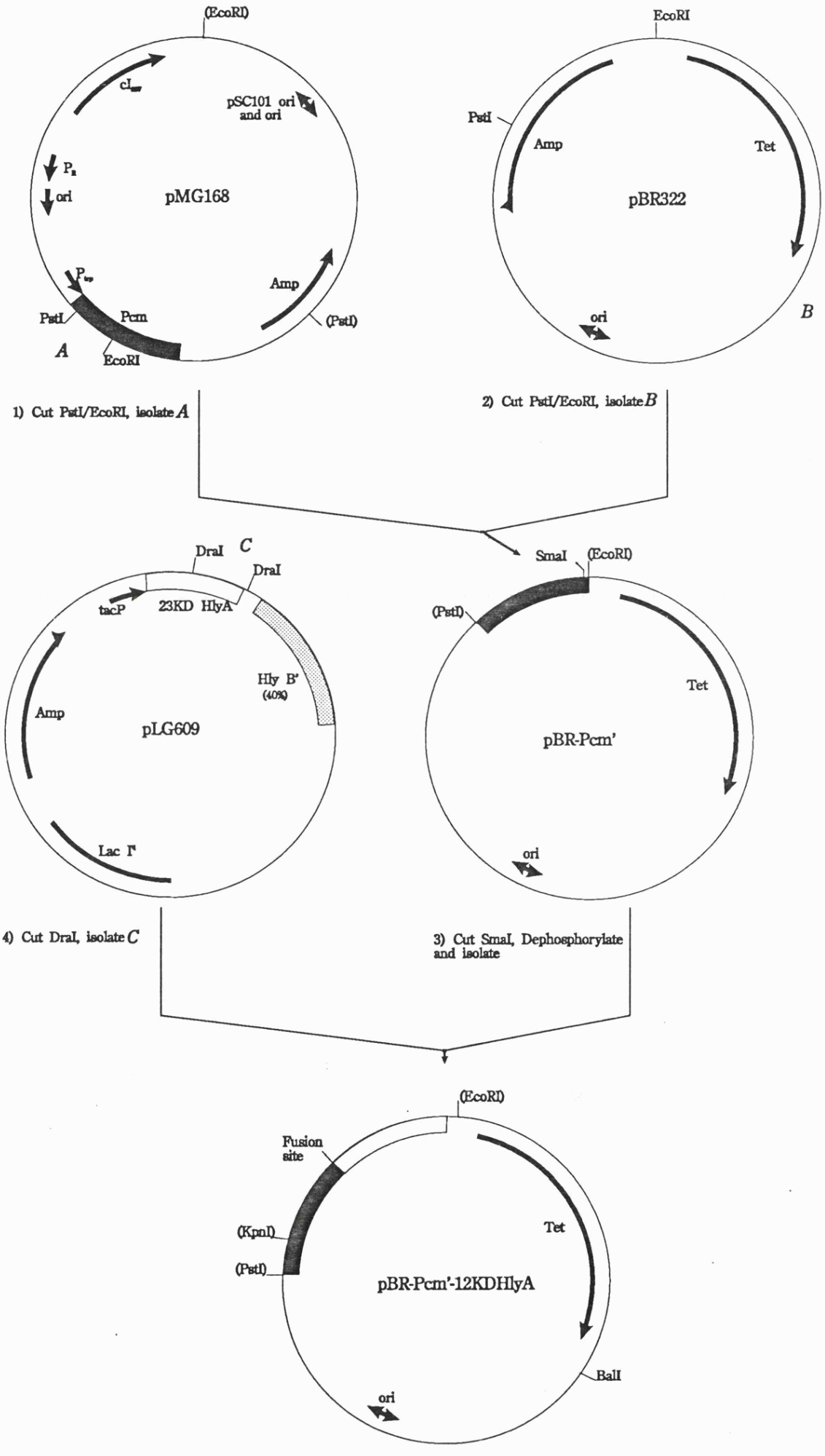
(see text for details of construction)

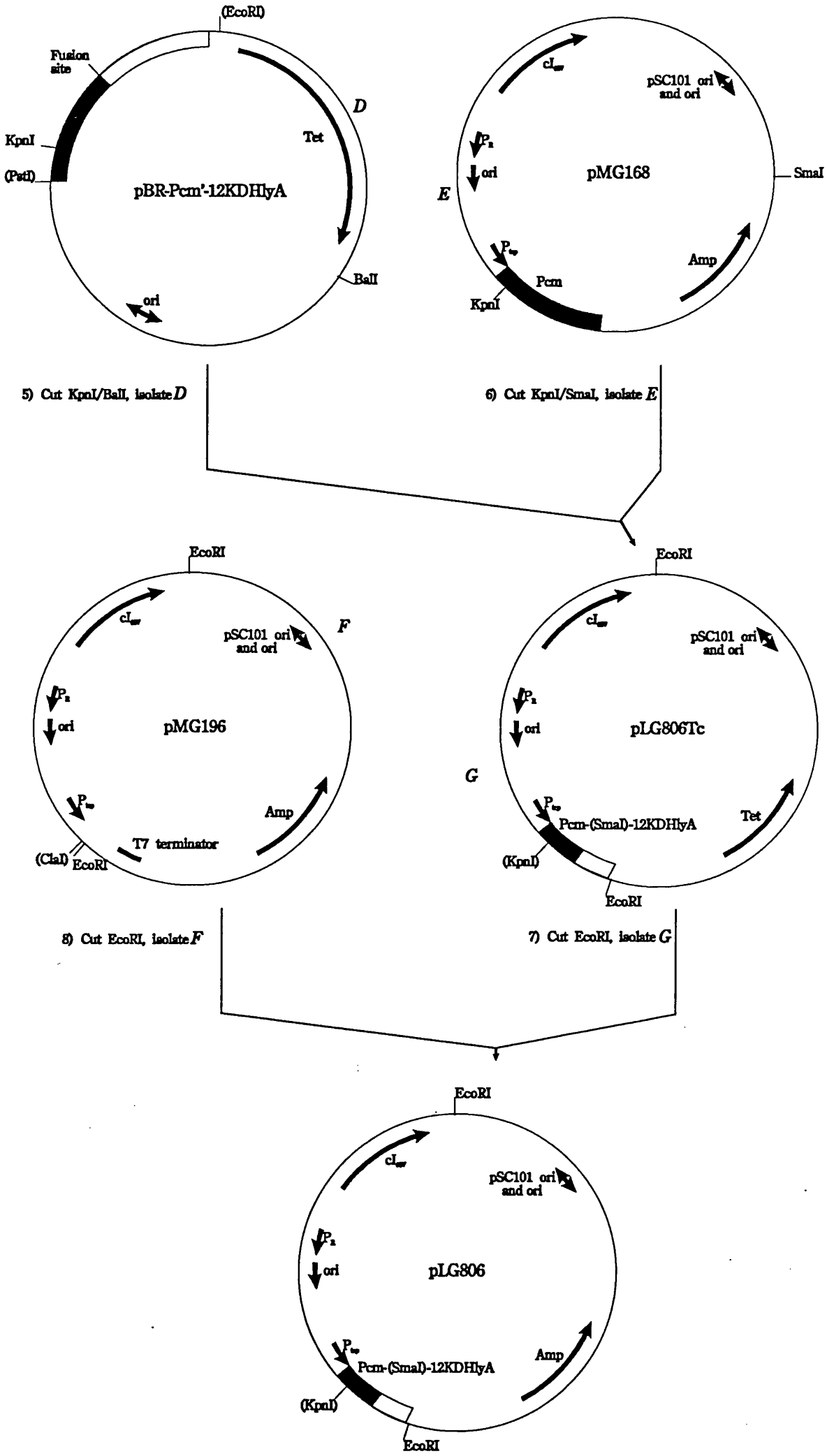
pMG168: see Figure 1 for details, Wright *et al.*, 1986

pLG800-2: see Section 3.2 and Appendix 1 for details


B
cI857
pSC101 ori and par
PR
ori
pMG168
Ptrp
Pcm
(BalI)
(BalI)
ApaI
BalI
Amp
PvuI
A
pLG800-2
5'Pcm
3'HlyA
Hly B'
(40%)
SmaI
C
1) Cut ApaI/BalI isolate (280bp) A
2) Cut ApaI/PvuI isolate B
3) Cut SmaI/PvuI isolate C
Three way ligation
pLG802
Pcm-(Bal)-23KDHlyA
Fusion site

Appendix 4a/b:
Construction of pLG806 (*pcm*-Sma-12KD*hlyA*)
pBR322: Bolivar *et al.*, 1978
pMG168: see Chapter 3; Figure 1 for details, Wright *et al.*, 1986
pLG609: see Chapter 3; Figure 2a for details
pLG800-2: see Section 3.2 and Appendix 1 for details
pMG196: Wright *et al.*, 1986
pLG806Tc: see text (Chapter 4; Section 4.2.2)


(EcoRI)
cI
pSC101 ori and ori
P
ori
pMG168
P
Amp
PstI
Pcm
(PstI)
A
EcoRI
EcoRI
PstI
Amp
Tet
pBR322
ori
B
1) Cut PstI/EcoRI, isolate A
2) Cut PstI/EcoRI, isolate B
DraI
C
DraI
tacP
23KD HlyA
Hly B'
(40%)
Amp
pLG609
Lac I'
SmaI
(EcoRI)
(PstI)
Tet
pBR-Pcm'
ori
4) Cut DraI, isolate C
3) Cut SmaI, Dephosphorylate and isolate
(EcoRI)
Fusion site
(KpnI)
(PstI)
Tet
pBR-Pcm'-12KDHlyA
BalI
ori


(EcoRI)
Fusion site
KpnI
(PstI)
pBR-Pcm'-12KDHlyA
Tet
BalI
ori
D
5) Cut KpnI/BalI, isolate D
cI
pSC101 ori and ori
pMG168
SmaI
Amp
Pcm
KpnI
E
6) Cut KpnI/SmaI, isolate E
EcoRI
pMG196
T7 terminator
Amp
(ClaI)
EcoRI
F
8) Cut EcoRI, isolate F
pLG806Tc
Pcm-(SmaI)-12KDHlyA
Tet
(KpnI)
EcoRI
G
7) Cut EcoRI, isolate G
pLG806
Pcm-(SmaI)-12KDHlyA
Amp
(KpnI)
EcoRI

Appendix 5a/b:

Construction of pLG805 (*pcm*-Sma-114bp*hlyA*)

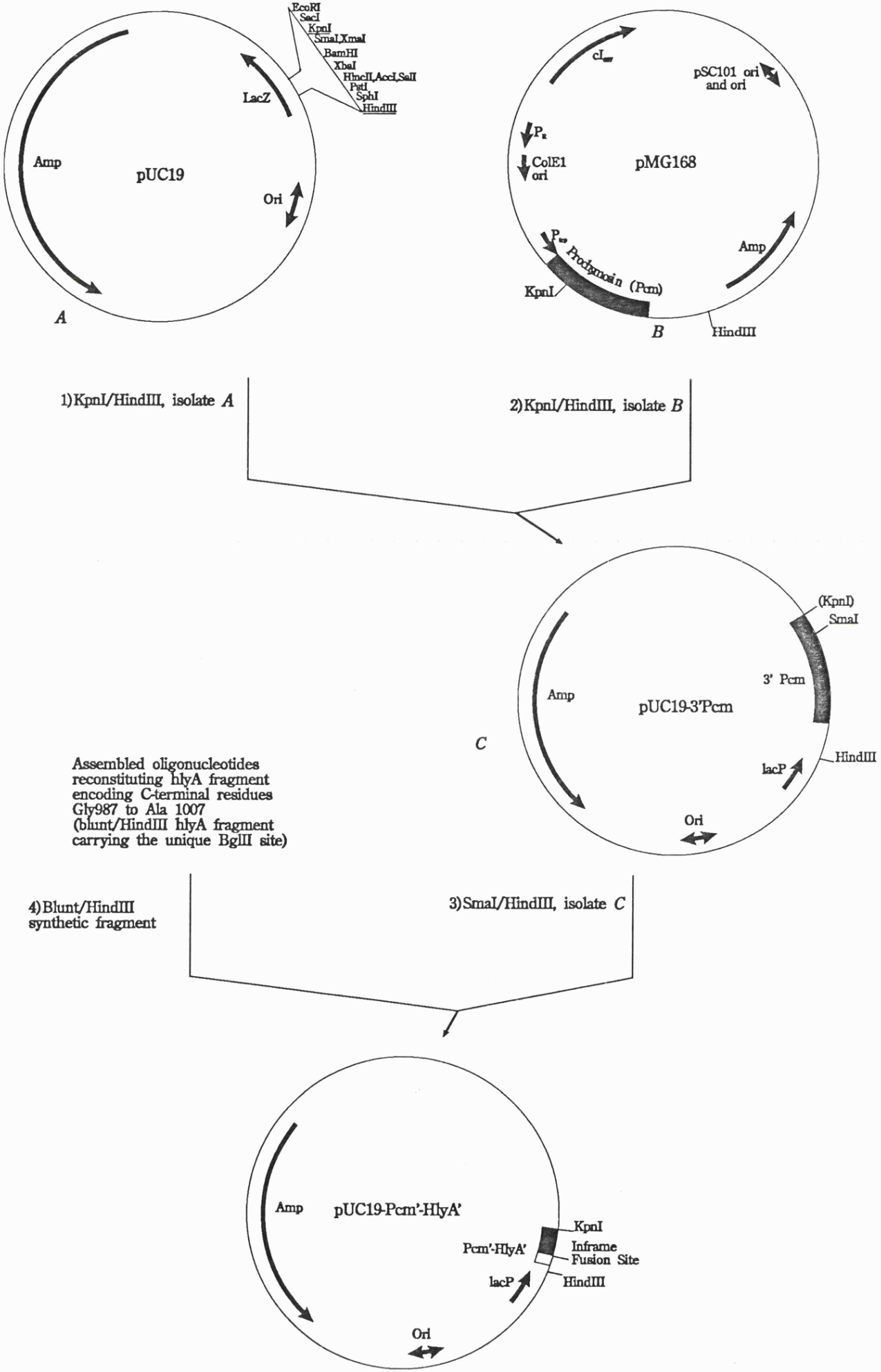
pUC19: Norrander et al., 1983

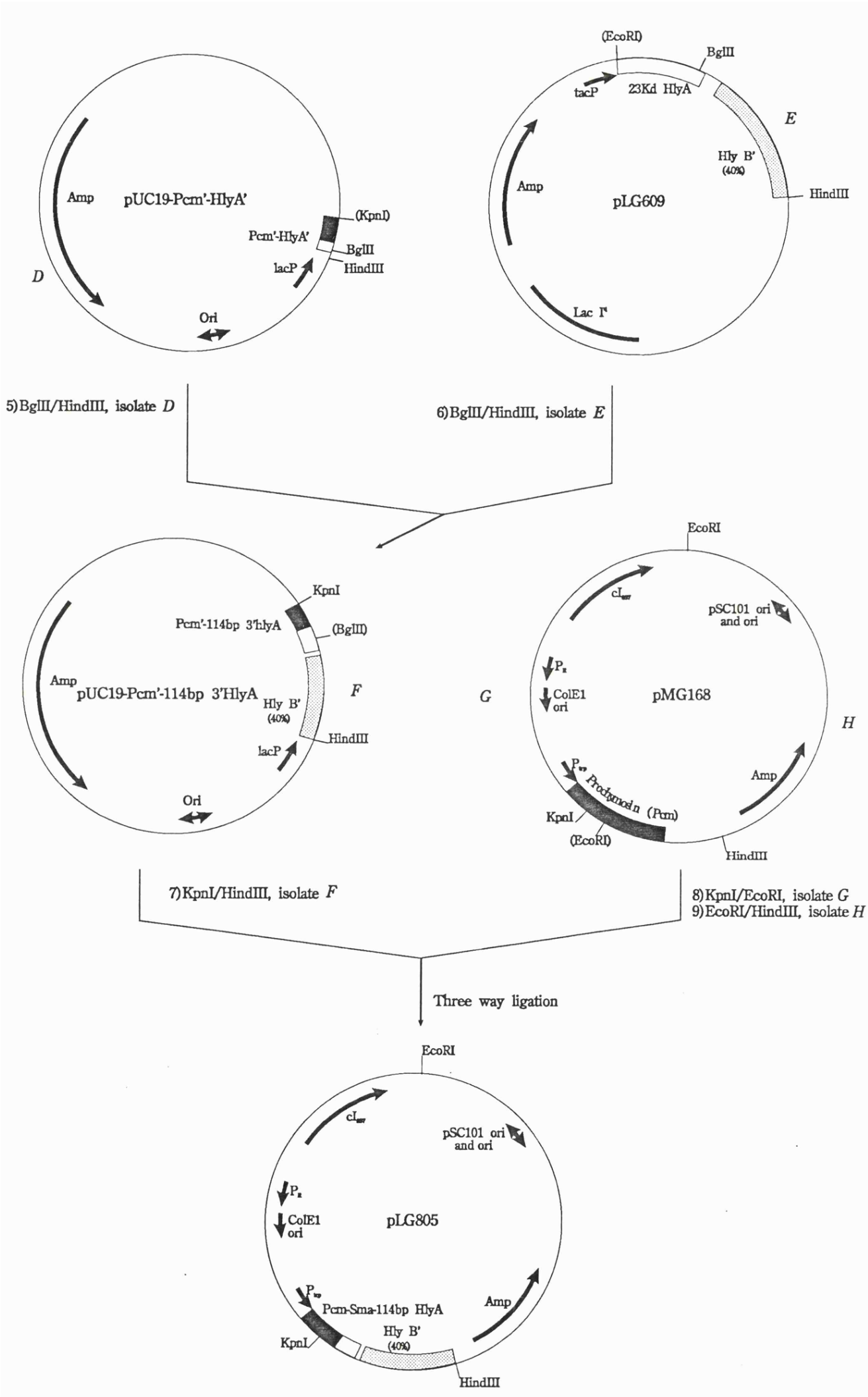
pMG168: see Chapter 3; Figure 1 for details, Wright *et al.*, 1986

Oligonucleotides reconstructing the 3' portion of *hlyA* LE2001 were kindly provided by Celltech Ltd.

pLG609: see Chapter 3; Figure 2a for details

Note: The correct assembly of the oligonucleotides used to rreconstruct the 3' of *hlyA* in this fusion was not verified by sequence analysis.


EcoRI
SacI
KpnI
SmaI,XmaI
BamHI
XbaI
HincII,AccI,SalI
PstI
SphI
HindIII
LacZ
Amp
pUC19
Ori
A
cI
pSC101 ori
and ori
P
ColE1
ori
pMG168
Amp
Prodigiosin (Pcm)
KpnI
B
HindIII
1) KpnI/HindIII, isolate A
2) KpnI/HindIII, isolate B
(KpnI)
SmaI
3' Pcm
Amp
pUC19-3'Pcm
C
HindIII
lacP
Ori
Assembled oligonucleotides
reconstituting hlyA fragment
encoding C-terminal residues
Gly987 to Ala 1007
(blunt/HindIII hlyA fragment
carrying the unique BglII site)
4) Blunt/HindIII
synthetic fragment
3) SmaI/HindIII, isolate C
Amp
pUC19-Pcm'-HlyA'
KpnI
Pcm'-HlyA'
Inframe
Fusion Site
lacP
HindIII
Ori

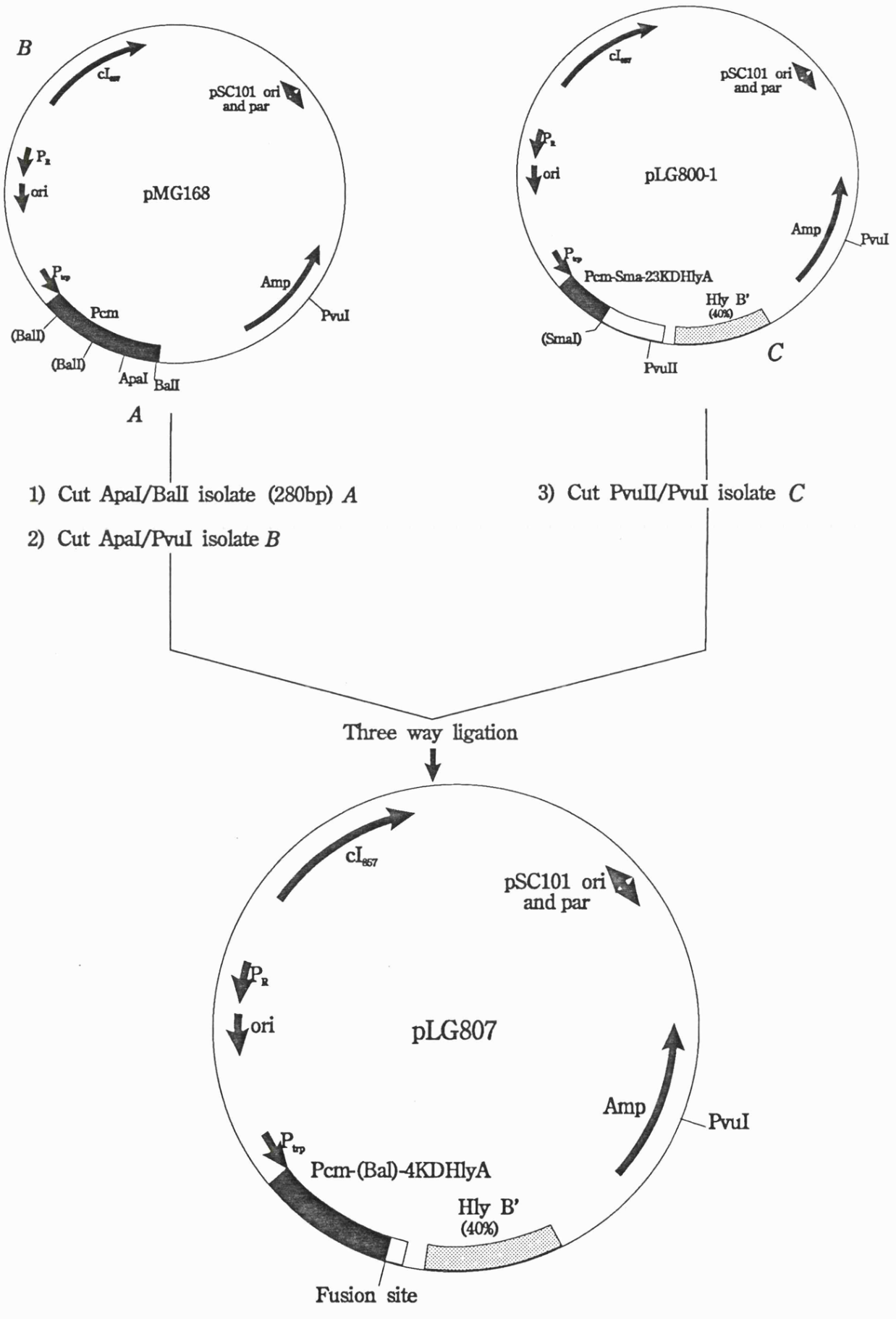

Amp
pUC19-Pcm'-HlyA'
(KpnI)
Pcm'-HlyA'
BglII
HindIII
lacP
Ori
D
(EcoRI)
BglII
tacP
23Kd HlyA
E
Hly B'
(40%)
HindIII
Amp
pLG609
Lac I'
5) BglII/HindIII, isolate D
6) BglII/HindIII, isolate E
KpnI
Pcm'-114bp 3'hlyA
(BglII)
Amp
pUC19-Pcm'-114bp 3'HlyA
F
Hly B'
(40%)
HindIII
lacP
Ori
G
EcoRI
pSC101 ori
and ori
ColE1
ori
pMG168
H
Amp
Prochymosin (Pcm)
KpnI
(EcoRI)
HindIII
7) KpnI/HindIII, isolate F
8) KpnI/EcoRI, isolate G
9) EcoRI/HindIII, isolate H
Three way ligation
EcoRI
pSC101 ori
and ori
ColE1
ori
pLG805
Pcm-Sma-114bp HlyA
Amp
Hly B'
(40%)
KpnI
HindIII

Appendix 6:

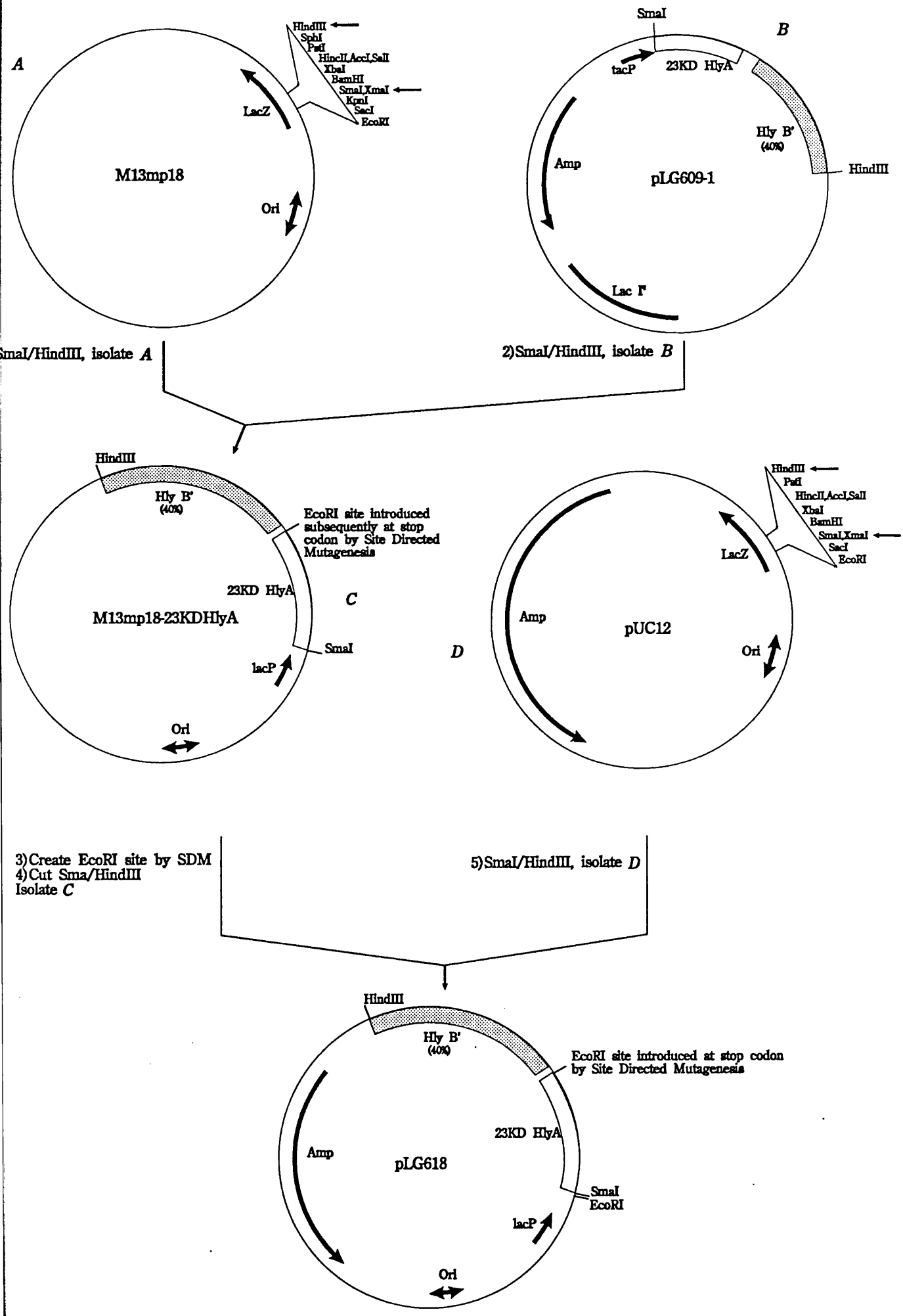
Construction of pLG807 (*pcm*-Bal-120bp*hlyA*)

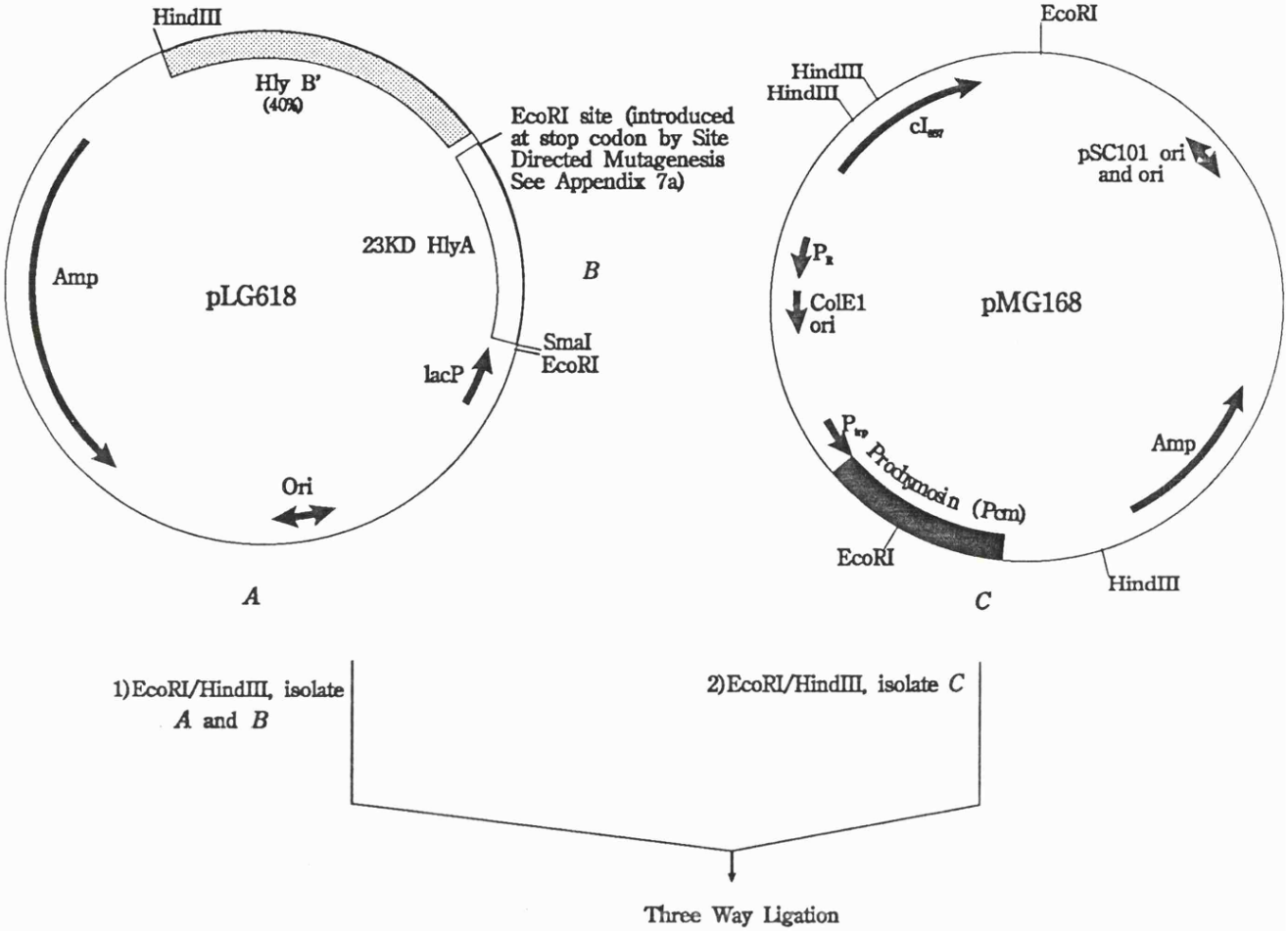
pMG168: see Chapter 3; Figure 1 for details, Wright *et al.*, 1986

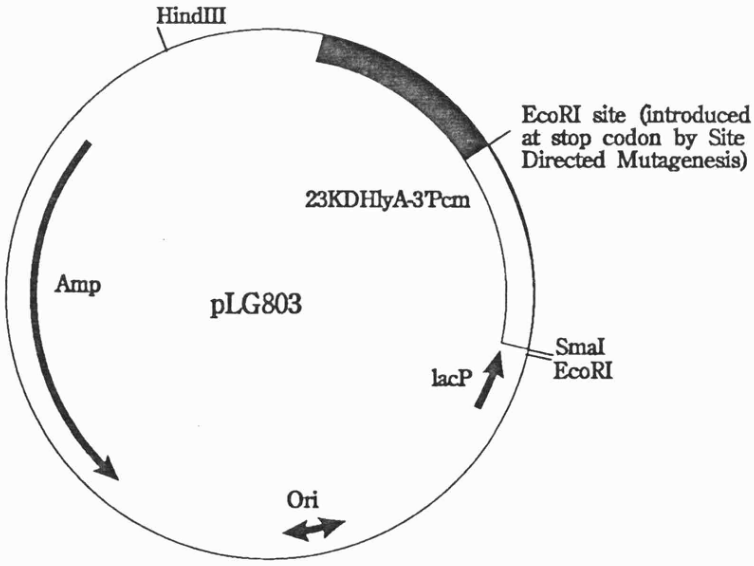
pLG800-1: see Section 3.2.1 and Appendix 1 for details


B
cI857
pSC101 ori
and par
PR
ori
pMG168
Ptrp
Pcm
(BalI)
(BalI)
ApaI
BalI
Amp
PvuI
A
cI857
pSC101 ori
and par
PR
ori
pLG800-1
Amp
PvuI
Ptrp
Pcm-Sma-23KDHlyA
Hly B'
(40%)
(SmaI)
PvuII
C
1) Cut ApaI/BalI isolate (280bp) A
2) Cut ApaI/PvuI isolate B
3) Cut PvuII/PvuI isolate C
Three way ligation
cI857
pSC101 ori
and par
PR
ori
pLG807
Amp
PvuI
Ptrp
Pcm-(Bal)-4KDHlyA
Hly B'
(40%)
Fusion site

Appendix 7a/b:
Construction of pLG618 (23KD*hlyA*EcoRI site direct mutant) and pLG803 (23DK*hlyA*-EcoRI-*pcm*)
see text for details: Chapter 4; Section 4.7.2
M13mp18: Yanish-Perron *et al.*, 1985
pLG609-1: see Chapter 3; Figure 2a/b
pUC12: Messing J, 1983
pMG168: see Chapter 3; Figure 1 for details, Wright *et al.*, 1986


A
HindIII
SphI
PstI
HincII,AccI,SalI
XbaI
BamHI
SmaI,XmaI
KpnI
SacI
EcoRI
LacZ
M13mp18
Ori
B
SmaI
tacP
23KD HlyA
Hly B'
(40%)
Amp
pLG609-1
HindIII
Lac I'
SmaI/HindIII, isolate A
2)SmaI/HindIII, isolate B
HindIII
Hly B'
(40%)
EcoRI site introduced
subsequently at stop
codon by Site Directed
Mutagenesis
23KD HlyA
C
M13mp18-23KDHlyA
SmaI
lacP
Ori
D
HindIII
PstI
HincII,AccI,SalI
XbaI
BamHI
SmaI,XmaI
SacI
EcoRI
LacZ
Amp
pUC12
Ori
3)Create EcoRI site by SDM
4)Cut Sma/HindIII
Isolate C
5)SmaI/HindIII, isolate D
HindIII
Hly B'
(40%)
EcoRI site introduced at stop codon
by Site Directed Mutagenesis
23KD HlyA
Amp
pLG618
SmaI
EcoRI
lacP
Ori


HindIII
Hly B'
(40%)
EcoRI site (introduced at stop codon by Site Directed Mutagenesis See Appendix 7a)
23KD HlyA
Amp
pLG618
B
SmaI
EcoRI
lacP
Ori
A
EcoRI
HindIII
HindIII
pSC101 ori and ori
ColE1 ori
pMG168
Amp
Prodigiosin (Pcm)
EcoRI
HindIII
C
1) EcoRI/HindIII, isolate A and B
2) EcoRI/HindIII, isolate C
Three Way Ligation


HindIII
EcoRI site (introduced at stop codon by Site Directed Mutagenesis)
23KDHlyA-3'Pcm
Amp
pLG803
SmaI
EcoRI
lacP
Ori

Appendix 8a/b:

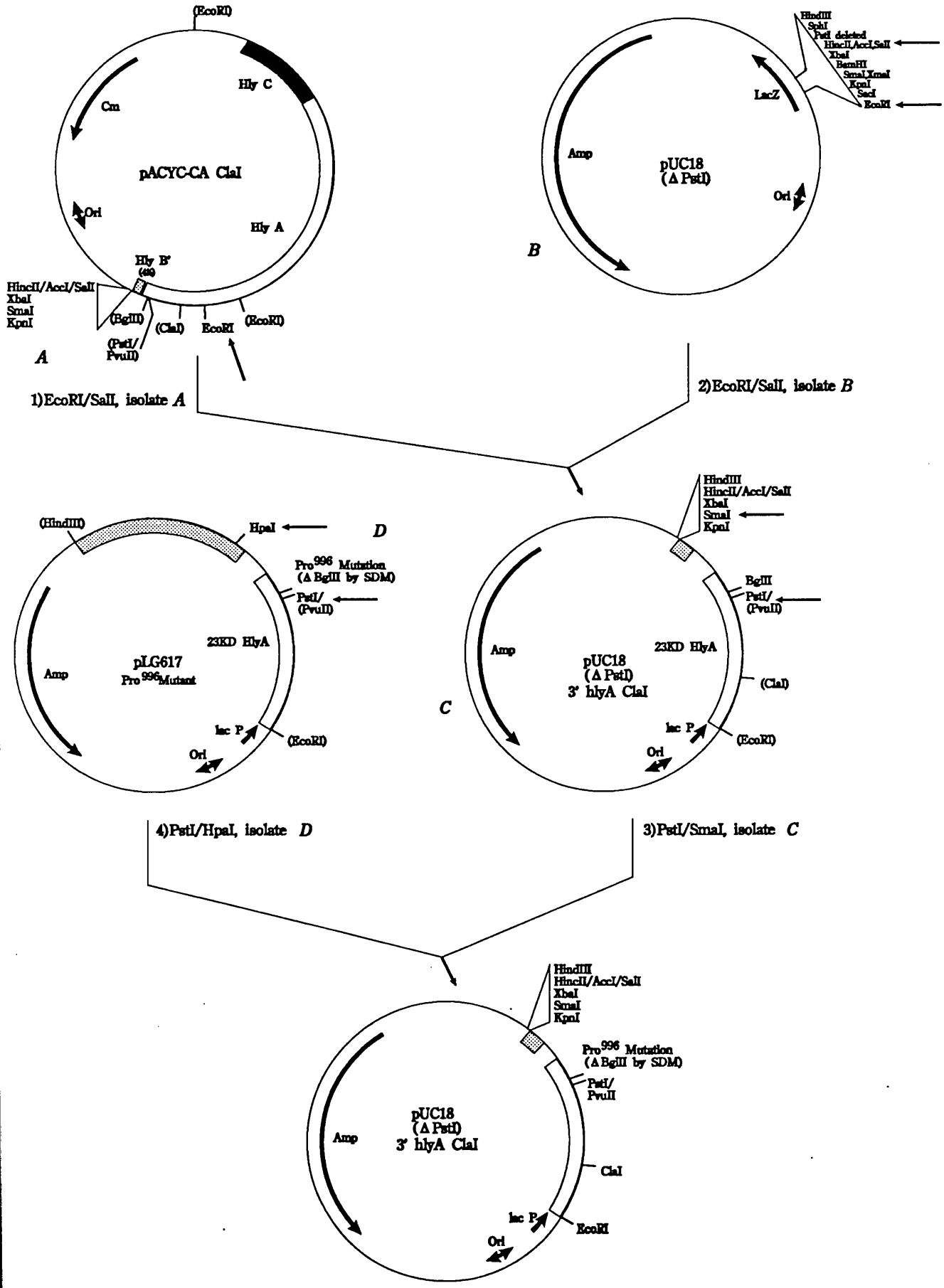
Transferring the Ser^{996} to Pro substitution from the 3' *hlyA* into the intact *hlyA* gene.

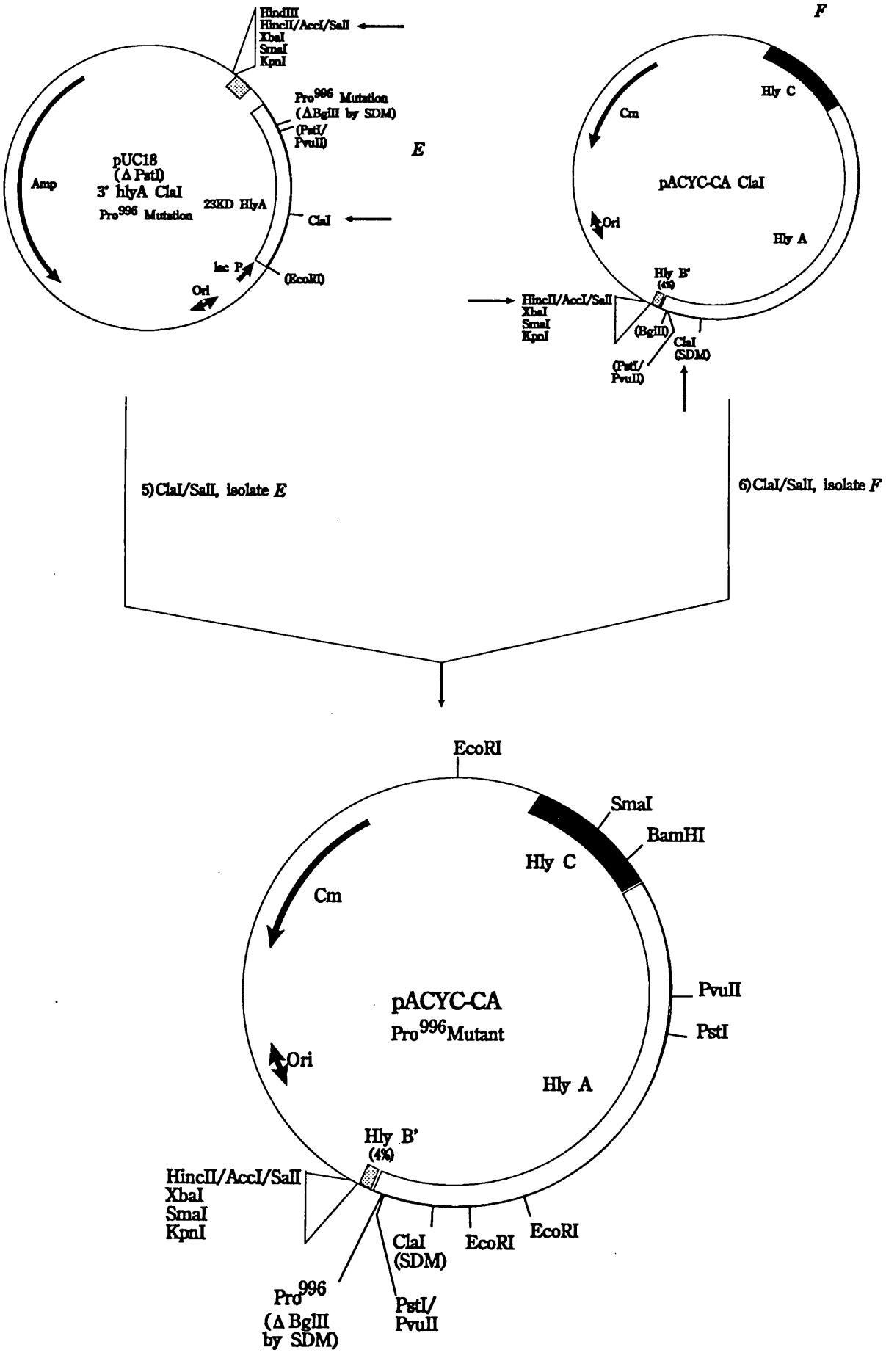
see text for details, Chapter 6; Section 6.2.1

pACYC-CA,ClaI: See Chapter 7; Section 7.5.1

pUC18 ΔPstI: pUC18 (Norrander *et al.*, 1983) with the PstI site deleted.

pLG617: pLG617 is identical to pLG618 shown in Appendix 7a except that they carry different site directed mutations in the 3' *hlyA* sequence.


(EcoRI)
Hly C
Cm
pACYC-CA ClaI
Ori
Hly A
Hly B'
(459
HincII/AccI/SalI
XbaI
SmaI
KpnI
(BglII)
(ClaI)
EcoRI
(EcoRI)
(PstI/
PvuII)
A
HindIII
SphI
PstI deleted
HincII/AccI,SalI
XbaI
BamHI
SmaI,XmaI
KpnI
SacI
EcoRI
LacZ
Amp
pUC18
(Δ PstI)
Ori
B
1)EcoRI/SalI, isolate A
2)EcoRI/SalI, isolate B
(HindIII)
HpaI
D
Pro996 Mutation
(Δ BglII by SDM)
PstI/
(PvuII)
23KD HlyA
Amp
pLG617
Pro996Mutant
lac P
(EcoRI)
Ori
HindIII
HincII/AccI/SalI
XbaI
SmaI
KpnI
BglII
PstI/
(PvuII)
Amp
23KD HlyA
pUC18
(Δ PstI)
3' hlyA ClaI
(ClaI)
C
lac P
(EcoRI)
Ori
4)PstI/HpaI, isolate D
3)PstI/SmaI, isolate C
HindIII
HincII/AccI/SalI
XbaI
SmaI
KpnI
Pro996 Mutation
(Δ BglII by SDM)
PstI/
PvuII
Amp
pUC18
(Δ PstI)
3' hlyA ClaI
ClaI
lac P
EcoRI
Ori


HindIII
HincII/AccI/SalI
XbaI
SmaI
KpnI
Pro996 Mutation
(ΔBglII by SDM)
(PstI/
PvuII)
E
pUC18
(Δ PstI)
3' hlyA ClaI
Pro996 Mutation
23KD HlyA
ClaI
Amp
lac P
Ori
(EcoRI)
F
Hly C
Cm
pACYC-CA ClaI
Ori
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
(BglII)
ClaI
(SDM)
(PstI/
PvuII)
5) ClaI/SalI, isolate E
6) ClaI/SalI, isolate F
EcoRI
SmaI
BamHI
Hly C
Cm
PvuII
PstI
pACYC-CA
Pro996Mutant
Ori
Hly A
Hly B'
(4%)
HincII/AccI/SalI
XbaI
SmaI
KpnI
ClaI
(SDM)
EcoRI
EcoRI
Pro996
(Δ BglII
by SDM)
PstI/
PvuII

REFERENCES

Literature Cited

Akiyama Y and Ito K (1985) The SecY membrane component of the bacterial protein export machinery: analysis by new electrophoresis methods for integral membrane proteins. EMBO J. 4:3351-3356

Andro T, Chambost J.P, Kotoujansky A, Cattaneo J, Bertheau Y, van Gijsengem F and Coleno A (1984) Mutations of Erwinia chrysanthemi defective in secretion of Pectinase and Cellulase. J. Bact. 160: 1199-1206

Baker K. (1987) Mechanism of protein translocation in *E. coli*. Ph.D Thesis, University of Leicester.

Baker K, Mackman N, Jackson M and Holland I.B (1987) Role of SecA and SecY in protein export as revealed by studies of TonA assembly into the outer membrane of *Escherichia coli*. J. Mol. Biol. 198: 693-703

Bankaitis V.A and Bassford Jr P.J (1985) Proper interaction between at least two components is required for efficient export of proteins to the *Escherichia coli* cell envelope. J. Bact. 161: 169-178

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

Bassford P and Beckwith J (1979) *Escherichia coli* mutants accumulate the precursors of a secreted protein in the cytoplasm. Nature 277: 538-541

Bassford Jr. P.J, Silhavy T.J, and Beckwith J.R (1979) Use of gene fusion to study secretion of Maltose-binding protein into the *Escherichia coli* periplasm. J. Bact 139: 19-32

Batenberg A.M, Brasseur R, Ruysschaert J-M, van Scharrenburg G.J.M, Slotboom A.J, Demel R.A and deKruijff B (1988b) Characterisation of the interfacial behaviour and structure of the signal sequence of *Escherichia coli* outer membrane protein PhoE. J. Biol. Chem. 263: 4204-4208

Batenberg A.M, Demel R.A, Verkleij A.K and deKruijff B (1988a) Penetration of the signal sequene of *Escherichia coli* PhoE protein into phospholipid model membrane leads to lipid specific change in signal peptide structure and alterations in lipid organisation. Biochemistry 27: 5678-5685

Baty D, Lloubes R, Geli V, Lazdunski C and Howard SP (1987) Extracellular release of colicin A is non-specific. EMBO J. 6: 2463-2468

Bayer ME (1979) The fusion sites between outer membrane and cytoplasmic membranes in bacteria: Their role in membrane assembly and virus infection. In Bacterial Outer Membranes (ed Inouye M) 167-202

Bayer M.E, Bayer M.H, Lunn C.A and Piglet V (1987) Association of thioredoxin with the inner membrane and adhesion sites in *Escherichia coli*. J. Bact. 169: 2659-2666

Beckwith J and Ferro-Novick S (1986) Genetic studies on protein export in bacteria. Curr Topics in Micro and Immun 125: 5-27

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

Bernard H-U and Helinski D.R (1979) Use of the lambda phage promoter P_L to promote gene expression of hybrid plasmid cloning vehicles. Meths. in Enzymol. 68: 482-492

Bieker K.L and Silhavy T.J (1989) PrlA is important for the translocation of exported proteins across the cytoplasmic membrane of *Escherichia coli*. Proc. Nat. Acad. Sci. 86:968-972

Bieker K.L and Silhavy T.J (1990) PrlA (SecY) and PrlG (SecE) interact directly and function sequentially during protein translocation in *E. coli*. Cell 61: 833-842

Birnboim H.C and Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nuc. Acids. Res. 7: 1513-1523

Bhakdi S, Mackman N, Nicaud J-M and Holland I.B (1986) *Escherichia coli* hemolysin may damage target cell membranes by generating transmembrane pores. Infect. and Immun. 51: 63-69

Blight M (1990) Analysis of the membrane localised transport protein haemolysin B. Ph.D Thesis, Dept. of Genetics, Leicester University, U.K.

Blight M and Holland I.B (1990) Structure and function of haemolysin B, P-glycoprotein and other members of a novel family of membrane translocators. Mol. Micro. 4: 873-880

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

Blobel G and Sabatini D.D (1971) Ribosome-membrane interaction in Eukaryotic cells. In "Biomembranes" ed Masion LA p193-195

Bochkaureva E.S, Lissen N.M and Girshovich A.S (1988) Transient assocation of newly synthesised unfolded proteins with the heat-shock GroEL protein. Nature 336: 254-257

Bolivar F, Rodriguez R.L, Greene P.J, Betlach M.C, Heyneker H.L and Boyer H.W (1977) Construction and characterisation of new cloning vehicles. II A multipurpose cloning system. Gene 2: 95-113

Bosch D, Leunissen J, Verbakel J, van Erp M and Tommassen J (1986) Periplasmic accumulation of truncate forms of outer-membrane PhoE protein of *Escherichia coli* K-12. J. Mol. Biol. 189: 449-455

Bosch D, Voorhout W and Tommassen J (1988) Export and localisation of N-terminally truncated derivatives of *Escherichia coli* K-12 outer membrane protein PhoE. J. Biol. Chem. 263: 9952-9957

Bourdineaud J-P, Howard S.P and Lazdunski C (1989) Localisation and assembly into the *Escherichia coli* envelope of a protein required for entry of colicin A. J. Bact. 171: 2458-2465

Boyd A and Holland I.B (1979) Regulation of the synthesis of surface proteins in the cell cycle of *Escherichia coli* B/r. Cell 18: 287-296

Braun V (1975) Covalent lipoprotein from the outer membrane of *Escherichia coli*. Biochim et Biophys Acta 415: 335-377

Briggs M.S, Cornell D.C, Dluthy R.A, and Gieresch C.M (1986) Conformation of signal peptide induced by lipids suggests initial step in protein export. Science 233: 206-208

Bremer H and Dennis P.P (1987) Modulation of chemical compositions and other parameters in the cell by growth rate. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular biology Part 1 (ed Neidhardt FC) p1527-1543

Bretscher M.S (1972) Asymmetrical lipid bilayer structure for biological membranes. Nat New Biol 236: 11-12

Bretscher M.S (1973) Membrane structures: Some general principles. Science 181: 622-629

Burnell E, van Alpen L, Verkley A and deKruijff B (1980) ^{31}P nuclear magnetic resonance and freeze-fracture electron microscopy studies on *Escherichia coli*. 1 Cytoplasmic membrane and total phospholipid. Biochim et Biophys Acta 597: 492-501

Cabelli R.J, Chen L, Tai P.C and Oliver D.B (1988) SecA Protein is required for secretory protein translocation in *E. coli* membrane vesicles. Cell 55: 683-692

Cavalieri S.J, Bohach G.A and Snyder I.S (1984) *Escherichia coli* alpha-haemolysin: characteristics and probable role in pathogenenicity. Micribiol. Rev. 48: 326-343

Chang A.C.Y and Cohen S.N (1978) Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bact. 134: 1141-1156

Chen L and Tai P.C (1985) ATP is essential for protein translocation into *Escherichia coli* membrane vesicles. Proc. Nat. Acad. Sci. 82: 4384-4388

Chen L, Rhoads D and Tai P (1985) Alkaline phosphatase ond OmpA protein can be translocated posttranslationally into membrane vesicles of *Escherichia coli*. J. Bact. 161: 973-980

Chirico W.J, Waters M.G and Blobel G (1988) 70K heat shock related proteins stimulate protein translocation into microsomes. Nature 332: 805-810

Chou P.Y and Fasman G.D (1978) Empirical predictions of protein conformation. Annual Reveiw of Biochem. 47: 251-276

Collier D.N, Bankaitis V.A, Weiss J.B and Bassford Jr P.J (1988) The antifolding activity of SecB promotes export of the *E. coli* Maltose-binding protein. Cell 53: 273-283

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

Creighton T.E (1984) Proteins: Structures and molecular properties. W.H Freeman and Company, New York

Cronan Jr J.E, Gennis B and Maloy S.R (1987) Molecular architecture and assembly of cell parts: Cytoplasmic membrane. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular biology Part 1 (ed Neidhardt FC) p 31-55

Crooke E and Wickner W (1987) Trigger factor: A soluble protein that folds pro-OmpA into a membrane-assembly-competent form. Proc. Nat. Acad. Sci. 84:5216-5220

Cunningham K, Lill R, Crooke E, Rice M, Moore K, Wickner W and Oliver D (1989) SecA protein, a peripheral protein of the *Escherichia coli* plasma membrane, is essential for the functional binding and translocation of proOmpA. EMBO J. 8: 955-959

de Graaf F.K and Oudega B (1986) Production and release of Cloacin DF13 and related colicins. Current Topics in Micro. and Immun. 125: 183-205

d'Enfert C and Pugsley A.P (1987) A gene fusion approach to the study of pullulanase export and secretion in *Escherichia coli*. Mol. Micro. 1: 159-168

d'Enfert C, Reyss I, Wandersman C and Pugsley A.P (1989) Characterisation of two membrane proteins required for pullulanase secretion by *Escherichia coli* K12. J. Biol. Chem. 264: 17462-17468

d'Enfert C, Ryter A and Pugsley A.P (1987) Cloning and expression in *Escherichia coli* of the *Klebsiella pneumoniae* genes for production, surface localisation and secretion of the lipoprotein pullulanase. EMBO J. 6: 3531-3538

deVrije T, de Swart R.L, Dowhan W, Tommassen J and de Kruijff B (1988) Phosphatidylglycerol is involved in protein translocation across *Escherichia coli* inner membranes. Nature 334: 173-175

Dalbey and Wickner W (1985) Leader peptidase catalyses the release of export protein from the outer surface of the *Escherichia coli* plasma membrane. J. Biol. Chem. 260(29): 15925-15931

Date T and Wickner W (1981) Procoat, the precursor of M13 coat protein, inserts post-translationally into the membrane of cells infected by wild type virus. J. Virology 37: 1087-1089

Delepelaire P and Wandersman C. (1989) Protease secretion by *Erwinia chrysanthemi*. J. Biol. Chem. 264: 9083-9089

Deshaies R, Koch B.D, Werner-Washburne M, Craig E.A and Schekman R (1988) A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. Nature 332: 800-805

Douglas C.M, Guidi-Rontani C and Collier R.J (1987) Exotoxin A of *Psuedomonas aeruginosa*: Active, cloned toxin is secreted into the periplasmic space of *Escherichia coli*. J. Bact. 169: 4962-4966

Dower W.J, Miller J.F and Ragsdal C.W (1988) High efficiency transformation of *E. coli* by high voltage electroporation. Nuc. Acids. Res. 16: 6127-6145

Duffaud G.D, Lehnhardt S.K, March P.E and Inouye M (1985) Structure and function of the signal peptide. Curr Topics in Membranes and Transport 24: 65-104

Economou A, Hamilton W.D.O, Johnston A.W.B and Downie J.A (1990) The *Rhizobium* nodulation gene *nodO* encodes a Ca^{2+}-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins. EMBO J. 9: 349-354

Eichenlaub R (1979) Mutants of the mini-plasmid pML31 thermosensitive in replication. J. Bact. 138: 559-562

Eilers M and Schatz G (1986) Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria. Nature 322: 228-232

Eilers M and Schatz G (1988) Protein unfolding and the energetics of protein translocation across biological membranes. Cell 52: 481-483

Eilers M, Hwang S and Schatz G (1988) Unfolding and refolding of a purified precursor protein during import into isolated mitochondria. EMBO J. 7: 1139-1145

Emr S.D and Bassford Jr P.J (1982) Localisation and processing of outer membrane and periplasmic proteins in *Escherichia coli* strains harbouring export-specific suppressor mutations. J.Biol. Chem. 257: 5852-5860

Emr S.D, Hall M.N and Silhavy T.J (1980a) A mechanism of protein localisation: The signal hypothesis and bacteria J. Cell. Biol. 86: 701-711

Emr S.D, Hanley-Way S and Silhavy T.J (1981) Suppressor mutations that restore export of a protein with a defective signal sequence. Cell 23: 79-88

Emr S.D, Hedgpeth J, Clement J.M, Silhavy T.J and Hofnung M (1980b) Sequence analysis of mutations that prevent export of lambda receptor, an *Escherichia coli* outer membrane protein. Nature 285: 82-85

Emr S.D, Schwartz M and Silhavy T.J (1978) Mutations altering the cellular location of the phage lambda receptor, an *Escherichia coli* outer membrane protein. Proc. Nat. Acad. Sci. 75(12): 5802-5806

Emr S.D and Silhavy T.J (1980) Mutations affecting localisation of an *Escherichia coli* outer membrane protein, the bacteriophage lambda receptor. J. Mol. Biol. 141: 63-90

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

Emtage J.S, Angal S, Doel M.T, Harris T.J.R, Jenkins B, Lilley G and Lowe P.A (1983) Synthesis of calf prochymosin (prorennin) in E.coli. Proc. Nat. Acad. Sci. 80: 3671-3675

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

Fandl J and Tai P.C (1987) Biochemical evidence for the *secY24* defect in *Escherichia coli* protein translocation and its suppression by soluble cytoplasmic factors. Proc. Nat. Acad. Sci. 84: 7448-7452

Feinberg A.P and Vogelstein B (1984) A technique for radiolabelling DNA restriction fragments to high specific activity. Anal. Biochem. 137: 266-267

Felmlee T, Pellet S, Lee E-Y and Welch R.A (1985b) *Escherichia coli* haemolysin is released extracellularly without cleavage of a signal peptide. J. Bact. 163: 88-93

Felmlee T, Pellet S and Welch R.A (1985a) Nucleotide sequence of an *Escherichia coli* chromosomal hemolysin. J. Bact. 163: 94-105

Felmlee T and Welch R.A (1988) Alterations of amino acid repeats in the *Escherichia coli* hemolysin affect cytolytic activity and secretion. Proc. Nat. Acad. Sci. 85: 5269-5273

Ferro-Novick S, Honma M and Beckwith J (1984) The product of gene *secC* is involved in the synthesis of export proteins in *E. coli*. Cell 38: 211-217

Fling S.P and Gregerson D.S (1986) Peptide and protein molecular weight determination by electrophoresis usig a high molarity Tris buffer system without Urea. Analytical Biochem. 155: 83-88

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

Freudl R, Schwartz H, Klose M, Movva N.R and Henning U (1985) The nature of information, required for export and sorting, present within the outer membrane porin OmpA of *Escherichia coli* K12. EMBO J. 4: 3593-3598

Freudl R, Schwarz H, Kramps S, Hindennach I and Henning U (1988) Dihydrofolate reductase (Mouse) and beta-galactosidase (*Escherichia coli*) can be translocated across the plasma membrane of *E. coli*. J. Biol. Chem. 263: 17084-17091

Freudl R, Klose M and Henning U (1990) Export and sorting of the *Escherichia coli* outer membrane protein OmpA. J. of Bioenergetics and Biomembranes 22: 441-450

Freudl R, Schwartz H, Stierhof Y-D. Gamon K, Hindennach I and Henning U (1986) An outer membrane protein (OmpA) of *Escherichia coli* K-12 undergoes a conformational change during export. J. Biol. Chem. 261: 11355-11361

Gardel C, Benson S, Hunt J, Michaelis S and Beckwith J (1987) *secD*, a new gene involved in protein export in *Escherichia coli*. J. Bact. 169: 1286-1290

Gardel C, Johnson K, Jacq A and Beckwith J (1990) The *secD* locus of *E. coli* codes for two membrane proteins required for protein export. EMBO J. 9: 3209-3216

Geller B.L, Movva N.R and Wickner W (1986) Both ATP and the electrochemical potential are required for optimal assembly of pro-OmpA into *Escherichia coli* inner membrane vesicles. Proc. Nat. Acad. Sci. 83: 4219-4222

Gennity J, Goldstein J and Inouye M (1990) Signal peptide mutant of *Escherichia coli*. J. of Bioenergetics and Biomembranes 22: 233-269

Gentz R, Kuys Y, Zwieb C, Taatjes D, Taatjes H, Bannwarth W, Stueber D and Ibrahimi I (1988) Association of degradation and secretion of three chimeric polypeptides in *Escherichia coli*. J. Bact. 170: 2212-2220

Gilmore R and Blobel G (1985) Translocation of secretory proteins across the microsomal membrane occurs through an environmental accessible to aqueous perturbants. Cell 42: 497-505

Gilmore R, Walter P and Blobel G (1982) Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. J. Cell. Biol. 95: 463-469

Glaser P, Ladant D, Sezer O, Pichot F, Ullmann A, and Danchin A (1988) The calmodulin sensitive adenylate cyclase of *Bordetella pertussis*: cloning and expression in *E. coli*. Mol. Microbiol. 2: 19-30

Gray L (1987) Functional analysis of the *Escherichia coli* haemolysin and its export. Ph.D Thesis, Dept. of Genetics Leicester University, U.K

Gray L, Baker K, Kenny B, Mackman N, Haigh R and Holland I.B (1989) A novel C-terminal signal sequence targets *Escherichia coli* haemolysin directly to the medium. J. Cell. Sci. Supp. 11; 45-57

Gray L, Mackman N, Nicaud J-M and Holland I.B (1986) The carboxy-terminal region of haemolysin 2001 is required for secretion of the toxin from *Escherichia coli*. Mol. Gen. Genet. 205: 127-133

Grunstein M and Hogness D.S (1975) Colony hybridisation: a method for the isolation of cloned cDNAs that contain a specific gene. Proc. Natl. Acad. Sci. 72: 3961-3965

Hall M.N and Silhavy T.J (1981) Genetic analysis of the major outer membrane proteins of *Escherichia coli*. Ann. Rev. Gen. 15: 91-142

Harris T.J.R, Lowe P.A, Lyons A, Thomas P.G, Eaton M.A.W, Millican A, Patel T.P, Bose C.C, Carey N.H and Doel M.T (1982) Molecular cloning and nucleotide sequencing of cDNA coding for calf prochymosin. Nuc. Acids. Res 10(7): 2177-2187

Hartl U-F and Neupert W (1990) Protein sorting to mitochondria: Evolutionary conservations of folding and assembly. Science 247: 930-938

Hartlein M, Schiessel S, Wagner W, Rdest U, Kreft J and Goebel W (1983) Transport of hemolysin by *Escherichia coli*. J. Cell. Biochem. 22: 87-97

Hashimoto F, Horigome T, Kanbayashi M, Yoshida K and Sugano H (1983) An improved method for separation of low molecular weight polypeptides by electrophoresis in SDS-polyacrylamide gels. Analytical Biochem. 129: 192-199

Hess J, Wels W, Vogel M, and Geobel W (1986) Nucleotide sequence of a plasmid-encoded hemolysin determinant and its comparison with a corresponding chromosomal hemolysin sequence. FEMS Microbiol. Lett. 34: 1-11

Higgins C.F, Hiles I.D, Whalley K and Jamieson D.J.(1985) Nucleotide binding by membrane components of bacterial periplasmic binding protein-dependent transport systems. EMBO J. 4:1033-1040

Higgins C.F, Hiles I.D, Salmond G.P.C, Gill D.R, Downie J.A, Evans I.J, Holland I.B, Gray L, Buckel S.D, Bell A.W and Hermodsen M.A (1986) A family of related ATP-binding subunits coupled to many distinct biological processes in *E. coli*. Nature 232: 448-450

Hines V. Brandt A, Griffiths G, Horstmann H, Brutsch H and Schatz G (1990) Protein import into yeast mitochondria is accelerated by the outer membrane protein MAS70. EMBO J. 9: 3191-3200

Hirst T.R and Holmgren J (1987a) Conformation of proteins secreted across bacterial outer membranes: A study of enterotoxin translocation from *Vibrio cholerae*. Proc. Nat. Acad. Sci. 84: 7418-7422

Hirst T.R and Holmgren J (1987b) Transient entry of enterotoxin subunits into the periplasm occurs during their secretion from *Vibrio cholerae*. J. Bact. 169: 1037-1045

Hobot J.A, Carlemalm E, Villiger W and Kellenberger E (1984) Periplasmic Gel: New concept resulting from the re-investigation of the bacterial cell envelope ultrastructure by new methods. J. Bact. 160: 143-152

Hoffman C.S and Wright A (1985) Fusions of secreted proteins to alkaline phosphatase: an approach to studing protein secretion. Proc. Nat. Acad. Sci. 82: 5107-5111

Holland I.B, Blight M and Kenny B (1990a) The mechanism of secretion of hemolysin and other polypeptides from Gram-negative bacteria. J. of Bioenergetics and Biomembranes 22: 473-491

Holland I.B, Kenny B, and Blight M (1990b) Haemolysin secretion from *E. coli*. Biochimie 72: 131-141

Holland I.B, Kenny B, Steipe B and Pluckthun A (1990c) Secretion of heterologous proteins in *Escherichia coli*. Meths. in Enzymol. 182: 132-143

Holland I.B, Mackman N, Gray L and Bhakdi S (1987) Secretion and mode of action of *Escherichia coli* haemolysin. Membrane-Mediated Cytotoxicity (Alan R. Liss, Inc) p9-18

Holland I.B, Wang R-C, Seror S.J, and Blight M (1989) Haemolysin secretion and other protein translocation mechanisms in Gram negative bacteria. Society General Microbiology Symposium 44. Cambridge University Press pp219-254

Horwich A.L, Kalousek E, Mellman I and Rosenberg L.E (1985) A leader peptide is sufficient to direct mitochondrial import of a chimeric protein. EMBO J. 4: 1129-1135

Howard S.P and Buckley J.T (1985) Protein export by a Gram negative bacterium: production of Aerolysin by *Aeromanas hydrophila*. J. Bact. 161: 1118-1124

Howard S.P and Buckley J.T (1986) Molecular cloning and expression in *Escherichia coli* of the structural gene for the haemolytic toxin aerolysin from *Aeromanas hydrophila*. Mol. Gen. Genet. 204: 289-295

Humphreys G.O, Willshaw G.A, Smith H.R and Anderson E.S (1976) Mutagenesis of plasmid DNA with hydroxylamine: isolation of mutants of multi-copy plasmids. Mol. Gen. Genet. 145: 101-108

Hurt E.C, Pesold-Hurt B and Schatz G (1984) The cleavage prepiece of an imported mitochondrial protein is sufficient to direct cytosolic dihydrofolate reductase into the mitochondrial matrix. FEBS Lett. 178: 306-310

Hurt E.C and van Loon A.P.C.M (1986) How proteins find mitochondria and intramitochondrial compartments. Trends Biol. Science 11: 204-207

Hussain M, Ozawa Y, Ichihara S and Mizushima S (1982) Signal peptide digestion in *Escherichia coli*. Eur. J. Biochem. 129: 233-239

Hutchisin III C.A, Nordeen S.K, Vogt K and Edgell M.H (1986) A complete library of point substitution mutations in the glucocorticoid response element of mouse mammary tumor virus. Proc. Natl. Acad. Sci. 83: 710-714

Hsu I-N, Delbaere L.T.J, James M.N.G and Hofmann T (1977) Penicillopepsin from *Penicillium janthinellum* crystal structure at 2.8A and sequence homology with porcine pepsin. Nature 266: 140-145

Inouye S, Soberon X, Franceshini T, Nakamura K. Itakura K and Inouye M (1982) Role of the positive charge on the animo-terminal region of the signal peptide in protein secretion across the membrane. Proc. Nat. Acad. Sci. 79: 3438-3441

Ito K, Wittekind M, Nomura M, Shiba K, Yura T, Muira A and Nashimoto H (1983) A temperature sensitive mutant of *E. coli* exhibiting slow processing of export proteins. Cell 32: 789-797

Jackson M.E and Pratt J.M (1987) An 18 amino acid amphiphilic helix forms the membrane-anchoring domain of the *Escherichia coli* penicillin-binding protein. Mol. Micro. 1: 23-28

Jackson M.E, Pratt J.M and Holland I.B (1986) Intermediates in the assembly of the TonA polypeptide into the outer membrane of *Escherichia coli* K12. J. Mol. Biol. 189: 477-486

Jackson M.E, Pratt J.M, Stoker N.G and Holland I.B (1985) An inner membrane protein N-terminal signal sequence is able to promote efficient localisation of an outer membrane protein in *Escherichia coli*. EMBO J. 4: 2377-2383

Jones J.D.G, Grady K.L, Suslow T.V and Redbrook J.R (1986) Isolation and characterisation of genes encoding two chitinase enzymes from *Serratia marcescens*. EMBO J. 5: 467-473

Josefsson L and Randall L.L (1981) Different export proteins in *Escherichia coli* show differences in the temporal mode of processing *in vivo*. Cell 25: 151-157

Kaput J, Goltz S and Blobel G (1982) Nucleotide sequence of the yeast nuclear gene for cytochrome c peroxidase precursor. J. Biol. Chem. 257: 15054-15058

Killian A, de Jong A.M.Ph, Bijvelt J, Verkleij J and de Kruijff B (1990) Induction of non-bilayer lipid structures by signal peptides. EMBO J. 9: 815-819

Keng T, Alani E and Guarente L (1986) The nine amino-terminal residues of γ-Aminolevulinate synthatase directs beta-galactosidase into mitochondria. Mol. Cell Biol. 6: 355-364

Kornberg R.D and McConnell H.M (1971) Inside-outside transitions of phospholipids in vesicle membranes. Biochemistry 10: 1111-1120

Koronakis V, and Hughes C (1988) Identification of promoters directing *in vivo* expression of hemolysin genes in *Proteus vulgaris* and *Escherichia coli*. Mol. Gen. Genet. 213: 99-104

Koronakis V, Cross M, and Hughes C (1988) Expression of the *E. coli* hemolysin secretion gene *hlyB* involves transcript anti-termination within the *hly* operon. Nuc. Acid. Res. 16: 4789-4800

Koronakis V, Cross M, Senior B, Koronakis E and Hughes C (1987) The secreted hemolysins of *Proteus mirabilis*, *Proteus vulgaris*, and *Morganella morganii* are genetically related to each other and to the alpha-hemolysin of *Escherichia coli*. J. Bact. 169: 1509-1515

Koronakis V, Koronakis E, and Hughes C (1989) Isolation of the C-terminal signal directing export of *E. coli* haemolysin protein across both bacterial membranes. EMBO J. 8: 595-605

Koshland D and Botstein D (1980) Secretion of beta-lactamase requires the carboxy end of the protein. Cell 20: 749-760

Koshland D and Botstein D (1982) Evidence for post-translational translocation of beta-lactamase across the bacterial inner membrane. Cell 30: 893-902

Krieg U.C, Walter P and Johnson A.E (1986) Photocrosslinking of the signal sequence of nascent preprolactin to the 54KD polypeptide of the signal recognition particle. Proc. Nat. Acad. Sci. 83: 8604-8608

Kuhn A and Wickner W (1985) Conserved residues of the leader peptide are essential for cleavage by leader peptidase. J. Biol. Chem. 260(29) 15914-15918

Kumamota C.A and Beckwith J (1983) Mutations in a new gene, *secB*, cause defective protein localisation in *Escherichia coli*. J. Bact. 154: 253-260

Kusukawa N, Yura T, Ueguchi C, Akiyama Y and Ito K (1989) Effects of mutations in heat-shock genes *groES* and *groEL* on protein export in *Escherichia coli*. EMBO J. 8: 3517-3521

Laemmli U.K (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685

Lazdunski A, Guzzo J, Filloux A, Bally M and Murgier M (1990) Secretion of extracellular proteins by *Pseudomonas aeruginosa*. Biochimie 72: 147-156

Lecker S, Lill R, Ziegelhoffer T, Georgopoulos C, Bassford Jr P.J, Kumamoto C.A and Wickner W (1989) Three pure chaperone proteins of *Escherichia coli* - SecB, trigger factor and GroEL - form soluble complexes with precursor proteins *in vitro*. EMBO J. 8: 2703-2709

Lee C.A and Beckwith J (1986) Suppression of growth and protein secretion defects in *Escherichia coli secA* mutants by decreasing protein synthesis. J. Bact. 166: 878-883

Lee C, Ping L, Inouye H, Brickman E.R and Beckwith J (1989) Genetic studies on the inability of beta-galactosidase to be translocated across the *Escherichia coli* cytoplasmic membrane. J. Bact. 171: 4609-4616

Letoffe S, Delepelaire P and Wandersman C (1990) Protease secretion by *Erwinia chrysanthemi*: the specific secretion functions are analogous to those of *Escherichia coli* alpha haemolysin. EMBO J. 9: 1375-1382

Lill R, Cunningham K, Brundage L.A, Ito K, Oliver D and Wickner W (1989) SecA protein hydrolyses ATP and is an essential component of the protein translocation ATPase of *Escherichia coli*. EMBO J. :961-966

Liss L.R and Oliver D.B (1986) Effects of *secA* mutations on the synthesis and secretion of proteins in *Escherichia coli*. J. Biol. Chem. 261: 2299-2303

Little S, Campbell C.J, Evans I.J, Hayward E.C, Lilley R.J and Robinson M.K (1989) A short N-terminal region of prochymosin inhibits the secretion of hybrid proteins from *Escherichia coli*. Gene 8: 321-329

Lo R.Y.C, Strathdee C.A, and Shewen P.E (1987) Nucleotide sequence of the leukotoxin genes of Pasteurella haemolytica A1. Infect. Immun. 55: 1987-1996

Lory S and Tai P (1983) Characterisation of the Phospholipase C gene of *Pseudomonas aeruginosa* cloned in *Escherichia coli*. Gene 22: 95-101

Lory S, Tai P and Davis BD (1983) Mechanism of protein excretion by Gram negative bacteria: *Pseudomonas aeruginosa* Exotoxin A. J. Bact. 156: 695-702

Ludwig A, Jarchau T, Benz R and Goebel W (1988) The repeat domain of *Escherichia coli* haemolysin (HlyA) is responsible for its Ca^{2+}-dependent binding to erythrocytes. Mol. Gen. Genet. 214: 553-561

Ludwig A, Vogel M and Goebel W (1987) Mutations affecting activity and transport of haemolysin in *Escherichia coli*. Mol. Gen. Genet. 206: 238-245

Lutenberg B and van Alphen L (1983) Molecular architecture and functioning of the outer membrane of *Escherichia coli* and other Gram negative bacteria. Biochim et Biophys Acta 737: 51-115

McCaman M.T, Andrews W.H and Files J.C (1985) Enzymatic properties and processing of bovine prochymosin synthesised in *E.coli*. J. of Biotechnology 2: 177-190

Mackman N (1984) Identification and secretion of the *Escherichia coli* haemolysin. Ph.D Thesis Dept. of Genetics, University of Leicester

Mackman N, Baker K, Gray L, Haigh R, Nicaud J.M and Holland I.B (1987) Release of a chimeric protein into the medium from Escherichia coli using the C-terminal secretion signal sequence. EMBO J. 6: 2835-2841

Mackman N and Holland I.B (1984a) Secretion of a 107K dalton polypeptide into the medium from a haemolytic *E.coli* K12 strain. Mol. Gen. Genet. 193: 312-315

Mackman N and Holland I.B (1984b) Functional characterisation of a cloned haemolysin determinant from *E. coli* of human origin, encoding information for the secretion of a 107KD polypeptide. Mol. Gen. Genet. 196: 123-134

Mackman N, Nicaud J-M, Gray L and Holland I.B (1985a) Genetical and functional organisation of the *Escherichia coli* haemolysin determinant LE2001. Mol. Gen. Genet. 201: 282-288

Mackman N, Nicaud J-M, Gray L and Holland I.B (1985b) Identification of polypeptides required for the export of haemolysin 2001 from *E.coli*. Mol. Gen. Genet. 201: 529-536

Mackman N, Nicaud J-M, Gray L and Holland I.B (1986) Secretion of Haemolysin by *Escherichia coli*. Current Topics in Micro. and Immun. 125: 159-181

Maniatis T, Fritsch E.F and Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor

Masure H.R, Au D.C, Gross M.K, Donovan M.G, and Storm D.R (1990) Secretion of *Bordetella pertussis* adenylate cyclase from *Escherichia coli* containing the hemolysin operon. Biochemistry 29: 140-145

Mayer P.A and Singer S.J (1986) Disulfide bonds and the translocation of proteins across membranes. Proc. Nat. Acad. Sci. 83: 9001-9005

Menestrina G (1988) *Escherichia coli* hemolysin permeabilises small unilamellar vesicles loaded with calcein by a single-hit mechanism. FEBS 232: 217-220

Messing J (1983) New M13 vectors for cloning. Meths. in Enzymol. 101: 20-78

Metcalfe M and Holland I.B. (1980) Synthesis of a major outer membrane porin by *Escherichia coli* spheroplasts. F.E.M.S Micro. Lett. 7: 111-114

Meyer D.I (1985) Signal recognition particle (SRP) does not mediate a translational arrest of nascent secretory protein in mammalian cell-free systems. EMBO J. 4: 2031-2033

Meyer D.I, Krause E and Dobberstein B (1982) Secretory protein translocation across membranes - the role of the "docking protein". Nature 297: 647-650

Michaelis S, Inouye H, Oliver D and Beckwith J (1983) Mutations that alter the signal sequence of alkaline phosphatase in *Escherichia coli*. J. Bact. 154: 366-374

Milstein C, Brownlee G.G, Harrison T.M and Mathews M.B (1972) A possible precursor of immunogloblin light chain. Nat. New. Biol. 239: 117-120

Mug-Opstelten D and Witholt B (1978) Preferential release of new outer membrane fragments by exponential growing *Escherichia coli*. Biochim. Biophys. Acta. 508: 287-295

Muhlradt P.F, Menzel J, Golecki J.R, and Speth V (1973) Outer membrane of *Salmonella*: site of export of newly synthesised Lipopolysaccharide on the bacterial surface. Eur. J. Biochem. 35: 471-481

Muller M and Blobel G (1984) Protein export in *Escherichia coli* requires a soluble activity. Proc. Nat. Acad. Sci. 81: 7737-7741

Muller D, Hughes C and Goebel W (1983) Relationship between plasmid and chromosomal hemolysin determinants of *Escherichia coli*. J. Bact. 153: 846-851

Murakami K and Mori M (1990) Purified presequence binding factor (PBF) forms an import-competent complex with a purified mitochondrial precursor protein. EMBO J. 9: 3201-3208

Nakahama K, Yoshimura K, Marumoto R, Kikuchi M, Lee I.S, Hase T, and Matsubura H (1986) Cloning and sequencing of *Serratia protease* gene. Nuc. Acids. Research 14: 5843-5855

Neame K.D and Homewood C.A (1974) Introduction to liquid scintillation.

Nicaud J-M, Mackman N, Gray L, and Holland I.B (1985a) Characterisation of HlyC and the mechanism of activation and secretion of the haemolysin from *E. coli* 2001. FEBS Letters 187: 339-344

Nicaud J-M, Mackman N, Gray L, and Holland I.B (1985b) Regulation of haemolysin synthesis in *E. coli* determined by HLY genes of human origin. Mol. Gen. Genet. 199: 111-116

Nicaud J-M, Mackman N, Gray L and Holland I.B (1986) The C-terminal, 23kDa peptide of E.coli haemolysin 2001 contains all the information necessary for its secretion by the haemolysin (Hly) export machinery. FEBS 204(2): 331-335

Noegel A, Rdest U, Springer W and Goebel W (1979) Plasmid cistrons controlling synthesis and excretion of the exotoxin alpha-haemolysin of *Escherichia coli*. Mol. Gen. Genet. 175: 343-350

Normington K, Kohno K, Kozutsumi Y, Gething M-J and Sambrook J (1989) S.cerevisiae encodes an essential protein homologous in sequence and function to mammalian BIP. Cell 57: 1223-1236

Norrander J, Kempe T and Messing J (1983) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene 26: 101-106

Oliver D.B (1987) Molecular architecture and assembly of cell parts: Periplasm and protein secretion. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular biology Part 1 (ed Neidhardt FC) p 56-69

Oliver D.B and Beckwith J (1981) *E. coli* mutant pleiotropically defective in the export of secreted proteins. Cell 25: 765-722

Oliver D.B and Beckwith J (1982) Regulation of a membrane component required for protein secretion in *Escherichia coli*. Cell 30: 311-319

Oropeza-Wekerle R.L, Muller E, Kern P, Meyermann R and Goebel W (1989) Synthesis, inactivation and localisation of extracellular and intracellular *Escherichia coli* hemolysins. J. Bact. 171: 2783-2788

Oropeza-Wekerle R.L, Speth W, Imhof B, Gentschev I and Goebel W (1990) Translocation and compartmentalisation of *Escherichia coli* hemolysin (HlyA). J. Bact. 172: 3711-3717

Ostermann J, Horwich A.L, Neupert W and Hartl F-U (1989) Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis. Nature 341:125-130

Owen P and Salton M.R (1975) Antigenic and enzymatic architecture of Micrococcus lysodeikticus membranes established by crossed immunoelectrophoresis. Proc. Nat. Acad. Sci. 72: 3711-3715

Pain D, Murakami H and Blobel G (1990) Identification of a receptor for protein import into mitochondria. Nature 347: 444-449

Pedersen S, Bloch P.L, Reeh S and Neidhardt F.C (1978) Patterns of protein synthesis in *Escherichia coli*: a catalog of the amount of 140 individual proteins at different growth rates. Cell 14: 179-190

Pedersen V.B, Christensen K.A and Foltmann V (1979) Investigations on the activation of bovine prochymosin. Eur. J. Biochem 94: 573-580

Pelham H.R.S, (1989) Heat shock and the sorting of luminal endoplasmic reticular proteins. EMBO J. 8: 3171-3176

Perara E, Rothman R.E and Lingappa V.R (1986) Uncoupling translocation from translation: Implication for transport of proteins across membranes. Science 232: 348-352

Pfaller et al (1988) Import pathways of precursor proteins into mitochondria: multiple receptor sites followed by a common membrane insertion site. J. Cell. Biol. 107: 2483

Pfanner N, Tropschug M and Neupert W (1987) Mitochondrial protein import: Nucleoside triphosphates are involved in conferring import-competence to precursors. Cell 49: 815-823

Phillips G.J and Silhavy T.J (1990) Heat-shock proteins DnaK and GroEL facilitate export of LacZ hybrid proteins from *Escherichia coli*. Nature 344: 882-884

Pohlner J, Halter R, Beyreuther K and Meyer T.F (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature 325: 458-462

Pratt J.M, Jackson M.E and Holland I.B (1986) The C-terminus of penicillin-binding protein 5 is essential for localisation to the *Escherichia coli* inner membrane. EMBO J. 5: 2399-2405

Pruss G.J and Drlica K (1989) DNA supercoiling and prokaryotic transcription. Cell 56: 521-523

Pugsley A.P, Chapon C and Schwartz M (1986) Extracellular Pullulanase of *Klebsiella pneumoniae* is a lipoprotein. J. Bact. 166: 1083-1088

Pugsley A.P, Kornacker M.G and Ryter A (1990) Analysis of the subcellular location of pullulanase produced by *Escherichia coli* carrying the *pulA* gene from *Klebsiella pneumoniae* strain UNF5023. Mol. Micro. 1: 59-72

Pugsley A.P and Reyss I (1990) Five genes at the 3' end of the *Klebsiella pneumoniae* pulC operon are required for pullulanase secretion. Mol. Micro. 4: 365-379

Pugsley A.P and Schwartz M (1985) Export and secretion of proteins by bacteria. FEMS Microbiol. Reveiws 32: 3-38

Randall L.L (1983) Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. Cell 33: 231-240

Randall L.L and Hardy S.J.S (1986) Correlation of competence for export with the lack of tertiary structure of the mature species: A study *in vivo* of Maltose-binding protein in *E.coli*. Cell 46:921-928

Randall L.L, Hardy S.J.S and Thom J.R (1987) Export of proteins: A biochemical view. Ann. Review Micro. 41: 507-542

Rassow R, Guiard B, Wienhues U, Hartl F-U and Neupert W (1989) Translocation arrest by reversible folding of a precursor protein imported into mitochondria. A means to quantitate translocation contact sites. J. Cell. Biol. 109: 1421-1428

Redman C.M and Sabatini D.D (1966) Vectorial discharge of peptides released by puromycin from attached ribosomes. Proc. Nat. Acad. Sci. 56: 608-615

Redman C.M, Siekevitz P and Palade G.E (1966) Synthesis and transfer of amylase in pigeon pancreatic microsomes. J. Biol. Chem. 241: 1150-1158

Riggs P.D, Derman A.I and Beckwith J (1988) A mutation affecting the regulation of a *secA-lacZ* fusion defines a new *sec* gene. Genetics 118: 571-579

Rothblatt J.A and Meyer D.I (1986) Secretion in yeast: translocation and glycosylation of prepro-alpha-factor *in vitro* can occur via an ATP-dependent post-translational mechanism. EMBO J. 5: 1031-1036

Rothman J.E and Kornberg R.D (1986) An unfolding story of protein translocation. Nature 322: 209-210

Sanger F, Nicklen S.K, and Coulson A.R (1977) DNA sequencing with chain termination inhibitors. Proc. Nat. Acad. Sci. 74: 5463-5467

Sato T, Ito K and Yura T (1977) Membrane proteins of *Escherichia coli* K-12: Two dimensional polyacrylamide gel electrophoresis of inner and outer membranes. Eur. J. Biochem. 78: 557-567

Schatz P.J, Riggs P.D, Jacq A, Fath M.J and Beckwith J (1989) The *secE* gene encodes an integral membrane protein required for export in *Escherichia coli*. Genes Dev. 3: 1035-1044

Schechter I, McKean D.J, Guyer R and Terry W (1975) Partial amino acid sequence of the precursor of immunoglobulin light chain programmed by mRNA *in vitro*. Science 188: 160-162

Schiffer M and Edmunson A.B (1967) Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. Biophys J. 7 121-135

Schleyer M and Neupert W (1985) Transport of proteins into mitochondria: Translocation intermediates spanning contact sites between outer and inner membranes. Cell 43: 339-350

Schoemaker J.M, Brassnett A.H and Marston F.A.O (1985) Examination of calf prochymosin accumulation in *E. coli*: disulphide linkages are a structural component of prochymosin containing inclusion bodies. EMBO J. 4: 775-780

Schwaiger M, Herzog V and Neupert W (1987) Characterisation of translocation contact sites involved in import of mitochondrial proteins. J. Cell. Biol. 105: 235-246.

Schwartz R.M and Dayhoff M.O (1978) Origins of prokaroytes, eucaryotes, mitochondria and chloroplasts: A perspective is derived from protein and nucleic acid sequence data. Science 199: 395-403

Scultz J, Silhavy T.J, Berman M.L, Fiil N and Emr S.D (1982) A previously identified gene in the *spc* operon of *Escherichia coli* K-12 specifies a component of the protein export machinery. Cell 31: 227-235

Sczakiel G, Wittinghofer A and Tucker J (1987) An efficient expression vector (pPLEX) for bacterial synthesis of non-fusion proteins and protein domains. Nuc. Acids. Res 15: 1878

Sen K and Nikaido H (1990) In vitro trimerization of OmpF secreted by spheroplasts of *Escherichia coli*. Proc. Nat. Acad. Sci. 87: 743-747

Sharma S.K (1986) On the recovery of genetically engineered proteins from Escherichia coli. Separation Science and Technology 21(8): 701-726

Shiba K, Ito K and Yura T (1986) Suppressors of the *secY*24 mutation: Identification and characterisation of additional *ssy* genes in *Escherichia coli*. J. Bact. 166: 849-856

Singer S.J, Maher P.A and Yaffe M.P (1987) On the translocation of proteins across membranes. Proc. Nat. Acad. Sci. 84: 1015-1019

Singer S.J and Nicolson (1972) The fluid mosaic model of the structure of cell membranes. Science 175: 720-731

Smit J and Nikaido H (1978) Outer membrane of Gram negative bacteria XVIII Electron microscopic studies on porin insertion sites and growth of cell surface of *Salmonella typhimurium*. J.Bact. 135: 687-702

Smith H.W (1963) The haemolysins of *Escherichia coli*. J. Pathol. Bacteriol. 85: 197-211

Smith W.P, Tai P.C, Thompson R.C and Davis B.D (1977) Extracellular labeling of nascent polypeptides traversing the membrane of *Escherichia coli*. Proc. Nat. Acad. Sci. 74: 2830-2834

Sollner T, Griffiths G, Pfaller R, Pfanner N and Neupert W (1989) MOM19, an import receptor for mitochondrial precursor proteins. Cell 59: 1061-1070

Springer W and Goebel W (1980) Synthesis and secretion of hemolysin by *Escherichia coli*. J. Bact. 144: 53-59

Stanley K.K and Luzio J.P (1984) Construction of a new family of high efficiency bacterial expression vectors: Identification of cDNA clones coding for human liver proteins. EMBO J. 3: 1429-1434

Stark M.J.R (1987) Multicopy expression vectors carrying the lac repressor gene for regulating high level expression of genes in E.coli. Gene 51: 255-267

Stoker N.G, Fairweather N.F and Spratt B.G (1982) Versatile low-copy-number plasmid vectors for cloning in *Escherichia coli*. Gene 18: 335-341

Strathdee C.A, and Lo R.Y.C (1989) Cloning, nucleotide sequence and characterisation of gene encoding the secretion functions of the *Pasteurella haemolytica* leukotoxin determinant. J. Bact. 171: 916-928

Strauch K.L, Kumamoto C.A and Beckwith J (1986) Does *secA* mediate coupling between secretion and translocation in *Escherichia coli*. J. Bact. 166: 505-512

Studier F.W, Rosenberg A.H, Dunn J.J and Dubendorff (1990) Use of T7 RNA polymerase to direct the expression of cloned genes. Meths. in Enzymol. 185: 60-89

Swinkels B.A, Evers R and Piet B (1988) The topogenic signal of the glycosomal (microsome) phosphoglycerate kinase of *Crithidia fasciculata* resides in a carboxy-terminal extension. EMBO J. 7:1159-1165

Tabe L, Krieg P, Strachan R, Jackson D, Wallis E and Colman A (1984) Segregration of mutant ovalbumins and ovalbumin-globin fusion proteins in Xenopus oocytes: Identification of an ovalbumin signal sequence. J. Mol. Biol. 180: 645-666

Talmadge K, Brosius J and Gilbert W (1981) An internal signal sequence directs secretion and processing of proinsulin in bacteria. Nature 294: 176-178

Talmadge K, Kaufman J and Gilbert W (1980b) Bacteria mature proinsulin. Proc. Nat. Acad. Sci. 77: 3988-3999

Talmadge K, Stahl S and Gilbert W (1980a) Eukaryotic signal sequence transports insulin antigen in *E. coli*. Proc. Nat. Acad. Sci. 77: 3369-3373

Tokunaga M, Tokunaga H and Wu HC (1982) Post-translational modification and processing of *E. coli* prolipoprotein *in vitro*. Proc. Nat. Acad. Sci. 79: 2255-2259

Tommassen J and Lugtenberg B (1984) Amino terminus of outer membrane PhoE protein: Localisation by use of *bla-phoE* hybrid gene. J. Bact. 157: 327-329

Towbin H, Staehelin T and Gordon J (1979) Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose: Procedure and some applications. Proc. Natl. Acad. Sci. 76: 4350-4354

Van Bogelen R and Neidhardt F.C (1990) Ribosomes as sensors of heat and cold shock in *Escherichia coli*. Proc. Natl. Acad. Sci. 87: 5589-5593

Vestweber D, Brunner J, Baker A and Schatz G (1989) A 42K outer-membrane protein is a component of the yeast mitochondrial protein import site. Nature 341:205-209

Vestweber D and Schatz G (1988a) A chimeric mitochondrial precursor protein with internal disulfide bridges blocks the import of authentic precursors into mitochondria and allows quantification of import sites. J. Cell. Biol. 107: 2037-2043

Vestweber D and Schatz G (1988b) Point mutations destabilizing a precursor protein enhance its post-translational import into mitochondria. EMBO J. 7: 1147-1151

Viebrook A, Perz A and Sebald W (1982) The import preprotein of the proteolipid subunit of the mitochondrial ATP synthatase from Neurospora crassa. Molecular cloning and sequencing of the mRNA. EMBO J. 1: 565-571

Vlasuk G.P, Inouye S, Ito H, Itakura K and Inouye M (1983) Effect of the complete removal of basic amino acids from the signal sequence on secretion of Lipoprotein in *Escherichia coli*. J. Biol. Chem. 258: 7141-7148

von Heijne G (1983) Patterns of amino acids near signal sequence cleavage site. Eur. J. Biochem. 133: 17-31

von Heijne G (1984) Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells. EMBO J. 3:2315-2319

von Heijne G (1985) Signal sequences: The limits of variation. J. Mol. Biol. 184: 99-105

von Heijne G (1986) Net N-C charge inbalance may be important for signal sequence function in bacteria. J. Mol. Biol. 192: 287-290

von Heijne G and Blomberg C (1979) Trans-membrane translocation of proteins. The direct transfer model. Eur. J. Biochem. 97: 175-181

Vogel M, Hess J, Then I, Juarez A and Goebel W (1988) Characterisation of a sequence (*hlyR*) which enhances synthesis and secretion of hemolysin in *Escherichia coli*. Mol. Gen. Genet. 212: 76-84

Vogelstein B and Gillespie D (1979) Preparative and analytical purification of DNA fron agarose. Proc. Natl. Acad. Sci. 76: 615-619

Voorhout W, de Kroon T, Leunissen-Bijvelt, Verkleij A and Tommassen J (1988) Accumulation of LamB-lacZ hybrid proteins in intracytoplasmic membrane like structures in *Escherichia coli* K12. J. Gen. Micro. 134: 599-604

Wagner W, Kuhn M and Goebel W (1988) Active and inactive forms of hemolysin (HlyA) from *Escherichia coli*. Biol. Chem. Hoppe-Seyler 369: 39-46

Wagner W, Vogel M and Goebel W (1983) Transport of hemolysin across the outer membrane of *Escherichia coli* requires two functions. J. Bact. 154: 200-210

Walter P and Blobel G (1981) Translocation of proteins across the endoplasmic reticulum. Parts I.II and III. J. Cell. Biol. 91: 545-561

Walter P and Blobel G (1982) Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. Nature 299: 691-698

Walter P, Gilmore R and Blobel G (1984) Protein translocation across the endoplasmic reticulum. Cell 38: 5-8

Wandersman C, and Delepelaire P (1990) TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. Proc. Natl. Acad. Sci. 87: 4776-4780

Watson M.E.E (1984) Compilation of published signal sequences. Nuc. Acids. Res 12(13): 5145-5164

Welch R.A, Dellinger EP, Minshew B and Falkow S (1981) Haemolysin contributes to virulence of extraintestinal *E. coli* infections. Nature 294: 663-668

Welch R.A, Hall R and Falkow S (1983) Molecular cloning and physical characterisation of a chromosomal hemolysin from *Escherichia coli*. Infect. Immun. 42: 178-186

Welch R.A and Pelett S (1988) Transcriptional organisation of the *Escherichia coli* hemolysin genes. J. Bact. 170: 1622-1630

Wickner N (1979) The assembly of proteins into biological membranes: The membrane trigger hypothesis. Ann. Review Biochem. 48: 23-45

Wiedmann M, Kurzchalia T.V, Hartmann E and Rapoport T.A (1987) A signal sequence receptor in the endoplasmic reticulum membrane. Nature 328: 830-833

Wolfe P.B, Silver P and Wickner W (1982) The isolation of homogeneous leader peptidase from a strain of *Escherichia coli* which overexpresses the enzyme. J. Biol. Chem. 257 (13): 7898-7902

Wolfe P.B, Rice M and Wickner W (1985) Effects of two *sec* genes on protein assembly into the plasma membrane of *Escherichia coli* J. Biol. Chem. 260: 1836-1841.

Wolfe P.B and Wickner W (1984) Bacterial leader peptidase; A membrane protein without a leader peptide, uses the same export pathway as pre-secretory pathways. Cell 36: 1067-1072

Wolfe P.B, Wickner W and Goodman J.M (1983) Sequence of the leader peptidase gene of *Escherichia coli* and the orientation of leader peptidase in the bacterial envelope. J. Biol. Chem. 258(19): 12073-12080

Wright E.M, Humphreys G.O and Yarranton G.T (1986) Dual-origin plasmids contain an amplifiable ColEI ori; temperature controlled expression of cloned genes. Gene 49: 311-321

Yanagida N, Uozumi T and Beppu T (1986) Specific excretion of *Serratia marcescens* protease through the outer membrane of *Escherichia coli*. J. Bact. 166: 937-944

Yanisch-Perron C, Vieira J and Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene 33: 103-119

Yarranton G.T, Wright E, Robinson M.K and Humphrey G.O (1984) Dual-origin plasmid vectors whose origin of replication is controlled by the coliphage lambda promoter pL. Gene 28:293-300

Yu F, Yamada H, Daishima K and Mizushima S (1984) Nucleotide sequence of *lsp*A gene, the structural gene for lipoprotein signal peptidase of *Escherichia coli*. FEBS Letters 173: 264-271

Yue K (1989) Assembly of outer membrane protein in *Escherichia coli*. Ph.D Thesis, Dept. of Biochemistry, University of Liverpool

Zimmerman R, Sagstetter M, Lewis M.J and Pelham H.R.B (1988) Seventy-kilodalton heat shock proteins and an additional component from reticulocyte lysate stimulate import of M13 procoat protein into microsomes. EMBO J. 7: 2875-2880