https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

APPLICATION OF PHAGE DISPLAY TO THE STUDY OF TOXIN-RECEPTOR INTERACTIONS

HECTOR ALEXANDER McLEAN

Thesis submitted for the degree of Doctor of Philosophy

Division of Infection and Immunity, University of Glasgow

September 1999

ProQuest Number: 10391479

All rights reserved

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

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

ProQuest 10391479

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

GLASGOW
U...VERSITY
LIBRARY

11769 (copy 2)

SUMMARY

In causing disease, most bacterial toxins must firstly interact specifically with one or more host receptors. Thus, discovery of the identity of these receptors is an important step towards elucidating a toxin's mode of action.

This project aimed to assess whether the technique known as phage display had any role to play in the identification of toxin receptors.

Libraries of DNA were supplied for this work. These libraries had degenerate DNA of defined sequence length cloned into genes encoding major or minor components of the capsid of filamentous phage of the Ff family. This resulted in random peptides being 'displayed' in the outer envelope of the phage.

Libraries of such phage were exposed to surfaces of immobilised toxin, with those carrying a random peptide sequence complementary to the receptor binding site interacting with, and being bound to, the toxin. Unbound phage were washed away and bound phage were removed by elution with a specific or non-specific agent.

The toxin used as a model in this work was the heat-labile toxin of enterotoxigenic *Escherichia coli* (LT), for which the natural receptor is already known. A diarrhoeal toxin almost identical to that from *Vibrio cholerae*, it recognises a glycolipid, ganglioside G_{M1}. This assessment therefore called for the peptide extension of the phage to mimic a natural carbohydrate moiety.

The structure of LT is designated AB_5, signifying that it has one **A**ctive and five **B**inding subunits. As this work was primarily concerned with binding, isolated LT-B pentamers, secreted by a genetically-modified marine *Vibrio* were used in the experimental work in preference to whole toxin. After examining the growth pattern of this strain, and relating it to B subunit production, the pentamers were purified by affinity chromatography using a galactose column, then biotinylated to facilitate attachment to streptavidin-coated plastic. The biotinylated product was tested for integrity and activity.

Three phage display libraries were assessed. After exposure to the toxin, the eluted phage were tested in binding assays and, where appropriate, were sequenced to determine the amino acid composition of the random insert .

The first library tested, the 'Smith hexamer', carried a six amino acid extension to minor capsid protein III and performed poorly: the selected clones showed no specific binding to the toxin. The second library, the 'Luzzago nonamer', which carried a nine amino acid extension to protein VIII , was an example of a library carried on a phagemid vector. It apparently had suffered a flaw in construction and appeared unable to synthesise coat protein. The third library, the 'Smith pentadecamer', carried a fifteen amino acid extension to protein VIII. Successive rounds of exposure to the toxin and amplification of binding phage demonstrated that, although selection of phage carrying specific peptides was occurring, the affinity was for elements of the test system other than LT-B itself.

Thus, these experiments were unable to identify peptides from a random, phage-borne library able to mimic the natural receptor for LT-B.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

A great debt of gratitude is owed to Dr. Robert Aitken for granting me the opportunity to undertake this project and for his assistance and encouragement over the past three years.

Thanks are also due to Dr. Harry Birkbeck for his support, not only during the course of this project, but in so many ways over the twenty years that I have known him.

Thanks must also be extended to Dr George Smith of the University of Missouri, Dr. Alfredo Nicosia of the Instituto di Ricerche di Biologia Molecolare P. Angelletti, Rome and to Dr. Russell Thompson and Mr. Jim Scott of the University of Glasgow, all of whom supplied me with phage display libraries and associated protocols.

I must also acknowledge the help and friendship of the staff of the Division of Infection and Immunity, particularly Janice Gilchrist, Ian Watt, Calum McCafferty and Allison Bertram.

On the home front, the love, patience and understanding of my wife Aileen and my children Lorna, Andrew and Emma, and the support of my mother and sisters has been in many ways the most important element of all.

This work was funded by the Wellcome Trust, Ecotoxicology grant number 041716/Z/94/Z/TMC/JC.

LIST OF FIGURES

LIST OF TABLES

PAGE No.

LIST OF ABBREVIATIONS

A	Adenine
A_{xxx}	Absorbance at xxx nanometres
B.	*Bordetella*
BAC	biotinamido-caproic acid
BBM	Brush border membrane
BIDS	Bath information and data services
C	Cytosine
cm	Centimetre
CT	Cholera toxin
Da	Daltons
°C	Degrees Celsius
Dsb	Disulphide binding protein
E. coli	*Escherichia coli*
EGF	Epidermal growth factor
ELISA	Enzyme-linked immunosorbent assay
eps	Extracellular protein secretion
ETEC	Enterotoxigenic *Escherichia coli*
F-pilus	Fertility pilus
g	Gravitational force
g	Gram
G	Guanine
Gal	Galactose
GalNAc	*N* - Acetylglucosamine
GDP	Guanosine diphosphate
Glc	Glucose
GMP	Guanosine monophosphate

gp	glycoprotein
GSP	General secretory pathway
GTP	Guanosine triphosphate
HABA	avidin-coupled 4'-hydroxyazobenzene-2-carboxylic acid
h	Hours
HRP	Horseradish peroxidase
IPTG	iso-propyl-thiogalactopyranoside
kb	kilobase
LB	Luria - Bertani
LBHS	High salt Luria - Bertani
LT	Heat-labile toxin
LT-B	B subunit of heat-labile toxin
LT-B^{biotin}	Biotinylated B subunit of heat-labile toxin
M	Molar
MAb	Monoclonal antibody
mg	Milligram
ml	Millilitre
ml^{-1}	per millilitre
Mol.	Molecular
mm	Millimetre
mM	Millimolar
M_r	Relative molecular mass
mRNA	Messenger ribonucleic acid
NANA	sialic acid (*N* - acetylneuraminic acid)
NHS	N-hydroxysuccinimide
nm	Nanometre
nmoles	Nanomoles
NNB	Codon bias involving C, G or T as third base

NNK	Codon bias involving G or T as third base
NNS	Codon bias involving C or G as third base
Oligo.	Oligonucleotide
OPD	ortho-phenylene diamine
ori	Origin of replication
ORT	Oral rehydration therapy
p	Protein
PBS	Phosphate buffered saline
PEG	Polyethylene glycol
RF	Replicative form
sp.	Species
ss	Single stranded
ST	Heat-stable toxin
T	Thymine
TBS	Tris buffered saline
tol	Tolerance of infection gene
T.U.	Transducing unit
µg	Microgram
µm	Micrometre
µM	Micromolar
UV	Ultra-violet
v / v	Volume per volume
V.	*Vibrio*
*V*60	Modified *Vibrio* species 60
WHO	World Health Organisation
wt	Wild type
w / v	Weight per volume
-ve	Negative
+ve	Positive

INTRODUCTION

PART 1 : PHAGE DISPLAY TECHNOLOGY

Methods for identification of ligands

The ability to determine the relationship between the amino acid sequence of an active protein and the functionality of that protein in mature configuration has long been recognised as being critical to the study of protein - ligand interactions.

For example, the quest for a deeper understanding of the nature of individual antibodies, hormones or bacterially-derived toxins requires knowledge of whether their natural ligands are peptides in which the critical binding portion (epitope) is a linear amino acid sequence (*ie.* continuous) or is a structural feature created by separated amino acids being brought together by convoluted folding of the protein (*ie.* discontinuous). In an attempt to ascertain this information, procedures such as assaying a bank of individual variants (*eg.* mutants) for the desired property, followed by sequencing of the variant, have proved useful. However, the practicality of this method is compromised by the vast number of variants that must be considered. Of more practical value are selective methods, in which a collection of variants are simultaneously exposed to the protein, and then treated in such a way as to eliminate those which are reacting only weakly and to isolate those that are reacting in a specific manner. The major difficulty that exists with this approach is the requirement to then identify conclusively the selected species. In recent years, however, this problem has been overcome by the development of selection systems in which the protein being tested is physically linked to the carrier of the DNA that encodes it. After selection for the protein, the DNA can be sequenced, thus giving valuable insight into the nature of the reactive protein, and, by implication, of the site that selected it (Clackson and Wells, 1994; Stephen *et al.*, 1995).

General principles of phage display

In 1988, Parmley and Smith introduced the idea of directly identifying receptor ligands by expressing them in bacteriophage vectors. Their work was inspired by the

variation of Geysen *et al.* (1986) to the traditional' method of screening for epitopes - digestion of the antigen with a range of specific proteases, the individual testing of each fragment for reactivity, and the piecing together of the sequences of the overlapping reactive fragments to reveal the peptide structure. In their method, Geysen *et al.* used random chemical synthesis to extend peptides which had bound monoclonal antibodies weakly, in order to generate peptides that were bound with higher affinity by the monoclonal antibody under investigation. This was done without knowledge of either the sequence or the conformation of the epitope. As well as simply duplicating the natural ligand, unrelated peptides ('mimotopes') were also discovered that were mimics of the discontinuous antigenic epitopes.

In the phage display technique that will be of concern in this project, such peptide mimics are generated not chemically but genetically, by the insertion of foreign DNA into those bacteriophage genes which encode external features of the phage capsid. On expression therefore, foreign proteins or peptides will be 'co-displayed'; fused within the capsid protein on the outer surface of the virion. Those which mimic important epitopes will therefore be antibody (and otherwise) selectable. If the displayed peptide is a specific ligand for a receptor, antibody, or other such molecule, that natural molecule can be used to capture the peptide and the phage displaying it. The captured phage can be used to reinfect a suitable host facilitating the cloning of sequences encoding peptides with interesting properties. The physical linkage of the peptide to the gene which encodes it represents one great advantage that phage display has over chemical synthesis techniques.

Ff class filamentous phage

The bacteriophage chosen as a display vector in phage display must remain infective despite the insertion of a foreign peptide. Equally, it must not kill the infected cell but rather should permit it to keep dividing thereby producing a high yield of phage.

Most of the phage display reported thus far has been carried out in the Ff (F-pilus specific) class of filamentous phage (reviewed by Smith, 1988b; Scott, 1992;

Makowski, 1993; Webster, 1996; Rodi and Makowski, 1999). This class comprises the f1, fd and M13 strains, which demonstrate 98% sequence homology, with few of the sequence differences resulting in the coding of alternative amino acids; those which are different are thought unlikely to affect biological function.

The wild-type phage has a tubular capsid, 7nm wide and 900 to 2000 nm long, that encloses a single-stranded DNA genome of 6407 or 6408 bases (strain dependent) encoding eleven proteins (pI - pXI). A list of the genes and gene products of Ff phage is given in Table 1 whilst a representation of the structure is shown in Figure 1.

The capsid primarily comprises 2700 molecules of the major coat protein pVIII (it constitutes 88% of the weight of the phage particle) arranged in a helical array. Translated with a 23 amino acid signal peptide, pVIII is cleaved to yield a mature molecule 50 amino acids in length, with a molecular weight of approximately 5235 Daltons (Da), that consists of three major domains. During phage assembly when the pVIII molecules assemble around the DNA strand, the first 10 - 13 residues of the carboxyl-terminal domain, four of which are lysine, interact with the sugar phosphate backbone of the chromosome. The hydrophobic central domain separates this region from the acidic amino-terminal domain which is exposed to the outer environment. Felici *et al.* (1991) report the amino-terminal sequence of the protein (for single letter coding see Appendix 5, page 188) to be :

A E G D D P A K A . . . .

All insertions of foreign protein into pVIII, thus far, have been into the $\mathbf{A}^1$ to $\mathbf{D}^5$ amino-terminal sequence that constitutes a mobile surface segment that is distinct from the alpha-helices that make up the rest of the protein (the full base and amino acid sequences of gene VIII are shown as Appendix 8, page 192). Site-directed mutagenesis (Felici *et al.*, 1991; Iannolo *et al.*, 1995) has shown that these residues play no role in either protein translocation or phage assembly.

At the 'proximal' end of the capsid (i.e. that end of the virus which is extruded last from the host during phage assembly) are clusters of five molecules of each of two minor capsid proteins pIII and pVI. The specificity of the Ff class of phage, all of

GENE	AMINO ACIDS *	MOL. WEIGHT *	FUNCTION
I	348	39502	Assembly
II	410	46137	DNA replication
III	406	42522	Capsid protein
IV	405	43476	Assembly
V	87	9682	Binding ssDNA
VI	112	12342	Capsid protein
VII	33	3599	Capsid protein
VIII	50	5235	Capsid protein
IX	32	3650	Capsid protein
X	111	12672	DNA replication
XI	108	12424	Assembly

* Given for mature protein excluding, where applicable, the signal peptide.

TABLE 1 : Size, weight and function of gene products of the Ff class of filamentous phage.

Adapted from Webster (1996).

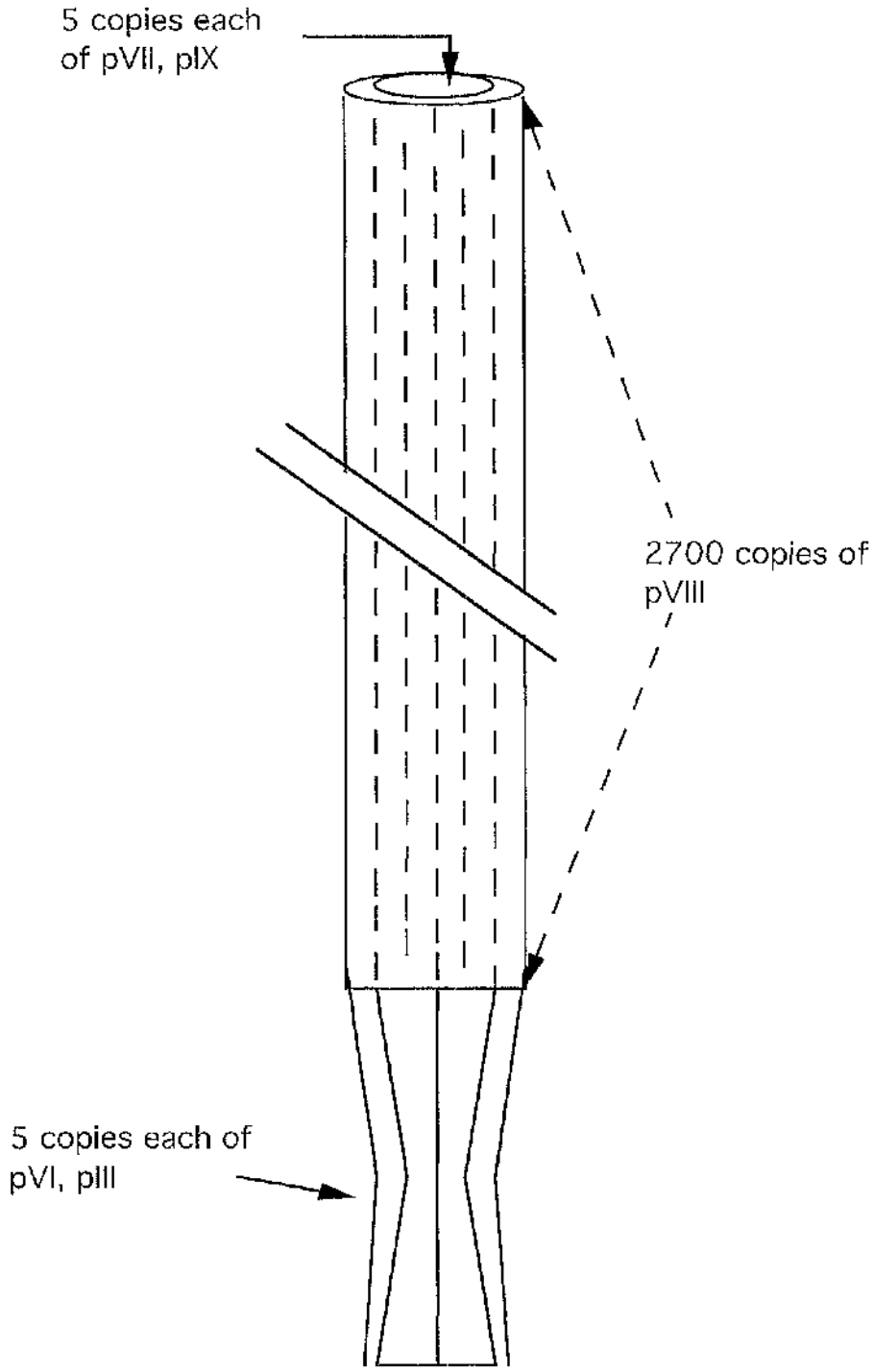


FIGURE 1: Structure of filamentous bacteriophage.

Adapted from Kishchenko *et al.*, 1994.

which infect *Escherichia coli* cells displaying F-pili, is a function of pIII. The largest protein in the virion, it is responsible for binding the F-pilus and thus initiating infection of *E. coli* hosts. The requirement for an F^+ host has been shown to be purely a function of the infection process, as phage DNA artificially introduced to the host by other routes was just as capable of establishing infection. Despite its vital role, pIII is nonetheless the most commonly used capsid protein for phage display, being able to tolerate large insertions with no loss of function. Like pVIII, it is translated with a signal peptide, in this case of 18 amino acids. The 406 amino acid mature protein is divided into three domains. The carboxyl-terminal domain acts as an anchor, holding the phage to the host membrane during virion assembly (Boeke and Model, 1982). It is thought that pIII is attached to the capsid *via* its C terminus and *via* pVI, which acts as a bridge, whilst its N terminus protrudes outwards. Proteins III and VI also serve to stabilise the phage against dissociation, remaining associated despite disruption of the phage by detergent (Hill and Stockley, 1996; Webster, 1996, Holliger and Riechmann, 1997). In 1995, Jespers *et al.* reported the first, and so far only, demonstration of a phage display fusion to pVI.

Five copies of each of two other minor capsid proteins pVII and pIX, which have roles in the initiation of phage assembly, form a highly hydrophobic plug at the 'distal' end.

Two factors make these phage particularly useful as cloning vectors: (a) Their great resistance to extremes of heat (70°C) and pH (2.0). This is exploited in the use of acid elution buffer to break the non-covalent bonds between peptide and ligate thus allowing recovery, without loss of function, of captured phage after affinity purification. (b) The fact that the length of the capsid is simply a function of the length of the DNA it contains. As there is no defined size, the genome can be lengthened by insertion in the knowledge that the helical array of pVIII will simply extend to cover this extra DNA, with an extra copy of pVIII being added for every additional 2.3 nucleotides.

Insertions into both pIII and pVIII are made at, or close to, the amino terminus of the mature protein. This terminus tends to be the most exposed part, thus increasing access for agents intended to select particular inserted sequences. This physical separation of the insert from the capsid also lessens the chances of cross-interference.

Infection cycle of Ff class of phage

Infection of *E. coli* by Ff phage (reviewed by Smith, 1988b; Webster, 1996; Model and Russel, 1988; Reichmann and Holliger, 1997; Rodi and Makowski, 1999) is a process which has only recently become defined with any clarity.

Infection is initiated by the interaction of the extreme amino terminal portion of pIII with the bacterial cell surface protein *TolA*, one of a group of membrane-spanning proteins (the products of the host *tolQRA* genes) that are also involved in the translocation of colicins; in mutants lacking them, phage and colicins are still able to bind to the host, but are unable to infect (the mutant host is therefore *tol*erant of the infecting agents). This interaction releases the next domain upstream of the terminus to bind to the F-pilus of the bacterium. The phage then commences disassembly as it is being drawn towards the host cell surface by retraction of the pilus. Whether this retraction is promoted by the attachment of the phage, or is the result of the regular cycle of pilus polymerisation and depolymerisation, is unknown. Translocation of the DNA genome through the membrane and into the cell is mediated by TolQ, Tol R and Tol A. Once in the cell, replication takes place thus :

Firstly, the host RNA polymerase and DNA synthesis apparatus combine to convert the single-stranded viral DNA to a supercoiled, double-stranded replicating form (RF) in a process that is independent of any virally-encoded proteins. This RF viral DNA will serve as a template for production of mRNAs for all viral proteins.

Then, nicking of the plus strand of the RF molecule by pII initiates rolling circle replication which results in the production of many circularised daughter RFs. Phage protein pII, and possibly the associated pX, in conjunction with host enzymes, mediate this stage of replication. The increasing number of RF's leads to a rapid

increase in production of all viral proteins. Those which aid DNA production are located in the cytoplasm, whilst those involved in phage assembly are integrated into the cell membrane.

The final stage occurs when the cytoplasmic concentration of phage protein pV reaches a critical level. Newly synthesised single stranded viral DNA is not replicated to form RFs, but rather individually associates with pV molecules to form virion-like precursors of the progeny phage.

The process by which these progeny are assembled and released, though still not completely understood, is seen as unique to filamentous phage in that assembly occurs not in the cytoplasm but rather at a specific site formed by pI, pIV and pXI at a point where the inner membrane is in close contact with the outer membrane. The DNA is passed through the membranes, distal end first, releasing the pV molecules which are recycled. The only capsid structural proteins with signal peptides, procoat pIII and pVIII, which, like pVI, pVII and pIX, exist as integral membrane proteins, are at this point cleaved by signal peptidase and all these mature proteins associate with the DNA. Phage assembly slows the growth rate of the host cell by about 50% while it proceeds. Typically, each first generation host cell will produce around 1000 phage, whilst each cell in subsequent generations will yield 100 - 200 particles.

First uses of filamentous fusion phage

Smith (1985) was the first investigator to report production of a fusion based upon the coat protein of a filamentous phage. Using wild-type f1 phage he inserted small fragments of DNA from a plasmid encoding the restriction endonuclease *Eco*RI into a *Bam* HI site in gene III to create a fusion protein with foreign amino acids between the two domains of pIII. The phage retained their infectivity and as *Eco*RI endonuclease, the expressed peptide, was accessible, they were also antibody selectable. The major problems with this system were that the phage could only be maintained by continual cycles of infection and that infectivity was severely depressed, with plaque forming (*ie.* infectious) units representing less than 1 in 50 of

the total virion content (as ascertained by spectrophotometry) compared with 1 in 2 for wild-type. Thus, possession of a fusion protein was clearly disadvantageous to the phage.

Parmley and Smith (1988) developed from this initial work many of the principal elements of phage display. They used a new vector, phage fd-tet, into which they inserted foreign protein near to the signal peptide at the amino-terminal of pIII. This is the exposed terminus of pIII (responsible for binding to the F-pilus), thus the resulting fusion protein was more accessible. They also found that this fusion had a less deleterious effect on phage function. A more detailed account of the development of fd-tet, and its uses in library work is given on page 51.

In assaying for the fusion, they developed the technique of 'biopanning' (described later) and proposed the construction of gene libraries by phage display of random oligonucleotides.

Classification of phage display vectors

Smith (1993) outlines six possible systems for displaying foreign proteins on the surface of filamentous phage. The type 3 vector has a single recombinant gene III with a foreign DNA insert. On expression, all five pIII molecules will carry the foreign insert. However a major problem with large insertions in pIII is that they can promote periplasmic degradation of the fusion protein and can interfere with its normal function of adsorption to the F-pilus, thus rendering the phage non-infective. This problem is overcome in the type 33 vector in which both a recombinant, and a wild-type gene III are present, resulting in a capsid displaying a mixture of pIII molecules, only some of which have foreign proteins fused. The presence of wild-type infective pIII on the capsid surface compensates for the presence of any non-infective fusion pIII. The recombinant gene is inducible, and careful control of expression level can produce monovalent display of pIII; McConnell *et al.* (1994) reported an average of 0.74 recombinant proteins per virion in their work. Whilst this does mean that a subset of virions will carry no recombinant protein, these wild-type

capsids will not feature in the selection process. The type 3 + 3 vector also has recombinant and wild-type genes present, however in this case the recombinant DNA is carried by a phagemid - a plasmid bearing antibiotic resistance markers with origins of replication both for the plasmid itself and for filamentous phage. Infection of phagemid-carrying host cells with a helper phage that carries all the genes to form a capsid but is defective in replication, results in both the helper genome and that of the phagemid being separately packaged into virions. The excess of wild-type pIII supplied by the helper means that around 10% of resulting virions will carry a fusion protein (Clackson and Wells, 1994).

Types 8, 88 and 8 + 8 are the gene VIII equivalents of these gene III types. It has been shown by Greenwood *et al.* (1991) that insertions greater in length than six amino acids, the typical size of a continuous peptide epitope, fail to produce viable phage using type 8 vectors. This is possibly due to disruption of the necessary interaction between pVIII and pVII and IX at the initiation stage of phage assembly (Makowski, 1993). If types 88 or 8+8 are used, each of the fusion proteins is probably locally monovalent, i.e. surrounded by wild-type proteins, as this has been shown to be necessary for capsid stabilisation.

An advantage in using pVIII for insertions, proposed by Folgori *et al.* (1994), is that 10 to 100 pVIII will display the protein on a single phage capsid as compared with a maximum of 5 copies when pIII is used.

The epitope library

As suggested by Parmley and Smith (1988), phage display has been widely used for the generation of libraries of random oligopeptides. A library can be created by synthesising degenerate oligonucleotides using random mixtures of nucleotides and cloning them into the appropriate coat protein genes of the phage or phagemid vector (Scott, 1992). Library size is probably determined by the efficiency with which the ligated mixture is transformed and by the length and composition of the inserted peptide (Folgori *et al.*, 1994), but will typically contain 10^7 to 10^8 clones that each

display a unique peptide sequence. The size of library required to contain every possible combination of amino acids can be calculated as 20^n, 20 being the number of amino acids and n being the length of the insert. A table showing the number of theoretical combinations for various insert sizes is shown as Table 2. It can be seen from this that a library of inserts of up to six amino acids could be 'complete' i.e. could contain every possible combination of amino acids, if 10^8 variants were produced. Although libraries of longer peptide insertions will be incomplete, each long peptide will contain within it several short epitopes. Thus, a library of long peptides may well contain more epitopes than an equivalently sized library of shorter peptides. For example, Kay *et al.* (1993) report that a 36 amino acid insert will, because of a sliding reading frame, contain 31 different hexamers (the formula $(X - N) + 1$, can be used to calculate how many N-mers an X-mer insert will contain). Longer inserts are also advantageous in that they may assume tertiary structure independent of the protein to which they are fused, thus providing multiple contact sites which may be required for ligand binding. However, it has been suggested that, even when enough variants have been generated to cover every theoretical combination, there will nonetheless remain some peptides which cannot feature in the library. Any inserts which exhibit toxicity towards the host, prevent the proper folding of host or phage proteins or which cannot be translated, secreted or processed by the host will necessarily be absent; for example, peptides that include codons not regularly used by *E. coli* have been seen to be under-represented in pVIII libraries produced using that host, possibly because the appropriate tRNAs are not present in sufficient numbers (Clackson and Wells, 1994; Rodi and Makowski, 1999).

Scott and Smith (1990) suggest that peptides encoded by DNA sequences that represent restriction sites of the endonucleases used in construction of the library will be missing. Any peptide that replicates the cutting sequence for the signal peptidase that releases the mature pIII or pVIII may also be eliminated. Moreover, it has been recognised that, as at least the first two amino acids of the mature protein exert some influence over signal peptidase action, then libraries where the insert commences at the extreme tip of the

Size of insert (amino acids)	Number of combinations
4	1.60×10^5
5	3.20×10^6
6	6.40×10^7
7	1.28×10^9
8	2.56×10^{10}
9	5.12×10^{11}
10	1.02×10^{13}
15	3.28×10^{19}

TABLE 2 : Number of different amino acid combinations in complete phage display libraries with various lengths of insert.

mature protein may also be subject to considerable bias in regard to the first few residues (Wilson and Finlay, 1998; Rodi and Makowski, 1999).

Selection of ligate-binding fusion phage by biopanning

'Biopanning' is the term coined by Parmley and Smith (1988) for the most widely used and adapted method for selection of ligate-binding phage (see Figure 2). In the original method, a dilution of biotinylated monoclonal antibody (MAb), which had previously been exposed to UV-irradiated phage (to block non-specific phage binding sites) was reacted with the library. This mixture was then applied onto a streptavidin coated petri dish where phage carrying fusion proteins which specifically reacted with the MAb were immobilised by the biotin-streptavidin bridge. Unreacted phage were removed by extensive washing of the plate, after which the bound phage were eluted with acidic buffer. These captured phage were amplified by infection of a fresh culture of the bacterial host and then subjected to a second round of panning, i.e. were exposed to a fresh surface of bound MAb. Only a small percentage of the phage in a library carrying random fusions would recognise and bind to the MAb. However, after eluting these few phage and amplifying them to high titre, a much higher number would be expected to bind on re-exposure. The ability, in this way, to enrich peptides with sequences of interest, represents the single greatest advantage that this technique has over manual synthesis of peptides. Theoretically, successive rounds of selective enrichment and amplification could be carried out indefinitely, but there is a danger that only the fastest growing phage are favoured by too much repetition. Scott and Smith (1990) report that a single round of panning resulted in a 10^5 enrichment in their application whilst Kay *et al.* (1993) used an excess of phage in a second round to increase competition and ensure selection of only those phage that bound with highest affinity. Individual clones can be isolated and the gene encoding the associated epitope sequenced.

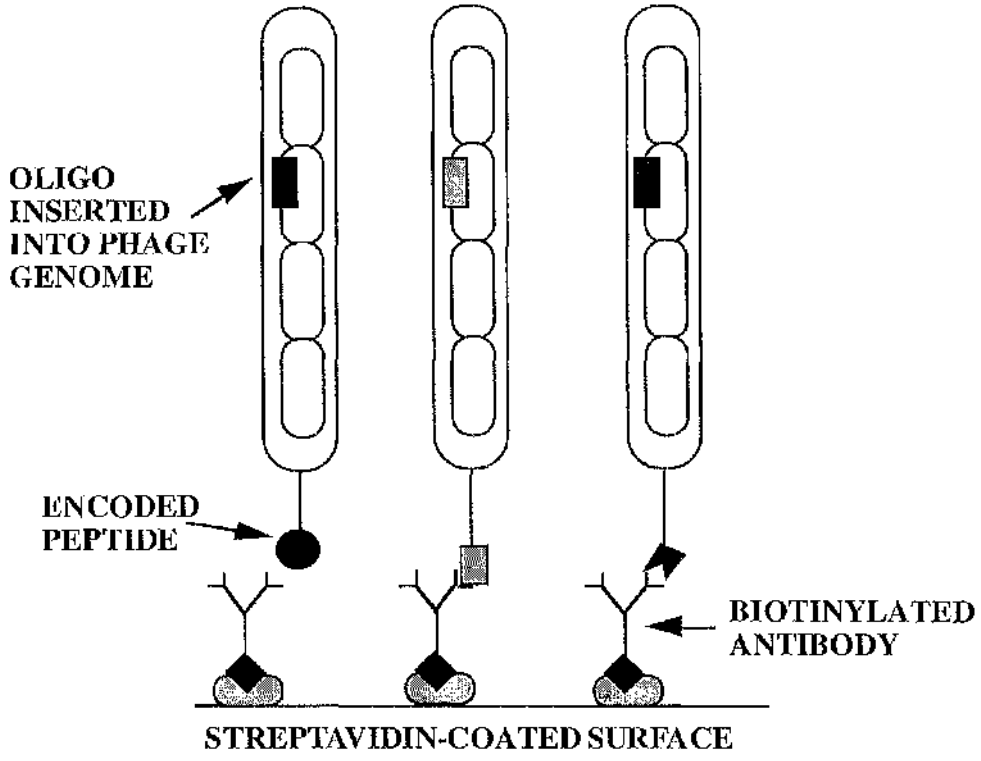


FIGURE 2 : Principle of biopanning.

Biotinylated antibody, captured by plate-bound streptavidin, can capture phage which display peptides compatible with the antibody. After washing to remove unbound phage, the captured material can be eluted off with acid. (After Parmley and Smith, 1988)

PART 2 : BACTERIAL PROTEIN TOXINS AND DIARRHOEAL DISEASE

The aim of this project was to assess whether phage display could be conveniently applied to the identification of receptors for bacterial protein toxins.

The diversity of bacterial protein toxins

Toxin is a difficult term to define comprehensively, but it is the enduring word for those important virulence factors of bacteria, the 'poisons' with which they damage their hosts. Toxins have been described as the 'smart weapons in the biological warfare waged between species, being precisely targeted against specific molecules' (Read and Stein, 1993).

The existence of toxins was first suggested, on purely speculative grounds, by Klebs in 1882 who gave them the name 'sepsins'. Koch, after isolating the causative organism of cholera in 1884, suggested that the diarrhoeal effect may be the result of an extracellular product of the bacterium. Finally, in 1886, Roux and Yersin precipitated pure diphtheria toxin, which they took to be 'a kind of enzyme' and for which the term toxin was first coined. Soon after this came the discoveries of tetanus toxin (1890) and botulinum toxin (1896). Toxin discovery continued through succeeding years; *e.g.* the dreadful consequences of gas gangrene in the First World War prompted investigation into the clostridial toxins. By the start of the 1950's, some 60 bacterial toxins had been identified. (van Heyningen, 1970; Stephen and Pietrowski, 1981; Alouf, 1994). A particularly important advance in the field of toxinology at this time came when Macfarlane and Knight (1941) showed that the α-toxin of *Clostridium welchii* (now *Clostridium perfringens*) type A had lecithinase (phospholipase C) activity which was specifically inhibited by *Clostridium welchii* type A antitoxic serum. For the first time a bacterial toxin in pure form was shown to be an enzyme, thus confirming Roux and Yersin's proposal of 55 years earlier.

Since the 1950's, knowledge in the field of toxinology has continued to increase. The most recent major advance in the field has been the elucidation of the 3-dimensional structures of various toxins, starting, in 1986, with *Pseudomonas aeruginosa* Exotoxin A. As to the present state of affairs, Alouf (1994) reported the following statistics in regard to protein and peptide toxins.

TOTAL No. OF TOXINS KNOWN 290

47% from gram +ve bacteria 53% from gram -ve bacteria

Of these, 75% are extracellular, 25% are intracellular.

At this stage, 6 had been characterised by crystallography.

Satisfactory classification of these toxins is rendered difficult by the lack of a single defining characteristic. They may be grouped according to the originating species of bacterium, the toxin's structure or its mode of action. *E. coli* strains alone produce α haemolysins, enterohaemolysins, verotoxins, cytotoxins and enterotoxins (Gyles, 1992). These include some toxins with the structural designation $A:B_5$ (explained later), a feature of those produced by, amongst others, *Bordetella pertussis* or *V. cholerae* strains. There are in excess of ten different species that produce toxins with ADP-ribosylating activity. Intoxication by a bacterial protein toxin may lead to effects which are cytotoxic, necrotic, haemolytic, neurotoxic, which may cause oedema of tissue, may affect protein synthesis or may have enteric effects. The toxins will act either at the plasma membrane of the host, from whence they affect trans-membrane signalling and /or membrane permeability, or within the cell, where they exhibit enzymatic effects on cytosolic functions of the cell (Montecucco *et al.*, 1994). Toxin genes may be chromosomal, phage-borne, plasmid-borne, or transposon-borne (Alouf, 1994). Lethality ranges from negligible to, in the case of botulinum D, an estimated 3,000,000 times more potent than strychnine (van Heyningen, 1970).

Host receptors for bacterial protein toxins.

An important step towards elucidating the mechanism of action of a toxin is to discover the host receptor to which it preferentially binds. As the concentrations of toxin at the binding site may be very low, and as that site may be in an environment that legislates against prolonged contact, such as the gut with its flowing content, it would therefore be most favourable to the toxin to be able to rapidly bind with high affinity (Gill, 1982). Given that it is unlikely that host cells possess receptors exclusively for the binding of toxins that will ultimately kill them (so-called 'suicide receptors'), it is therefore more likely that the toxins must bind to existing receptors normally utilised in host cellular processes. If these receptors trigger uptake by the cell of the binding protein, the entry of the toxin to the cell is thereby facilitated. Ledley *et al.* (1976) presented evidence that the B subunit of cholera toxin shares sequence similarity with some glycoprotein hormones, the suggestion being that this represents shared recognition mechanisms for host cell receptors. In addition, the process of binding to the receptor may actually trigger uptake into the cell. Among known receptors, gangliosides (ceramide-linked oligosaccharides, discussed in detail on page 35) feature strongly, as do proteins and glycoproteins. Diphtheria toxin has been positively identified as binding to a heparin-binding EGF-like growth factor precursor. The diversity of receptor types and the extent of the gaps in our knowledge are illustrated in a short list of some toxins, and their receptors, shown as Table 3.

The role of toxins in diarrhoeal disease

Enteric pathogens producing one or more toxins are one of the principal aetiologies of diarrhoeal disease. Sears and Kaper (1996) list some 60 clearly identifiable enterotoxins, produced by 21 bacterial species. Toxins have been related to disease by monitoring the effect of ingestion on volunteers and / or the presence of toxin in stool samples, or by identifying intestinal or serum antibody responses to toxins. Another frequently applied technique in enteric toxin research is injection into the exposed

TOXIN	RECEPTOR
Cholera	G_{M1}*
C. perfringens δ	G_{M2}*
Diphtheria	heparin-binding EGF-like growth factor precursor
E. coli Heat-labile (LT) I	G_{M1}*
E. coli Heat-labile (LT) II	G_{D1a}*
E. coli Heat-stable (ST) I	guanylate cyclase type C (GC-C)
E. coli Heat-stable (ST) II	unknown
Pertussis	glycoprotein receptor
Shiga	Gb3*
Shiga-like	Gb3 or Gb4 (strain dependent)*
Pseudomonas exotoxin A	α_2MR / LRP [1]

1 = α_2-macroglobulin receptor / low density lipoprotein receptor-related protein

TABLE 3 : Diversity of toxin receptors.

Ganglioside receptors (marked *) are discussed later (page 35).

ileal loops of experimental animals; this was, for example, used when De *et al.* demonstrated that net secretion could be induced by cell-free culture supernatant fluids of *V. cholerae*, a finding which led to the discovery of cholera toxin (discussed in greater detail below).

Sears and Kaper also define four distinct diarrhoeal activities.

A strict enterotoxin is defined as one capable of stimulating either (a) net secretion in intestinal models, in the absence of cellular damage, or (b) ion secretion in an Ussing chamber. A cytoskeleton-altering toxin produces net secretion *via* the non-lethal alteration of cell shape, usually as a result of rearrangement of F-actin filaments. A cytotoxin lethally damages the cell or tissue, while a neurotoxin alters intestinal smooth muscle activity. Few toxins, however, can be neatly slotted into any one of these categories as it seems increasingly likely that amongst those acting enterically, there are toxins with the capability of simultaneously subverting more than one cellular mechanism in order to create the intestinal secretion characteristic of disease.

Significance and occurrence of diarrhoeal disease

Despite a century of progress in understanding the causes of diarrhoea, and advances in treatment, it remains responsible for around 25% of infant deaths annually (3.7 - 4.6 million). As well as being potentially fatal, the disease is a major contributory factor to the widespread malnutrition in the 'developing' world, where a 5 year old child may have suffered episodes of diarrhoea for up to 15% of his or her life. The most vulnerable are the youngest, with the frequency of attack declining thereafter if the child reaches the age of two. An incidence of five attacks, per child under two years old, per year was observed in Africa. Even in the developed world, infant diarrhoea is a public health issue; in the U.S.A., an estimated 38 million cases requiring medical attention are seen annually in patients under five years of age, resulting in some 425 deaths (Black, 1993).

Another factor adding to the total annual incidence of diarrhoea world-wide is traveller's diarrhoea. This is defined as an individual from a 'low risk' area suffering a

sudden increase, by two-fold or greater, in the frequency of unformed bowel movements, often with associated abdominal cramps, nausea and malaise, during, or immediately after, a visit of up to one months duration to a 'high risk' area. High risk areas are taken to be the developing countries of Latin America, Africa, the Middle East and Asia, whilst low risk countries are those of Northern Europe, North America and Australasia. Traveller's diarrhoea may strike up to 40% of the 12 million or so who travel annually, although in only around 1% of cases will the symptoms persist longer than a few days and only extremely rarely will it ever lead to death (Anonymous, 1986; Black, 1990; 1993).

History and epidemiology of cholera

Cholera, the most severe diarrhoeal disease, has been known since ancient times. Initially restricted mainly to Asia and India, it gradually spread along trade routes into Europe and thence to the New World in the early years of the 19th Century. Commerce and colonisation in succeeding years resulted in frequent epidemics throughout the world. Manifestations range from *Cholera gravis* (severe dehydration, cramps and ruptured blood vessels which can result in death just 6h after onset) to mild or even asymptomatic infection. The shrivelled appearance and skin discoloration characteristic of the most severe manifestation of the disease meant that cholera was feared as much in the 19th Century as the bubonic plague had been before it. Since 1817 there have been seven pandemics, the most recent of which commenced in 1961 (McNeill, 1976; Black, 1986; Hirst, 1995).

Cholera is contracted primarily through ingestion of faecally-contaminated water, a fact first demonstrated in 1854 by Dr. John Snow who, by preventing the use of the Broad Street water pump in London, arrested the spread of the disease in the locality (McNeill, 1976). Cholera remains endemic to this day in areas where water supplies are susceptible to contamination such as coastal areas of the Indian sub-continent. Although practically unknown in the developed world for most of this century, a recent surge of isolated outbreaks have been reported in 17 Mediterranean and

Eastern European countries (Anonymous, 1994b). In South America a major epidemic commenced in 1991. At its peak in 1992, this resulted in 354000 cases being notified to the World Health Organisation (WHO) (Anonymous, 1994a) .

Although there is a long-proven link with contaminated drinking water, the use of dirty water for purposes such as bathing, food preparation, utensil cleaning, *etc.* can continue the spread of the disease even amongst those to whom clean drinking water is available. Seafood too is a notorious vector for this disease. The WHO, in 1995, issued the following guidelines for travellers to cholera-affected areas (Anonymous, 1995a).

- Drink only water that has been boiled or disinfected with chlorine or iodine.
- Avoid ice, unless you are sure that it is made from safe water.
- Eat food that has been thoroughly cooked and is still hot when served.
- Avoid raw seafood and other raw foods.
- Boil unpasteurised milk before drinking it.
- Ice cream from unreliable sources is frequently contaminated and can cause illness. If in doubt, avoid it.
- Be sure that meals bought from street vendors are thoroughly cooked in your presence and do not contain any uncooked foods.

If contracted, cholera can result in the loss of up to 1.25 litres of fluid diarrhoea per hour and up to a total of 60 litres in the course of a severe intoxication. The stool is low in protein and high in Na^+, K^+, Cl^- and HCO_3^- (Levine *et al.*, 1983).

In the absence of counteractive measures, case fatality rates can exceed 20%. However, cholera is treatable with antibiotics and, more importantly, with oral rehydration therapy (ORT), in which isotonic solutions of glucose and salts are continually fed to the victim, replacing those lost through diarrhoea. This simple remedy is estimated to save the lives of one million infants each year and is thought to be responsible for the relatively low death rate in the South American outbreak of the1990's. In Tanzania in 1977, case fatality rates were reduced from 55% to 1.6% by

the introduction of ORT, whilst WHO figures show a fall in world-wide case fatality rates from 49.3% in 1961 to 1.7% in 1993 (Black, 1986; Hirschhorn and Greenough, 1991; Anonymous, 1994b).

Vibrio cholerae

Snow's demonstration of the source of cholera in 1854 (see above) did not lead to isolation of the causative organism until advances in bacteriological technique made such breakthroughs possible. Robert Koch, the originator of many such technical improvements, turned his attention to cholera in 1883, being appointed leader of an investigation into the disease by a German government alarmed at the prospect of an Egyptian outbreak spreading to mainland Europe. Having arrived in Egypt as the epidemic was receding, he moved on to Calcutta in order to continue his work (Brock, 1988). It was here that he isolated an organism which, he said, was '*not quite straight, as other bacilli are, but slightly curved, comma-like.*' As these organisms were present only in the stool or intestines of cholera victims and not in healthy people, nor those suffering any other illness, he concluded that *Vibrio cholerae* was the causative organism of the disease (Anonymous, 1884).

The epidemic and pandemic outbreaks of cholera seen thus far are due to the strain *V. cholerae* O1 for which man is the only natural host. Two distinct biotypes of O1 have been identified, namely Classical and El Tor (named after the field station where it was first isolated), each of which has Ogawa and Inaba serotypes. The ratio of mild to severe cases with the Classical biotypes is 4 or 5 : 1 whilst that with El Tor is 40 or 50 : 1. The normal infective inoculum is thought to be around 10^2 to 10^3 organisms, and in fully healthy volunteers 10^3 organisms of the less severe El Tor Inaba strain, when administered with bicarbonate to neutralise stomach acid, caused a mild diarrhoeal disease in four out of six subjects. Administration of 10^6 organisms of the same biotype and serotype without bicarbonate failed to infect (Levine *et al.*, 1983; Black, 1986).

A new serogroup of *V. cholerae* which is immunologically unrelated to either the O1 biotypes or to any of the non-O1 serotypes, and designated O139, emerged in India at the end of 1992. Although still centred in Asia, the WHO express the fear that, should it spread, it may cause an excess of deaths amongst adults in those areas where *V. cholerae* O1 is presently endemic, and may reverse the trend of ever-decreasing case fatality rates. A vaccine against this strain is thus one of their major research priorities (Anonymous, 1994b; 1995b; Hirst, 1995).

Escherichia coli diarrhoeal disease

All cases of diarrhoea displaying the classic symptoms of cholera were recorded as such regardless of the absence of *V. cholerae* from stool samples until, in the decade following the Second World War, *E. coli* strains (today called enterotoxigenic *E. coli* or ETEC strains) were seen to be associated with similar diarrhoeal disease (Sack, 1975). De *et al.* (1956) demonstrated that such strains were capable of inducing identical symptoms to cholera in ileal loops of rabbits.

E. coli strains are responsible for around 20% of severe diarrhoeal cases in infants and were found in studies in Bangladesh, Zaire, Brazil and Thailand, to be the most frequent pathogen in samples taken from 2-year-olds receiving clinical treatment for diarrhoea (Hirschhorn and Greenough, 1991; Black, 1993).

The major syndrome attributable to ETEC in numerical if not pathological terms, is the aforementioned traveller's diarrhoea. The cause of the most common manifestation, a watery stool containing no blood or body tissue (i.e. non-dysenteric diarrhoea) of three to four days duration (Gorbach *et al.*, 1986) had remained a matter of speculation until *E. coli* was implicated by Kranendonk (1960). Being part of the normal gut flora, it had, until then, been overlooked as a potential cause of disease. ETEC strains have now been identified, in all countries where epidemiological surveying has been done, as the most common cause of the syndrome. In twenty-seven studies of visitors to Latin America, Asia and Africa, diarrhoea was shown to have affected 50% either during or immediately after their visit. Staying in a higher

standard of accommodation did nothing to lessen the incidence of disease. ETEC strains are thought to be responsible for at least 50% of these cases. Aside from drinking water, salads, uncooked vegetables and ice cubes are the most frequent sources of contamination. *E. coli*-mediated diarrhoea ranges in severity from cholera-like to mild. In terms of case numbers it is far more significant than cholera, afflicting 650 million people annually and resulting in 800000 deaths. (Black, 1986; 1990; Kean, 1986; Spangler, 1992; Hirst, 1995).

Role of toxins in cholera-like diarrhocal disease

As already stated, Robert Koch had proposed that the symptoms of cholera were the result of a toxin released by the causative organism when he first related *V. cholerae* to diarrhoeal disease in 1884. This theory was not proved until 1959 when Dutta *et al.*, using oral administration to infant rabbits previously subjected to gastric lavage, and De, using rabbit ileal loops, independently demonstrated toxic activity by cell-free culture supernates. The enzymically-active toxin is released into the gut lumen by organisms which attach to the mucosa of the proximal small intestine. No damage to, nor invasion of, the mucosa by the organism accompanies toxin production, the bacteria remaining in the intervillus spaces or crypts until washed out by natural flow. Cholera toxin (CT) was isolated in pure form by Finkelstein and LoSpalluto (1970). As little as 5µg of pure CT (administered with bicarbonate) was shown to be sufficient to cause mild diarrhoea in four out of five adult volunteers in just 6 - 14h while a dose of 25µg resulted in the loss of around 20 litres of fluid (Levine *et al.*, 1983).

In the 1960's *E. coli* was also shown to be a cause of diarrhoea in young animals and Smith and Halls (1967), using the ileal loop method of De (1959), were able to show that cell-free filtrate of a porcine strain was capable of causing loop distension, thus implying the production of a toxin by this bacterium too. A 'heat-labile, non-dialysable, immunogenic enterotoxin' was the first to be identified, from an *E. coli* human strain, by Sack *et al.* (1971), and was named 'heat labile toxin' (LT).

CT or LT intoxication causes massive loss of water and electrolytes from affected cells into the lumen of the small intestine thus producing the characteristic 'rice water' diarrhoea and potentially leading to the death of the host from dehydration.

Types of *E. coli* enterotoxins

Purification of ETEC toxins was to prove a difficult process because of the simultaneous production of both heat-labile and a heat-stable enterotoxin (ST) by *E. coli* strains (Smith and Gyles, 1970), and because, unlike CT, these toxins are not secreted but are retained within the bacterial periplasm. Sack (1975) defined ST as being able to withstand a temperature of 100°C for 15 minutes while LT is destroyed by treatment at 60°C for 30 minutes. Around 35% of strains isolated worldwide cause severe diarrhoea, through the production of both ST and LT. Less severe symptoms are caused by the 30% of strains that are LT^+ / ST^- and by the 35% that are LT^- / ST^+.

As this work will primarily be concerned with LT, ST will be briefly summarised here. There are two types, designated ST-I and ST-II which are related only in that they are both heat-stable, plasmid-encoded, low relative mass peptides capable of causing diarrhoea. ST-I is a single peptide of 18 or 19 amino acids (depending on host) with a relative molecular mass (M_r) of around 2000 which recognises a host intestinal epithelial receptor, localised in the brush border membrane. After attachment to this guanylate cyclase type C receptor, diarrhoea results from the raising of cellular levels of cGMP, a cellular signal molecule. ST-II is 48 amino acids in length and appears to be produced by porcine ETEC only. The receptor and mechanism are as yet unknown although diarrhoeal effects may be produced by raising levels of cellular bicarbonate. ST types are not antigenically related to either CT or to LT (Acheson, 1992; Gyles, 1992; Sears and Kaper, 1996).

Green *et al.* (1983) identified a second distinct serotype of LT and Pickett *et al.* (1986) proposed the names LT-I for the classical type and LT-II for the new type. Despite sharing a similar structure and mode of enzymic action, LT-II is not

neutralised by antisera against either LT-I or CT, while its active portion shows only around 56% sequence homology. Its binding portion shows no homology with those of CT and LT-I and recognises a different receptor. It is thus probably only distantly related to CT and will not be given further coverage in this report.

Comparison of LT and CT

Since De *et al.* (1956) reported that ETEC strains were capable of inducing cholera-like symptoms, substantial similarities between LT-I and CT have been demonstrated. Both porcine LT-I (LTp-I) and human LT-I (LTh-I) were shown to be antigenically related to cholera toxin (CT) (Gyles and Barnum, 1969; Smith and Sack, 1973) and to each other. Although not identical, LTh-I and LTp-I show 95% (Honda *et al.*, 1981) or 96% (Yamamoto and Yokota, 1983) sequence similarity and for the purposes of this report will be considered, unless otherwise specified, as the single entity 'LT-I'. *E. coli* -mediated diarrhoeal diseases of other animals have been reported; for example, Inoue *et al.* (1993) sequenced a chicken strain (LTc-I); however, they concluded that this was in fact identical to LTh-I.

Clements and Finkelstein (1978a; 1978b) demonstrated immunologically that LT shared identity with both choleragen (the cholera holotoxin) and choleragenoid (the non-enzymic binding portion of cholera toxin) suggesting that LT shared a similar structure. This was confirmed by electrophoretic separation of various cross-linked moieties (Gill *et al.*, 1981). It has also been reported that LT-I and CT catalyse similar reactions, share 79% sequence homology and recognise the same host cell receptor (Holmgren, 1973; Moss and Richardson, 1978; Dallas and Falkow, 1980). Each of these points will be expanded upon later in this report. In addition, both CT and LT are potent immunogens. On contacting mucosal surfaces they can stimulate production of both secretory IgA and systemic IgG and when administered along with another antigen they can act as adjuvants, stimulating the response to that antigen (Dougan, 1994).

These similarities allow the classification of LT and CT as 'cholera-like toxins', however, differences do exist. LT had been shown to be plasmid-encoded by Smith and Halls (1968) whilst Vasil *et al.* (1975), by a series of conjugation experiments, showed that CT is chromosomally-encoded. Location of the toxin is also a point of difference. Whilst both De (1959) and Dutta *et al.* (1959) had demonstrated the presence of CT in culture-free filtrates, it was only when Richardson (1969) showed that cell lysis did not increase its activity that secretion of the toxin into the external environment was confirmed. In contrast, Clements and Finkelstein (1979) had shown that LT-I was not present in supernatant fluid of carefully prepared cultures, suggesting that it was cell-associated. Hirst *et al.* (1984a) proved that it is periplasmically located by demonstrating release of toxin on sphaeroplasting of cells (i.e. rupturing of cell walls with lysozyme whilst maintaining the integrity of the plasma membrane).

Structure of LT-I

The earlier-mentioned difficulty in purifying LT-I made characterisation of the toxin problematical (an account of purification methods is given on page 48). Prior to definitive purification by Clements and Finkelstein (1979), reported M_r's showed variations indicative of the lack of purified material, however, it is now known that LT-I is assembled from heavy and light subunits.

The heavy 'A' subunit and the light 'B' subunit are transcribed from a di-cistronic mRNA, translated, and processed to mature proteins of 240 and 103 amino acids respectively (Dallas and Falkow, 1980). The A subunit has M_r of 28000 whilst the B subunit has M_r of 11400, with LTp-I being larger due to minor amino acid variations (Hofstra and Witholt, 1984; Leong *et al.*, 1985). These relative masses correspond closely with those reported for CT (Lonnroth and Holmgren, 1973). A comparison of subunit M_r of CT, LT-I and LT-II is shown in Table 4.

The A subunit can be proteolytically nicked by tryptic digestion *in vitro* to yield A_1 and A_2 peptides (Mr of 22000 and 5500 respectively) which remain joined by a disulphide bond between a **C** in each of the two fragments.

Toxin	Holotoxin	A subunit	A1 fragment	A2 fragment	B subunit
CT	85000	27000	22000	5500	11600
LT-I	88000	28000	22000	5500	11400
LT-II	83000	28000	21000	7000	11800

TABLE 4 : Relative masses of components of cholera - like toxins

After Gyles, 1992.

Gill (1976) predicted from results obtained on polyacrylamide gels that the CT holotoxin consisted of five B subunits bound in the form of a ring which facilitated binding of the toxin to the host epithelium while the enzymically active A subunits lay in the centre of the ring. He subsequently predicted this to be the case in LT-I also (Gill *et al.*, 1981). X-ray crystallography has since confirmed the predicted conformation for LT-I to be correct with the holotoxin consisting of two subunits which, while functionally independent, are closely bonded. The A_1 portion is triangular and is hinged to the long α-helix that is A_2 by a disulphide bridge near to the carboxyl- terminus. The A_2 α-helix lies along one face of the A_1 triangle and interacts extensively with it. The B subunits form a pentameric ring bonded by six β–sheets at each intersection (three from each monomer) with the central pore of the ring lined with α-helices. The ring has a highly charged 'upper' surface with a flat appearance whilst the 'underside' is convoluted due to clefts between adjacent monomers. The carboxyl-terminal region of A_2 passes through the pore in the centre of the five B subunits, interacting with those subunits and holding the A_1 subunit above the 'flat' surface of the B ring (Sixma *et al.*, 1991; 1993; van Heyningen, 1991). Figure 3 is a schematic diagram of the B subunit pentamer showing the secondary structure features revealed by X-ray crystallography.

Export of LT to the periplasm

Hofstra and Witholt (1984), with a series of pulse chase experiments using (^{35}S) methionine, demonstrated that LT-I originates in the cytoplasm as pre-LT-A (M_r of 29500) and pre-LT-B (M_r of 13500). Pre-LT-A and pre-LT-B were shown to have signal sequences of 18 and 21 amino acids respectively. Such sequences serve to target proteins across the cytoplasmic membrane into the periplasm, although they have no influence on whether or not they remain there (Hirst and Welch, 1988). The discovery of the LT signal sequence confirmed the conclusion that Dallas and Falkow (1980) had drawn from a study of the amino acid sequence of LT. They had presented four lines of evidence that a signal sequence was present: (i) the lengths of the

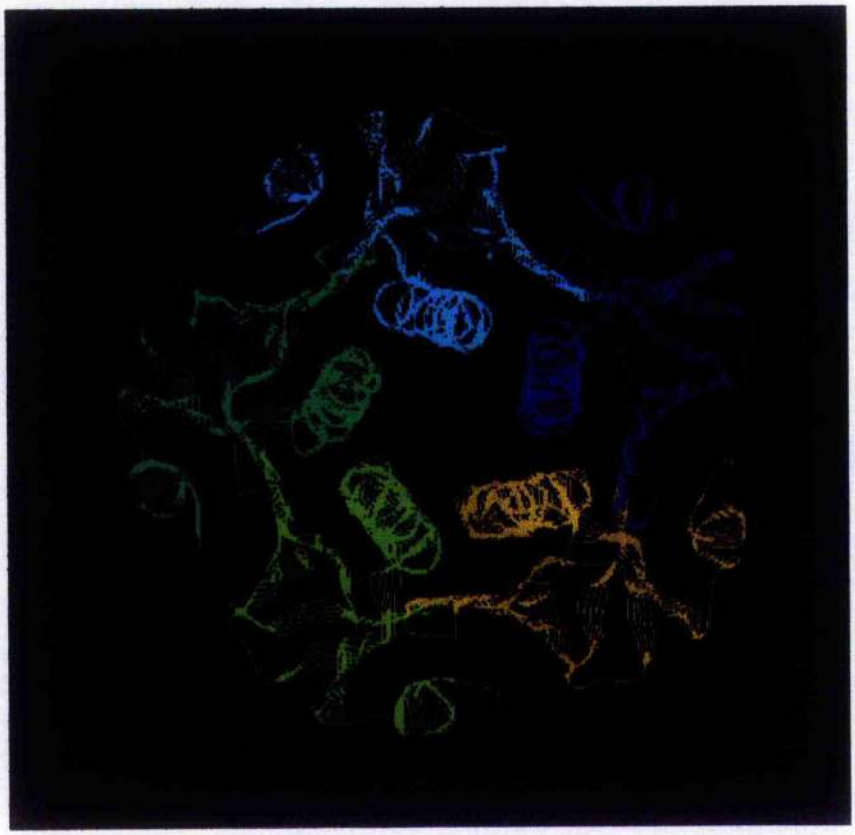


FIGURE 3 : Secondary structure of the B subunit pentamer

Structural features as revealed by X-ray crystallography include the α-helix-lined central pore formed by the five monomers (each differentially coloured). (After Sixma *et al*, 1991)

putative signal sequences were consistent with those previously reported (*ie.* approximately 20 amino acids); (ii) the position of the cleavage point, after a glycine, was also consistent with others; (iii) the sequence resembled known signal sequences, especially in regard to its high hydrophobicity; (iv) LT was thought at that time (subsequently proven, see above) to be a periplasmic protein. Hirst *et al.* (1994) report that LT export is unaffected in mutants lacking either the Sec or GroEL 'chaperones' most frequently associated with export, suggesting that in this case another less well defined system is involved.

Hirst and Holmgren (1987a) have demonstrated that CT holotoxin is secreted from the cell in its complete form. In a further paper, (Hirst and Holmgren, 1987b) they used pulse labelling to confirm that monomeric B is exported into the periplasm, that it associates in around 1 minute to form pentameric B rings and that it is these intact rings which, after a pause, are secreted into the surrounding medium. From these results they concluded that folding of the proteins into their mature conformations must occur in the periplasm. This work thus confirmed that the secretion of CT was necessarily a two-stage process (cytoplasm to periplasm then periplasm to outside) and was thus typical of the general secretory pathway, (GSP) seen in several species of Gram negative bacteria (*eg. Klebsiella, Pseudomonas, Erwinia*) rather than the single-step, signal peptide-independent type secretion typified by *E. coli* α-haemolysin.

Genetic alterations influencing the three or four carboxyl-terminal amino acids profoundly affected the ability of the B subunit protein to assemble and secrete due to the effect this had on the ability to fold (Sandkvist *et al.*, 1990). That at least some of the folding is mediated by disulphide bonding was suggested by the work of Hardy *et al.* (1988) who demonstrated that the addition of dithiothreitol, which reduces disulphide bonds, to growing LT-producing cells arrested formation of B subunits. This finding was confirmed by Yu *et al.* (1992; 1993) who reported that disulphide-bonding protein A (DsbA) was produced by both CT- and LT-producing strains. This protein, first described by Bardwell *et al.*(1991), is an enzyme (M_r of 21000) with

thiol disulfide oxidoreductase activity and is thus a promoter of disulphide bond formation. Hirst *et al.* (1994) have shown that disruption of the *dsbA* gene causes a 100-fold reduction in LT production.

Although single B subunits can spontaneously form pentameric rings independent of A subunit (thus forming the above mentioned choleragenoid), Hardy *et al.* (1988) showed that incorporation of A subunit can elicit a three fold acceleration in oligomerisation of B subunits and that, in solution, mature A subunits cannot associate with complete B pentamers. Orlandi and Fishman (1993), however, have shown that A subunit can associate with B pentamers which are membrane bound to form active holotoxin.

Secretion of cholera-like toxins from the periplasm

Extracellular secretion is relatively rare in *E. coli*, but much less so in *Vibrio* species. *V. cholerae* for example, is known to secrete a number of extracellular proteins including at least three types of protease (Young and Broadbent, 1982) as well as CT itself.

It was shown by Hunt and Hardy (1991) that LT could be released from the producing cell by growth in medium containing physiological concentrations of bile salts, an effect which could be increased by augmentation of the medium with trypsin. The conclusion drawn was that host factors may facilitate the release of LT by non-specifically permeabilising the outer membrane of the cell. Other reported results suggest that a more specific system is also involved. Firstly, Pearson and Mekalanos (1982) created the plasmid pJM17 by cloning the genes for CT (A and B) into plasmid pBR322. *E. coli* K12 strains carrying this plasmid produced CT, however 94% of this normally exported toxin remained cell associated. Following on from this, plasmid-encoded LT holotoxin (Neill *et al.*, 1983) and either the LT-A or LT-B subunit alone (Hirst *et al.*, 1984b) were expressed in *V. cholerae*. In these instances, the normally periplasmic LT was secreted, as was the B subunit; the A subunit alone remained cell associated. This suggests that there is a specific secretory apparatus in

V. cholerae which recognises structural elements of the B pentamer. Further to this, Hirst and Holmgren (1987a) showed that no other periplasmic protein was released along with CT in this system.

Secretion of CT is therefore, most probably, a complex process, involving as it does, translocation of the proteins across both cell membranes. Subunits appear to cross the cytoplasmic membrane as monomers which then fold and form the mature toxin that will be secreted to the outside. The clear implication is that the failure of *E. coli* to secrete LT is primarily due to its lack of a mechanism for passage across the outer membrane, as it can be secreted without any modification by *V. cholerae*.

Overbye Michel *et al.* (1995) used pMMB68 (see page 45) and other plasmids to transfer the *etxB* gene (under the control of an inducible promoter) into a range of such Gram -ve species. On testing for the secretion of LT-B, only members of the *Vibrionaceae* and *Aeromonadaceae* appeared capable of doing so (retention of β-lactamase in the periplasm was also monitored to confirm that the outer membrane was intact). This result is reported as being consistent with an observed similarity in the amino acid sequences of the GSP proteins of *V. cholerae* and *Aeromonas hydrophila*.

The GSP in *V. cholerae* is thought to involve accessory proteins. The gene coding for the first of these to be identified was named *e*xtracellular *p*rotein *s*ecretion (*eps*) E (Sandkvist *et al.*, 1993). Now, twelve such *eps* gene products have been identified and the suggestion is that they form a multi-protein complex that permits specific translocation across the outer membrane. Mutations in 6 *eps* genes, including *epsE*, have been shown to abolish translocation, as demonstrated by a build-up of CT in the periplasm. Further *eps* gene products have been shown to be necessary for export in other species and are likely to be so in *V. cholerae* also (Sandkvist *et al.*, 1995). The exact mechanism involved remains poorly understood, but it has been clearly demonstrated that it is specific to designated proteins; periplasmic proteins such as β-lactamase or alkaline phosphatase are not concomitantly secreted. The location of the *eps* gene products is similarly uncertain: sequence data shows that *epsE* lacks both a

leader sequence (suggestive therefore of a cytoplasmic location) and a long stretch of hydrophobic amino acids, capable of spanning a membrane (Sandkvist *et al.*, 1993). Thus it is probably membrane-associated, although not necessarily permanently so (Overbye Michel *et al.*, 1995).

This issue of secretion of LT-B by *Vibrio* species is further discussed on page 45.

Host receptor for cholera-like toxins

The first breakthrough towards elucidating the nature of the host cell moiety to which the B subunit pentamer of cholera-like toxins bind came from van Heyningen *et al.* (1971) who showed that a mixture of gangliosides produced a blocking effect on CT action. Gangliosides are oligosaccharides linked to hydrophobic ceramide backbones. The ceramide anchors the ganglioside into the cell membrane leaving the oligosaccharide projecting from the cell and thus available for interaction with ligands (Moss and Vaughan, 1991). Holmgren *et al.* (1973) showed that cholera toxin bound with great avidity to the cell surface ganglioside G_{M1} (for structures of gangliosides see Table 5) then further showed that LT did likewise (Holmgren, 1973). It is now established that G_{M1} is the functional binding site for CT (Griffiths and Critchley, 1991) with the CT recognising the terminal GalNAc - Gal sequence.

Holmgren *et al.* (1985) removed the lipid from human brush border membrane (BBM) and whole cell membrane. CT could no longer bind to such a membrane suggesting that it is almost completely specific for G_{M1}, however LT-I binding was unaffected. This suggested that, as well as ganglioside, LT-I was also able to bind to non-lipid (presumed glycoprotein) receptors. Holmgren confirmed this by showing that free CT-B subunit, which blocks the activity of CT, had no effect on the activity of LT-I bound to delipidised membrane, thus indicating that the glycoprotein is an actual receptor rather than simply a non-specific binding site. Free LT-B subunit blocked both CT and LT-I activity, presumably by blocking both ganglioside and glycoprotein receptors. Fukuta *et al.* (1988) demonstrated weak binding by CT and LT-I to G_{D1b}, and by LT-I only to G_{M2} and asialo G_{M1}, by immunostaining of thin

G_{M1}	Gal - GalNAc - Gal - Glc - Ceramide | NANA
asialo-G_{M1}	Gal - GalNAc - Gal - Glc - Ceramide
G_{M2}	GalNAc - Gal - Glc - Ceramide | NANA
G_{D1b}	Gal - GalNAc - Gal - Glc - Ceramide | NANA | NANA

TABLE 5 : Structures of gangliosides.

Abbreviations : Gal - Galactose ; GalNAc - *N* - Acetylglucosamine ; Glc - Glucose ; NANA - sialic acid (*N* - acetylneuraminic acid)

layer chromatograms. As the B subunits of CT and LT-I show 82% sequence similarity there is scope for different receptor recognition. Griffiths and Critchley (1991) used Wistar rat BBM to study binding to CT and LT-I. From this study they concluded that LT-I binding could be blocked by galactose and the galactose-specific lectin, ricin, suggesting that LT-I binds to BBM galactoproteins, these showing only weak affinity for CT. They also suggested that there are some 20 to 30 times more binding sites for LT-I than for CT in rat BBM, with strain and species differences in expression of these moieties. Holmgren *et al.* (1985) reported no difference in numbers of LT-I and CT binding sites in human BBM.

With the holotoxins of CT and LT-I containing five B subunits, binding to receptors is multivalent with five potential binding sites per molecule. The fluid nature of the membrane and the flexible nature of the oligosaccharide extension probably facilitates this (Moss and Vaughan, 1991).

Orientation of bound pentameric rings

One of the most surprising aspects of the structure of LT, as determined by Sixma *et al.* (1991; 1993) was the finding that it was the convoluted underside of the B ring which presented to the membrane of target cells. This orientation places the A_1 subunit on the opposite side of the B pentamer from the host membrane which, given that the pore in the centre of the ring is too small to easily allow A_1 to pass through (as originally predicted by Gill (1976)), poses a question as to how the A_1 subunit can contact the membrane in order to insert. This arrangement was confirmed by Orlandi and Fishman (1993) who demonstrated that membrane bound holotoxin could still bind anti-A_1 subunit specific antibody and that intact holotoxin with anti-A_1 FAb fragments bound to it was not prevented from binding to a membrane. From these findings it was apparent that the A_1 subunit of bound holotoxin must be oriented away from, and not towards the host membrane (see Figure 4).

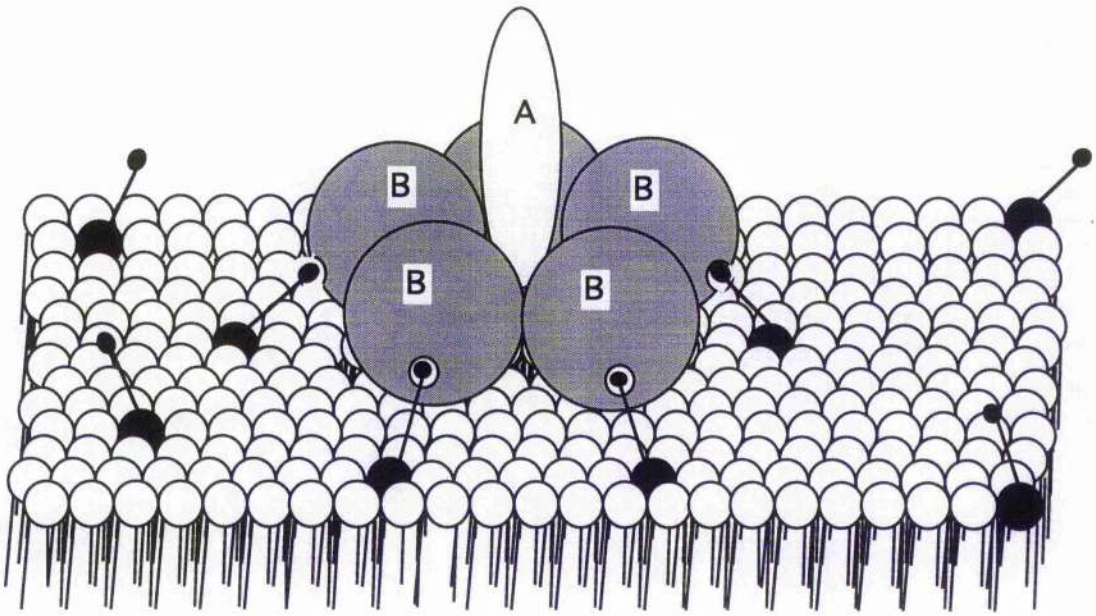


FIGURE 4 : Cholera-like toxin anchored to a membrane.

After Montecucco *et al.*, 1991. Attachment is *via* ganglioside molecules (shown in black). Note that the B subunit intervenes between the A subunit and the membrane.

Binding of cholera-like toxins to G_{M1}

Further crystallography was performed on CT and LT bound to both galactose and G_{M1} (Merritt *et al.*, 1994a; 1994b). The work on G_{M1} binding showed each of the five binding sites to be located almost entirely within a single B subunit, with only a single hydrogen bond to a moiety on an adjacent subunit. A schematic representation of the binding of G_{M1} in the cholera toxin pocket, taken from Merritt *et al.* (1994a), is seen in Figure 5.

Bonding ideally involves the terminal galactose and the sialic acid, with GalNAc and the second galactose playing a minor role and the glucose playing no role at all. It is worth noting that H^{13} of CT is not conserved in LT. Being part of the binding site, it is possible that it accounts for the different range of gangliosides that each can bind to.

Binding is primarily to one subunit, with a single G being the only residue involved from an adjacent subunit. Hence, a toxin molecule can bind five G_{M1} molecules, one per subunit.Although sialic acid is a feature of all but the asialo-gangliosides, LT can bind to galactose, lactose or agarose alone with sufficient avidity to allow these sugars to be used as ligands in affinity chromatography, as discussed on page 48.

Another outcome of these investigations (Merritt *et al.*, 1995) is of great interest. When investigating receptor binding of CT-B, B-subunit pentamers featuring a G33 to D mutation were produced. These pentamers unexpectedly joined together by the insertion of a pentapeptide loop, (consisting of residues 10 to 14) of one, into the GM1-binding site of another. As each pentamer acted as both 'donor' and 'receptor' a lattice of conjoined pentamers was formed. The amino acids involved in this loop (AEYHN), bonded in a manner analogous to the natural receptor, with the imidazole ring of H13 both stacking over the indole ring of W88 and donating a hydrogen bond to E51 in the manner of GM_1 (see Figure 5). Although no equivalent of the sialic acid portion of G_{M1} was seen, by being able to insert and bind in the receptor-binding site of CT-B, this peptide was therefore possibly mimicking G_{M1}.

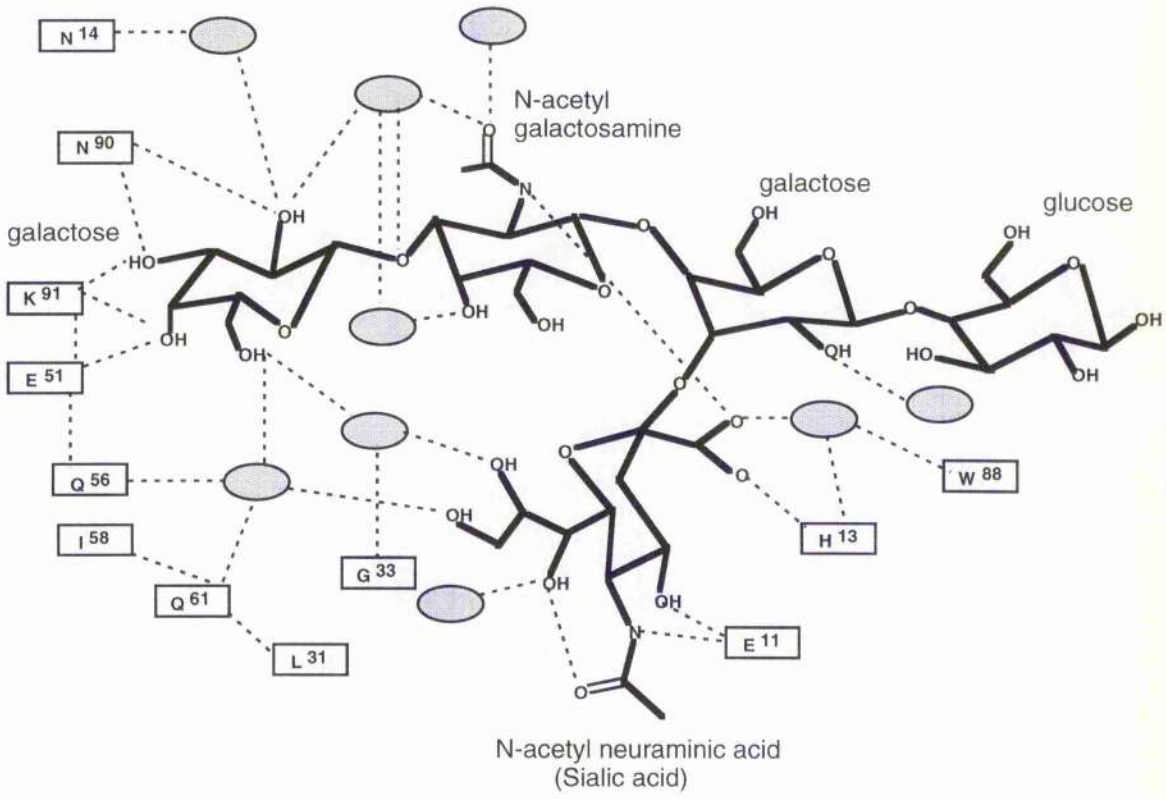


FIGURE 5 : Hydrogen bonding between CT and ganglioside G_{M1}.

Residues in the CT binding pocket are shown in rectangles, solvent molecules as ovals. All of the binding pocket residues belong to B subunit 5, except for G^{33} which is from an adjacent subunit.

Enzymic effect of cholera-like toxins

Investigations into the way in which cholera-like toxins induced diarrhoea led to the finding reported by Field *et al.* (1969) and Greenough *et al.* (1969) that CT had the same effect on rabbit intestine, and could be blocked by the same inhibitors, as chemical agents which caused an increase in the levels of cellular cAMP. Gill *et al.* (1976) showed that an increase in cAMP in pigeon erythrocytes when treated with the A_1 subunit of LT was due to NAD- and ATP-dependent ADP ribosylation of a regulatory protein of the host cell. Moss and Richardson (1978) subsequently demonstrated that A_1 had NAD-glycohydrolase and ADP-ribosyltransferase activity, and thus could catalyse the reactions

$$\text{NAD} \xrightarrow[\textit{NAD glycohydrolase}]{} \text{ADP-ribose} + \text{nicotinamide} + H^+$$

$$\text{NAD} + \text{arginine} \xrightarrow[\textit{ADP ribosyltransferase}]{} \text{ADP-ribose-L-arginine} + \text{nicotinamide} + H^+$$

Structure and function of G proteins

The regulatory protein affected by A_1 of CT and LT is a signal transducing guanine nucleotide-binding (G) protein which would normally function by coupling hormone and neurotransmitter receptors to their specific intracellular effectors. The particular G protein targeted by CT and LT-I for ADP-ribosylation is G_s. Adenylate cyclase activity is modulated by G_s (stimulatory) and G_i (inhibitory) proteins. G_s, in common with other G proteins, is a heterotrimer which consists of three subunits termed α, β and γ. It is the α subunit which stimulates adenylate cyclase, while the β and γ subunits exert an inhibitory effect on its activity. G_s, in common with other G proteins, is activated by the binding of GTP thus :

(1) $\quad G_s + GTP \longrightarrow \{G_s.GTP\}$

(2) $\{G_s.\ GTP\}$ + adenylate cyclase ----------> $\{G_s.GTP.$adenylate cyclase$\}$

ATP---------->cAMP

The binding of GTP results in the dissociation of the α from the β and γ subunits. G_s does not remain in a state of permanent activation, however, as it is intrinsically a GTPase, an enzyme which catalyses the conversion of GTP into GDP, thus promoting reassociation of α, β and γ.

Hormones, on binding to receptors, would normally accelerate the release of GDP from G_s thus creating new binding sites for stimulatory GTP on the α subunit. The intrinsic GTPase activity of the α, and the inhibitory nature of the β and γ subunits would serve to modulate this activity (Neer and Clapham, 1988; Fishman, 1990; Moss and Vaughan, 1991).

Export and activation of cholera-like toxins

Although the nature of the binding of the B subunit to the cell and the enzymic activity of the A subunit are well understood, the specific mechanism that links the two processes is much less clear. After the binding of the B subunits to the ganglioside receptor, the A subunit of CT or LT-1 must be internalised and activated. It is probably the insertion and activation events which account for the lag period observed between binding of the toxin to the host cell and the increase in cellular cAMP: with disrupted cells, no such delay occurs (Fishman, 1980).

In order to activate the enzymic subunit, it must firstly be nicked to A_1 and A_2 subunits (Clements and Finkelstein, 1979). Although it had been suggested that cholera *vibrios* produce a haemagglutinin / protease which is capable of mediating this cleavage (Gill and Rappaport, 1979; Booth *et al.*, 1984), recent evidence suggests that host factors may be involved. Such cleavage is not, however, essential for secretion as both porcine and human LT were released from a genetically modified *V. cholerae* without being nicked (Hirst *et al.*, 1984b).

Gill and King (1975) were the first to show that the A_1 subunit of CT is enzymically active only after reduction of the disulphide bond linking it to the A_2 subunit. A_1.

once activated, is enzymically functional in the absence of CT-B but is unable to penetrate target cells, therefore confirming that the B subunit is necessary for receptor binding. When this reduction occurs is not yet known.

The results from photolabelling experiments using liposomes suggest that, triggered by the reduction of the disulphide bond, the A_1 subunit of CT is inserted into the target cell membrane with A_2 and the B pentamer remaining on the outer surface (Tomasi and Montecucco, 1981) The crystallography work of Ribi *et al.* (1988) led to a similar conclusion. Using intact cells, however, Janicot *et al.* (1988) found that on toxin insertion, whole toxin, subunit A and subunit B were no longer accessible to antibodies, and therefore proposed that the holotoxin molecule is taken up by endocytosis.

Any theory that seeks to give an account of events within the cell must explain the fact that, whilst the majority of G_s is located in the brush border (apical) membrane, adenylate cyclase is found at the opposite end of the cell, on the basolateral membrane. Based on work carried out using the epithelial cell line T84, Lencer *et al.* (1995a; 1995b, 1997) and Wolf *et al.* (1997) have suggested the following mechanism for CT.

Binding to the cell is at caveolae (invaginations) in specialised detergent-insoluble microdomains (DIGs), where G_{M1} is clustered. The toxin-G_{M1} complex is drawn into an apical endosome, where it is activated by being nicked by a host serine / protease (when entering the cell *via* the basolateral membrane, this activation does not occur). The presence at the carboxyl-terminus of A_2 of an 'address' signal (**KDEL** in CT, **RDEL** in LT), directs the complex through the Golgi to the endoplasmic reticulum and thence from apically recycling to basolaterally targeted membrane structures. Fusion of these vesicles to the basolateral membrane is at a site near the G_s - adenylate cyclase complex. Reduction of the disulphide bond between A_1 and A_2, in this model, presumably occurs at the Golgi / endoplasmic reticulum stage of processing (Sears and Kaper, 1996).

What is clear is that once in the target cell, the toxin is unable to intoxicate any adjoining cells and is thus sloughed off with the dying cell. This imposes self-limitation on the extent of CT-mediated diarrhoeal diseases (Spangler, 1992).

Activation of G_S by cholera-like toxins

Once activated and *in situ*, A_1 catalyses the transfer of ADP-ribose from NAD to $\mathbf{R}^{201}$ on the α subunit of G_S. This promotes an increase in its sensitivity to GTP and loss of its GTPase activity thereby stabilising the G protein in its GTP bound form. As a result, GDP is released and binding of further GTP is promoted. Dissociation of the α subunit from the inhibitory β and γ subunits is also irreversible. The production of cAMP is now activated in an unregulated manner as adenylate cyclase activity undergoes a several hundred-fold increase. It is this uncontrolled and irreversible increase in cellular cAMP which was thought to cause the loss of water and electrolytes symptomatic of *V. cholerae* or *E. coli* intoxication (Fishman, 1990; Lee *et al.*, 1991; Moss and Vaughan, 1991). Contributory to ADP-ribosylation are ARF's, a family of GTP-binding host proteins, (M_r of ~20000) that have been shown to stimulate the activity of CT and LT. In both cases the activity was directly associated with the A_1 subunit, and therefore was not simply the acceleration of a physical function such as the release of A_1 from A_2. Exactly how these proteins fit in to the overall picture is still, however, unclear (Moss and Vaughan, 1991).

There does remain some doubt as to whether increased cAMP alone is responsible for all the observed physiological changes resulting from intoxication. It has been demonstrated that cAMP on its own does not cause fluid accumulation in ileal loops while prostaglandins (which show a marked increase in concentration during CT intoxication) can cause such accumulation. The suggestion is thus made that prostaglandins may also play a role in disease pathogenesis (Peterson and Ochoa, 1989). The fact that two other G protein subunits, $G_i\alpha$ and $G_o\alpha$, may also be ADP-ribosylated by CT (Rappuoli and Pizza, 1991) again suggests that the fluid

accumulation effects seen in cholera-like intoxication are the result of a rather more complex series of events than originally envisaged.

PART 3 : LTB PRODUCTION, PURIFICATION AND ANALYSIS

In view of the detailed characterisation of its reaction with G_{M1}, LT was an ideal model system for testing phage display. A surface of purified, immobilised LT-B needed to be prepared for panning the phage display libraries.

Origin of B-subunit producer *Vibrio* 60

Initially LT was purified from isolated whole cells (see below). However, more recently recombinant expression systems have been developed in order to learn more about the processes of export and secretion in toxin-producing strains. Bacterial strains capable of over-expressing and/or secreting toxin subunits have uses in their own right.

Vibrio species can export a number of products out of the cell including LT-B, the *etxB* gene product (Hirst *et al.*, 1984b; see page 33). LT-B has been expressed by a modified marine strain, *Vibrio* 60 pMMB68, developed by Leece and Hirst (1992). Expressing LT-B in this strain, rather than *V. cholerae*, circumvents the problems of pathogenicity or simultaneous production of CT (Marcello *et al.*, 1994).

Vibrio species 60 is related to *Vibrio anguillarum* and was one of a number of bacteria collected from sites along the Pacific coastline of Japan in June 1977 during a study of haemagglutinin, an important virulence factor that gives *Vibrio* species adhesive properties (Oishi *et al.*, 1979). In common with most of the other strains analysed, species 60 produced a non-lethal exohaemagglutinin as well as protease, amylase and DNase (Ichige *et al.*, 1988). Analysis of a spontaneous streptomycin resistant mutant of this strain, MVT606, suggested that it secreted protease by a specific pathway. Treatment of MVT606 with *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (NTG) created extracellular protein production (*epr*) mutants that were incapable of

secreting protease, which remained in the periplasm. MVT606 and one of its *epr* mutants (MVT1192) were used by Leece and Hirst for secretion of LT-B subunits.

The original source of the B subunit was a human ETEC, H74 - 114, from which Dallas (1983) purified two plasmids. From these he excised the *etx* genes by restriction with *Pst*I. This fragment was inserted into pBR322 yielding a plasmid, pWD600, that was used for mapping of the *etx* gene. Deletions from pWD600 were engineered *via* three internal restriction sites to yield further plasmids, which were transformed into *E. coli* hosts. One of these plasmids, pWD605, resulted from the deletion of four base pairs from the recognition sites of *Sst*I, and had the phenotype *etxA*$^+$ *etxB*$^-$, whilst another, pWD615, the result of adding four bases to the middle of the *Xba*I recognition site, was *etxA*$^-$, *etxB* $^+$.

Sandkvist *et al.* (1987) excised *etx* genes from both of these plasmids into a controlled expression vector, pMMB66EH. This was developed by Furste *et al.* (1986), who modified an IncQ replicon, RSF1010, by inserting a polylinker with an upstream *tac* promoter, a *lacI*^{q} ('super-repressor') gene and the *bla* (β-lactamase) gene for selection. IncQ replicons have a broad host-range and can be mobilised by members of the IncP group, (*eg.* the commonly used RP4) which can transfer genes between, and can themselves be maintained in, most Gram -ve bacteria.

A 0.59 kb *Eco*RI / *Hin*dIII fragment of pWD615 that carried the ribosome binding site and the entire coding sequence of the *etxB* gene as well as six codons, including the stop codon, from the 3' end of the *etxA* gene, was excised and inserted into the pMMB66EH polylinker, yielding pMMB68, in which expression of *etxB* was under the control of the *tac* promoter (see Figure 6).

Leece and Hirst transformed pMMB68 into *E. coli* S17-1 which has RP4 integrated chromosomally, and thence mobilised the plasmid into the marine vibrios MVT606 and MVT1192.

Growth of *Vibrio* 60 pMMB68 in a standard broth, and induction with the lactose analogue, iso-propyl thiogalactopyranoside (IPTG), resulted in the expression of *etxB* leading to the secretion of intact B pentamers into the medium (Amin and Hirst,

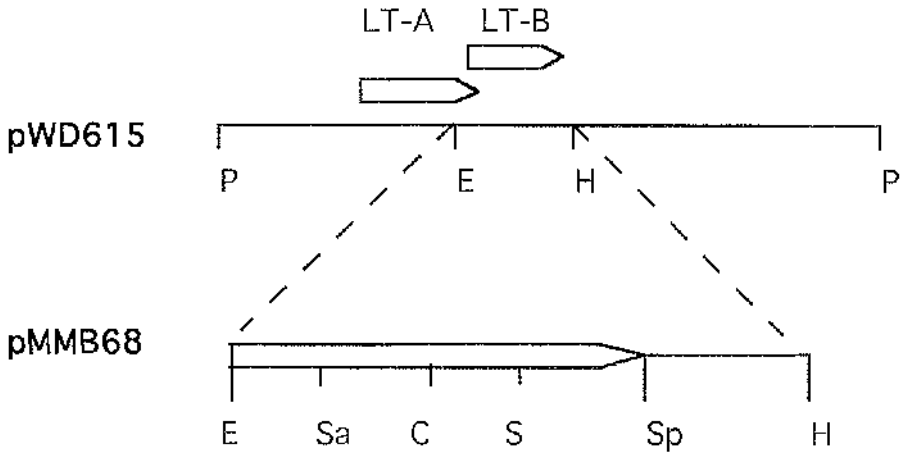


FIGURE 6 : Construction of plasmid pMMB68.

The 0.59kb fragment of pWD615 containing *etxB* was inserted into pMMB66EH yielding pMMB68 (Restriction sites : P, *Pst* I ; E, *Eco*RI ; H, *Hin*dIII ; Sa, *Sac*I ; C, *Cla*I ; S, *Sma*I ; Sp, *Spe*I). After Sandkvist *et al.*, 1987.

1994).

Purification of LT

As stated, the earliest attempts at purification of LT were hampered by its periplasmic location, and the presence of ST, giving results that were less than clear and occasionally contradictory. Generally, in excess of 20 litres of culture was grown in a fermentor, then centrifuged and filtered to remove cells. Supernatant fluids were concentrated by ultrafiltration and lyophilisation, then further purified by various combinations of ammonium sulphate precipitation, gel filtration and ion exchange chromatography, isoelectric focusing and preparative polyacrylamide electrophoresis (Moon *et al.*, 1970; Dorner, 1975; Dorner *et al.*, 1976; Finkelstein *et al.*, 1976; Clements and Finkelstein, 1978b). Definitive purification was finally achieved by Clements and Finkelstein (1979) who processed whole cell lysate, cell free supernatant fluid and a concentrated, ammonium sulphate precipitate of a sodium chloride extract of whole cells. Each of these was precipitated with 60% ammonium sulphate, dialysed, concentrated by ultrafiltration, and then, exploiting the ability of LT to bind weakly to saccharides (see page 39), affinity purified on an agarose column. LT was eluted from the column with galactose, the eluate concentrated by ultrafiltration and finally purified by gel filtration chromatography. Approximately 30 times more purified material resulted from the whole cell extract than from the other two sources. Around the same time, Kunkel and Robertson (1979) were purifying LT from 80 litres of culture by pH extraction of the whole cells, ammonium sulphate precipitation, affinity chromatography using hydrophobic and hydroxylapatite columns, and gel filtration. The Clements and Finkelstein method of LT purification remained favoured (*eg.* it was used by the University of Groningen group who first crystallised the LT toxin (Pronk *et al.*, 1985; Sixma *et al.*, 1991)) until the advent of expression systems such as that described above.

The method by which Amin and Hirst (1994) separated LT-B, produced by *Vibrio* 60 pMMB68, from other secreted proteins and from culture medium components was similar to those employed by Clements and Finkelstein (1979) and Kunkel and

Robertson (1979) (see above), and involved ultrafiltration of 24h culture supernatant fluids, precipitation with ammonium sulphate and hydrophobic interaction chromatography. In a more recently published procedure, LT-B fusion products were purified by isolating material from culture supernatant fluid by differential solubility in 20% and 60% ammonium sulphate and hydrophobic interaction chromatography (Loregian *et al.*, 1996).

A novel method was developed by Uesaka *et al.* (1994), from an earlier technique of Clements and Finkelstein, who had bound LT to an agarose affinity column. The supply of the appropriate agarose being apparently inconsistent, Uesaka *et al.* captured LT on commercially available, bead-immobilised galactose. They passed culture supernatant fluids from an *E. coli* strain that carried the LT genes through such a column, confirmed that it had bound, and then recovered the toxin by competitive elution with galactose solution. This method is similar in concept to that of Tayot *et al.* (1981), who had used a column of immobilised G_{M1} for affinity extraction of CT from *V. cholerae* supernatant fluid. Uesaka *et al.*, however, report that the Tayot *et al.* method failed to work with LT whereas their method successfully purified both CT and LT.

Analysis of purified LT-B

The success of any protein purification protocol can be conveniently assessed by sodium lauryl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli, 1970) and by testing the reaction of the product of purification with an antibody specific for the target protein. The antibody specificity of proteins separated by SDS-PAGE can be tested by Western blotting, in which the proteins on the gel are electrophoretically transferred to more flexible, less fragile nitrocellulose membrane for processing. After blocking non-specific binding sites, the membrane is probed immunologically with a mouse monoclonal antibody and an anti-mouse antibody coupled to horse radish peroxidase (HRP). Detection of reacting protein is by the

addition of an HRP substrate that results in a colour change. In the G_{M1} ELISA technique of Svennerholm and Holmgren (1978), LT-B, specifically captured by plastic-coupled G_{M1}, can be subjected to identical immunological detection.

Protein concentration is traditionally estimated by the method of Lowry *et al.* (1951), based on alkaline copper labelling of the protein. However Markwell *et al.* (1978) found that this method was adversely affected when test samples had a high sugar or EDTA content. They reported that sugar-stability could be restored by increasing the concentration of copper tartrate. An alternative method of protein estimation is that of Bradford (1976), which detects the shift in the absorption maximum of Coomassie Blue (from 465nm to 595nm) on combination with protein.

Immobilisation of LT-B for biopanning

As reported earlier, (see page 14), biopanning with phage display libraries has, traditionally, been carried out on biotinylated protein immobilised on streptavidin. In terms of effectiveness and convenience, this remains a method of choice.

Biotin, a small water-soluble vitamin, spontaneously bonds with avidin, a protein of hen egg white. Although suitable in many applications, avidin is basic and may therefore bind proteins non-specifically. Its long carbohydrate chains may also interact with sugar-binding sites. Use of the avidin derivative streptavidin, a protein of *Streptomyces griseus*, can circumvent these obstacles. The exact nature of the interaction between streptavidin and biotin remains unknown, although it is believed that the tryptophan rings of the streptavidin molecule are the target for biotin attachment.

A protein such as LT-B can be rapidly coupled to biotin, which forms with streptavidin the strongest non-covalent bond in nature, by use of commercially available kits. Biotinylation is facilitated by the N-hydroxysuccinimide ester of biotinamido-caproic acid (BAC-sulfoNHS), which couples spontaneously to primary and secondary amines, including the amino acid lysine, by nucleophilic attack at the ester linkage. Biological activity is unaffected as this system allows for biotinylation in aqueous solution at neutral pH.

The caproic acid arm of the BAC molecule acts as a spacer molecule, reducing steric hindrance between the protein and the biotin / streptavidin interface (see Figure 7).

There are alternative methods of immobilisation for LT-B. For example, to do so via the specific ligand for LT-B, G_{M1} would be highly effective, and is indeed the basis of G_{M1} ELISA (described on page 82), however, it would not be an appropriate method when phage are sought that bear peptides that will occupy the G_{M1} binding site.

LT-B can also be immobilised by capture by specific antibodies. The anti-LT antibody 118.87 (see page 81) or a commercially-available anti-cholera toxin antibody, when attached to plastic, will both immobilise added LT-B without compromising the binding of G_{M1} and therefore, presumably, G_{M1}-mimicking phage.

Although antibodies will adhere directly to naked plastic, a more consistent coating can be obtained by binding them indirectly via Protein G. This Staphylococcal cell wall protein, first isolated by Björck and Kronvall (1984) will attach to a plastic surface by non-specific hydrophobic interaction and is capable of non-immune binding to most immunoglobulins .

PART 4 : SMITH HEXAMER LIBRARY

The first library tested against LT-B^{biotin}, that of Scott and Smith (1990), displayed a hexamer fusion on protein III.

The Smith hexamer library

The Ff class are the phage of choice for display library work as they satisfy the principal criterion for phage propagation - they do not normally kill their host (see 'Ff class filamentous phage', page 3). Such phage, however, must be propagated by continual cycles of infection.

A new strain, fd-tet, was developed from phage fd by inserting the tetracycline resistance segment of transposon Tn*10* into the negative strand origin of replication,

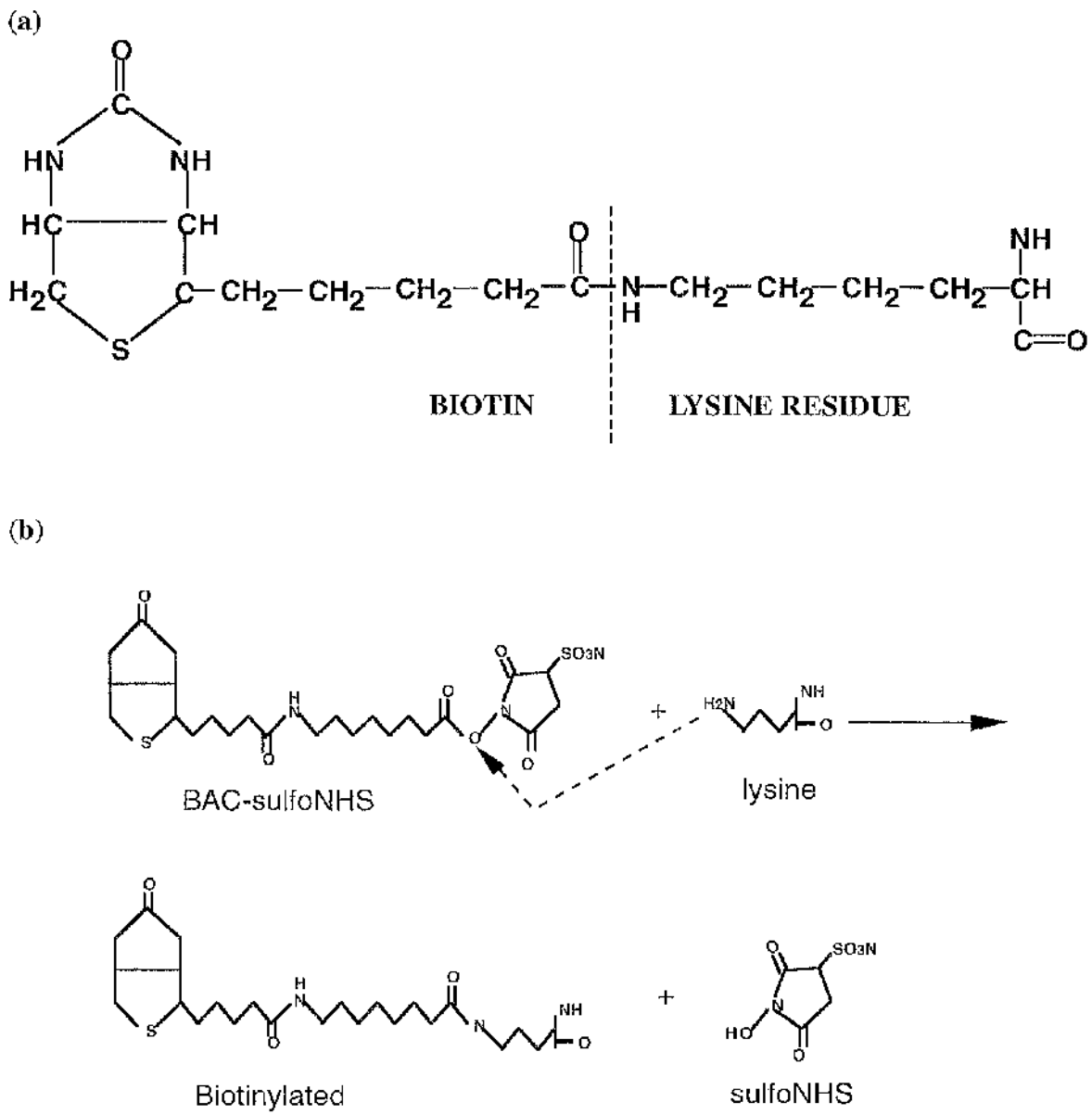


FIGURE 7 : Biotinylation of lysine residues.

(a) Direct coupling to lysine; (b) Coupling *via* an N-hydroxysuccinimide (NHS) ester of biotinamido-caproic acid (BAC). Nucleophilic attack by the lysine amino group is indicated by the dashed line.

meaning that infected host cells could be propagated on tetracycline agar in bacterial colonies, in the manner of a plasmid, without the need for infection (Zacher *et al.*, 1980; Smith, 1988a).

The Smith hexamer library was developed in a vector, fUSE5, that was based on fd-tet (see Figure 8). Firstly, gene III of fd-tet (**A**) was modified by the introduction, by oligoucleotide-directed mutagenesis, of *Sfi*I restriction sites between the codons for $\mathbf{A}^1$ and $\mathbf{E}^2$, to yield the phage vector fUSE5. Intracellular RF DNA (see page 8) of fUSE5 (**B**) was prepared from a 2 litre stationary phase culture by boiling and caesium chloride centrifugation. A 33 base pair degenerate *Bgl*I fragment (**C**), prepared by polymerase chain reaction, and consisting of 32^6 different nucleotide sequences was cloned into the *Sfi*I site of fUSE5 (**D**) (the designation NNK will be explained in 'Biasing construction of an epitope library', below). The amino acid sequence of the amino-terminus of the mature form of pIII is shown in part **E**. The underlined portions are present in wild-type fd-tet and the part designated $\mathbf{X_6}$ represents the random hexapeptide insert. The ligation product was electroporated into an *E. coli* host that was then used to inoculate 64 tetracycline-containing agar plates, on which only those that had taken-up phage could grow. After incubation, each plate contained around 3 million clones which were eluted and the phage precipitated and purified (Scott and Smith, 1990).

The final library consisted of 1.3×10^{14} phage, representing 2×10^8 different clones. Hosts infected with this phage under normal growth conditions give a yield of progeny that is around 25% of that seen when wild-type fd are used. However, in contrast to fd, yields cannot be increased by vigorous aeration in rich media. Host cells infected with fd-tet are resistant to tetracycline at a concentration of at least 40μg ml^{-1} but result in the production of near-invisible plaques. Enumeration by plaque number is therefore not feasible. Rather, transductant colonies are counted, these being host cell colonies that have been rendered capable, by fd-tet infection, of propagation on tetracycline-containing media (Smith, 1988a).

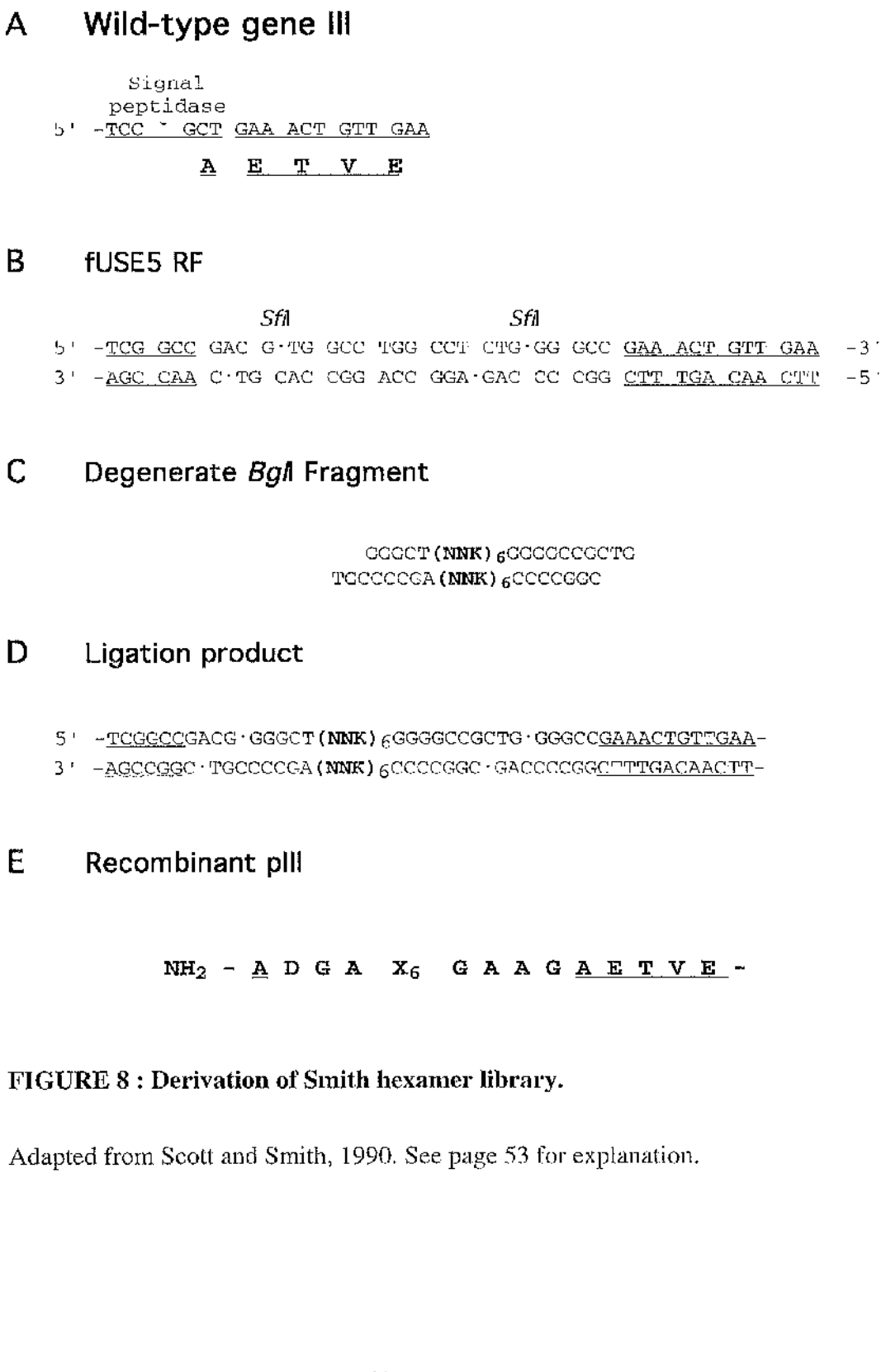


FIGURE 8 : Derivation of Smith hexamer library.

Adapted from Scott and Smith, 1990. See page 53 for explanation.

Biasing construction of an epitope library

A basic element of phage display library construction is the insertion of random oligonucleotides into the DNA coding for a coat protein of the phage. However, within the genetic code there are not equal numbers of codons coding for each of the amino acids. As a result, nucleotide sequences consisting of genuinely randomised bases will not code for peptides featuring an equal distribution of the 20 amino acids, but will rather exhibit a bias in favour of those for which there are more codons. In an attempt to overcome this, both Scott and Smith (1990), and Cwirla *et al.* (1990), produced phage display libraries carrying random hexapeptides by inserting the coding sequence $(NNK)_6$, where N represents an equal mixture of the four bases A, C, G and T while K represents an equal mix of G and T only. This will produce single codons for 12 of the amino acids, two codons for 5 more and three codons for the remaining 3 amino acids as well as the amber stop codon, TAG. Thus in randomised hexapeptides produced under NNK biasing, amino acids encoded by three codons will be incorporated 3^6 times more often than those with only one codon. This ratio increases threefold with each additional peptide included in the fusion (*ie.* heptapeptide, 3^7; octapeptide, 3^8;....*etc.*). The full genetic code, the NNK abbreviation and codon frequencies are shown in Appendices 5 and 6 (pages 188 and 189).

Devlin *et al.* (1990), in their pentadecamer (15 amino acid) library used NNS with S representing C and G. Like NNK, NNS can yield only 32 of the possible 64 codons, and also codes for all 20 amino acids and only one stop codon. Kay *et al.* (1993) used NNB, with B representing C, G and T, to construct their library of 36 amino acid inserts. This particular bias allows 48 codons but as this still includes only one stop codon the frequency with which it should occur is less than that seen with NNK or NNS. A disadvantage of NNB is the disparity in amino acid representation *eg.* five of the possible codons are for **S** whilst **M**, **E**, **Q**, **K** and **W** are coded for only once meaning that **S** could appear 5^6 times more frequently in a 'random' hexapeptide

library than those others. The number of possible combinations is, however, sufficiently high to ensure that each displayed peptide should be unique.

Another strategy is that employed by McLafferty *et al.* (1993). Their gene III library phage carried 18 amino acid fusions, and had this somewhat unusual pattern:

A E **G X C X Y Y X C X S Y I E G R I V E** T V E S

where the insert is shown in bold letters. The variable part of the insert is described as being variegated rather than fully randomised, in that **X** is produced by NNT and **Y** by NNG. The result of this bias is that **X** can be any one of fifteen amino acids and **Y** can be any one of thirteen. Eight amino acids can appear in either position. The reported advantage of this bias is a reduction in the number of redundant codons, with 8.55×10^6 different peptides resulting from only 1.68×10^7 DNA sequences. By comparison, the NNK bias employs 64-fold more DNA sequences to produce only 8-fold more possible peptides. The obvious disadvantage is that not every possible peptide is coded for. The two **C** residues are placed to attempt to induce disulphide constraint (see page 71).

Other library construction strategies

Devlin *et al.* (1990) employed a further tactic when constructing their library. The amino-terminal amino acids of the mature pIII produced was predicted to be -

$$\mathbf{A\ E\ (XXX)_{15}\ P_6\ A\ E}$$

with first **A** and **E** being those seen at the amino-terminus of wild-type phage. They were included here to overcome the potential signal sequence processing difficulties mentioned earlier (see page 12). The second **A E** was included in order that, after the artificial insert, the wild-type pIII sequence would be present in an uninterrupted form. The poly-proline sequence was added as its rigid conformation might serve to distance the added peptide fusion from the rest of the protein, thus rendering it more accessible to binding surfaces. This library was used to identify peptides which would bind streptavidin and was thus the first demonstration that libraries can include peptide mimics of non-protein ligands.

Uses of the Smith hexamer library

In the original paper by Scott and Smith, two monoclonal antibodies (MAb's), specific for the **DFLEKI** hexapeptide epitope of the protein myohemerythrin, were biotinylated, immobilised and probed with a small library (2.3×10^7 clones) that had been prepared alongside the major library. In three biopanning rounds, yields of eluted phage, relative to the number of phage originally applied, rose from 3×10^{-5}% in the first round to 1% in the third round. In sequencing the final round clones, **DFLEKI** did not occur, however, the most likely explanation for this fact is pure chance. In a complete NNK-biased hexamer library there are 32^6 possible random nucleic acid sequences. However, only three of these will code for **DFLEKI**. Amongst the relatively small number of phage tested by Scott and Smith (only 23 million clones) there was little possibility of the three sequences actually occurring. However, although **DFLEKI** did not occur, ELISA showed that the phage that were selected did bind to the MAb's as effectively as a phage manufactured to display that peptide. Inhibition by a peptide including the **DFLEKI** sequence also followed a similar pattern. By way of a negative control, original library material and preparations of non-binding phage showed no binding in similar assays.

Both Smith and Scott continued to apply the library in other applications, and in 1993 used it to attempt to find agonists or antagonists of the S-peptide, one of the two products (the other is S-protein) of the partial digestion of ribonuclease (Smith *et al.*, 1993). Although neither of the fragments is independently enzymic, when re-combined activity is restored. This is due to the interaction of a 15 amino acid segment of the S-peptide with a cleft in the S-protein. Smith *et al.* used the library to look for other peptides capable of modulating enzymic activity. Panning revealed a series of peptides with near-consensus sequences. One of these, **YNFEVL**, was shown to be capable of blocking the enzymic activation normally resulting from S-peptide/S-protein interaction, indicating its antagonistic quality.

The use to which Scott *et al.* (1992) put the library is, perhaps, of more relevance to this project as they bound library phage to a ligand which does not naturally bind peptides. Their work involved Concanavalin A (ConA), a globular protein lectin of the Jack bean (*Canavlia ensiformis*) which, at neutral pH, consists of four identical monomeric units, each of which has a site specific for the binding of saccharides. These binding sites are most complementary to α-D-mannopyranosyl residues, which are thought to bind principally through hydrogen bonds. This high avidity, metal ion-dependent, attachment to the saccharide is thought to be part of a cell recognition system (Becker *et al.*, 1976; Goldstein, 1976).

By use of the library in a standard panning experiment, they were able to identify peptides which bound at, or near, the sugar-binding site of ConA with an avidity equivalent to that of methyl α-D-glucopyranoside, a natural substrate for the lectin. Phage eluted by acid from the Con A surface were characterised in four ways : (i) Sequencing revealed a consensus motif (**YPY**); (ii) Phage selectively bound to ConA in ELISA assays; (iii) Phage could be captured by immobilised ConA and (iv) Phage-binding to ConA was inhibited by the addition of methyl α-D-mannopyranoside.

The most avidly binding phage, which carried the sequence **MYWYPY**, not only bound ConA with an affinity comparable with that of the natural ligand, but showed remarkable selectivity for ConA alone and not for related lectins.

Simultaneously with Scott *et al.*, Oldenburg *et al.* (1992) also reported the isolation of phage which bind to the sugar-binding site of Con A. Their work, utilising an octapeptide library, drew similar conclusions, and identified the same consensus sequence (**YPY**).

Other uses of phage display in non-protein systems

In a paper which is contemporary with Scott and Smith, Devlin *et al.*, (1990) (see page 56) were the first to demonstrate that phage display could be used to identify peptides which could mimic the binding of non-proteins. The library was exposed to blocked streptavidin, a protein that has no reported affinity for peptides, and binding

phage were eluted. Phage stocks were prepared from 60 resultant plaques, then used in binding assays and in sequencing work.

The binding assay involved comparing ratios of selected phage and wild-type M13mp19 phage eluted from a streptavidin surface after exposure to a mixture of the two. The results suggested that 56 of the 60 selected phage were enriched by a factor of 10 relative to wild-type M13mp19. Adding biotin to the assay system to block the streptavidin surface significantly reduced this enrichment. Sequencing of 20 of these enriched phage revealed nine different peptide sequences, all of which contained the sequence **HP**, and seven of which contained **HPQ** at some point in the 15 amino acid random insert. This consensus was thus taken to represent the streptavidin-binding sequence. The authors point to the fact that both histidine and biotin contain carbon- and nitrogen- containing rings.

Hoess *et al.* (1993) investigated peptide binding to a monoclonal antibody, B3, which normally recognises the carbohydrate antigen, $Lewis^y$, which appears on the surface of adenocarcinoma cells. On sequencing twenty binding clones isolated from an octapeptide library, they were all found to contain a sequence, **PWLY**.

In a review of this branch of the subject Hoess (1993) suggests that the similarity of this sequence to that found in the Con A work may indicate that certain amino acids are best able to mimic specific sugars. His suggestion is that aromatic amino acids, with their cyclic structures, are able to mimic the saccharide rings, and that **Y** may be able, in addition, to establish hydrogen bonding through its -OH group. This theory concurs with an idea presented by Scott *et al.*, (1992) who supported the proposal with evidence that the replacement of **Y** with **A** in ConA-binding peptides resulted in a ~50% reduction in inhibitory strength, and the fact that almost all of the selected peptides contained at least one residue which was aromatic and one which was hydroxylated. In a more recent report, it has been proposed by Phalipon *et al.* (1997) that **P** can serve to orient aromatic peptides in a manner similar to the branching naturally observed in carbohydrates.

However, as the lab-work stage of this project neared completion a paper by the Scott and Oldenburg group (Bonnycastle *et al.*, 1996) cast doubt upon their earlier finding, reporting that the peptide that had apparently mimicked the binding of α–D-methylmannoside had not reacted with the sugar-binding site, but rather with an adjacent site. In a paper produced after further study of the phenomenon, (Harris *et* al., 1997) they stated that true carbohydrate-mimicking peptides appear to be relatively rare. The reasoning behind and implications of these findings for this project are discussed on page 180.

PART 5 :LUZZAGO CONSTRAINED NONAMER LIBRARY

pEMBL family of single stranded vectors

Dente *et al.* (1985) sought to create a new family of single-stranded cloning vectors without the limitations of insert size that were a feature of the M13 phage derivatives prevalent at that time. In order to achieve this, they modified one of the widely available double-stranded plasmids to single-stranded DNA form, following the lead given by Dotto and Horiuchi (1981) and Dotto *et al.* (1981), who had cloned into pBR322 the intergenic region of the Ff phage f1 responsible for phage replication and packaging (*ie.* the origin of replication or f1 *ori*). Replication of single-stranded phage involves the export of single-stranded phage genomes derived from the double-stranded RF. By superinfecting cells carrying this modified pBR322 with f1 phage, progeny virion capsids carrying either pBR322 ssDNA or f1 ssDNA were produced at roughly the same frequency.

In order to create their family of cloning vectors, Dente *et al.* chose to insert f1 *ori* into pUC plasmids. The plasmids they chose all carried an ampicillin resistance gene and the gene that codes for synthesis of the α-peptide of β-galactosidase (*lacZ*). Differentiation between the pUC vectors chosen for modification by Dente *et al.* was dependent on the size and content of a polylinker sequence located within *lacZ*. Insertion of DNA into this polylinker resulted in inactivation of *lacZ*, and

recombinants could be easily identified by growth of the host cell on medium containing X-gal (5-chromo-4-chloro-3-indolyl-β-D-galactosidase); clones producing β-galactosidase were blue, those in which *lacZ* was insertionally inactivated were white.

Dente *et al.* inserted the 1300 base-pair f1 *ori* sequence adjacent to *lacZ* in four pUC plasmids. This insertion did not affect *lacZ* function. The orientation of the insert determined which of the two strands of *lacZ*, from the pUC vector, were encapsidated in the virions resulting from infection of a bacterial cell. In pEMBL(+) vectors, the antisense strand of *lacZ* was present, in pEMBL(-) vectors, the coding strand. One such phagemid, pEMBL19^{+}, was used by Felici *et al.* to create the vector with which subsequent pVIII libraries were produced.

Figure 9 shows a map of a generic pEMBL vector.

Construction of Felici *et al.* phagemid vector pC89

To create their own library, Felici *et al.* (1991) isolated a 481 nucleotide section of the f1 genome by restriction with *Sna*BI and *Bsm*I. The 2 base extension left by *Bsm*I was removed by Klenow treatment. This fragment contained DNA sequences for the entire gene VIII and the amino-terminus of gene III. This was inserted into pEMBL19^{+} at *Sma* I, one of the fourteen restriction sites in the polylinker that defines this member of the pEMBL family. Clones in which gene VIII was placed under the control of the *plac* promoter were easily differentiated as they formed blue colonies on X-gal. The residual gene III DNA was removed and gene VIII was joined directly to the gene for the α-peptide of β-galactosidase in such a way as to introduce an 'amber' stop codon (*ie.* TAG) between the two. The resulting plasmid (named pC83) could be distinguished from parent plasmids by its ability to form a blue colony in a *supE* or amber suppressing host, whilst forming a white colony in a *sup*$^{\circ}$ host. The pEMBL polylinker was removed and an *Eco* RI and a *Bam* HI restriction site introduced, these flanking a second amber codon, and phagemid vector pC89 was thus complete. A *supE* host carrying this

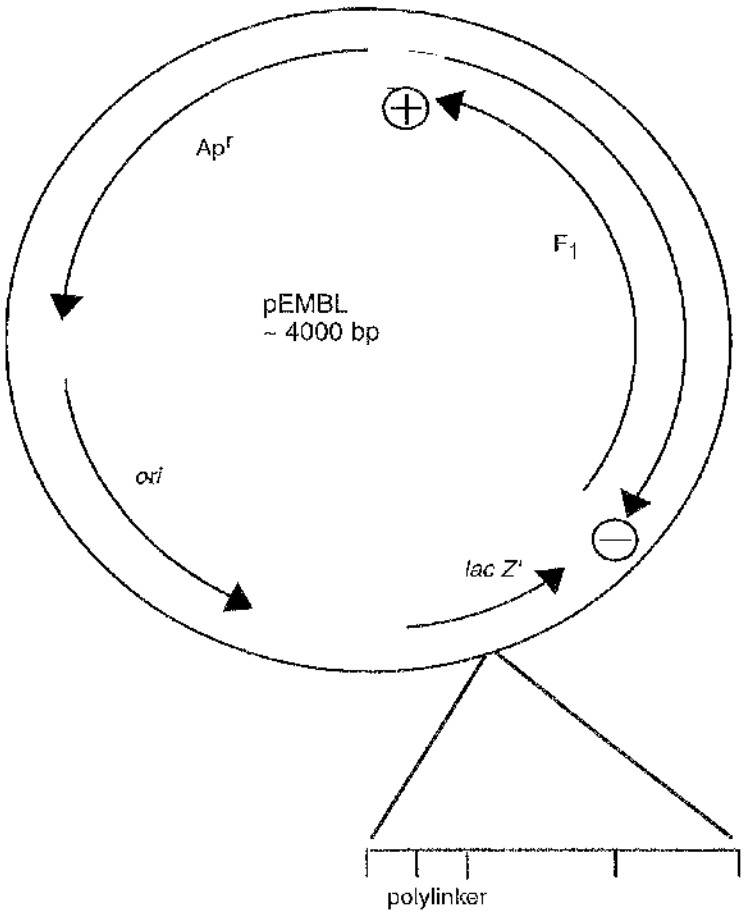


FIGURE 9 : Map of a generic pEMBL plasmid.

Markers : Ap^r, ampicillin resistance; *ori*, ColE1 origin of replication; *lacZ'*, β-galactosidase; F1, origin of replication of phage f1 in its two possible orientations (+ and -). Family members are defined by the size and composition of the polylinker.

phagemid, with its double-amber, would appear to be only a white to pale blue colour, as lower amounts of α-peptide would be produced in this case. However, oligonucleotide insertions at the *Eco* RI - *Bam* III restriction site would remove the second amber (see Figure 10) codon so *supE* hosts carrying pC89 modified in this way would be an intense blue on X-Gal.

Felici *et al.* created a phage display library by synthesising a double stranded 46-mer that included a random 26-mer oligonucleotide (produced by the adding equimolar quantities of each of the four nucleotides to the synthesis reaction) positioned between *Eco*RI and *Bam*HI restriction sites. Restriction with these two endonucleases yielded a 32 base pair fragment with cohesive ends, suitable for ligation into gene VIII of a similarly digested pC89 phagemid. The recombinant phagemids were then electrotransformed into *Escherichia coli* cells and antibiotic resistant cells selected for on agar containing ampicillin. In this way, approximately 9.4 x 10^7 independent clones were generated, with around 40% of them giving the dark blue appearance on X-gal that indicated insertion of random oligonucleotide into the *Eco*RI - *Bam*HI restriction sites of gene VIII. This process is summarised in Figure 11.

As pC89 only contained the genes for pVIII and the f1 *ori*, a second so-called helper phage was required to produce infectious phage virions from these selected clones. Although wild-type phage can be used as a helper, (see the work of Dotto *et al.* above) it is preferable to use a packaging-deficient helper that encodes all the capsid proteins but lacks f1 *ori* in order to promote complementation between phagemid and helper. Felici *et al.* used F1 IRI to co-infect plasmid-harbouring cells to produce infectious virions that carried both wild-type pVIII and the coat protein carrying a random amino-terminal peptide.

Luzzago constrained nonamer library

A library based on that prepared by Felici *et al.* was prepared by Luzzago and co-workers by introducing a **C** codon at either end of the random oligonucleotide. The full implications of this addition are discussed on page 71.

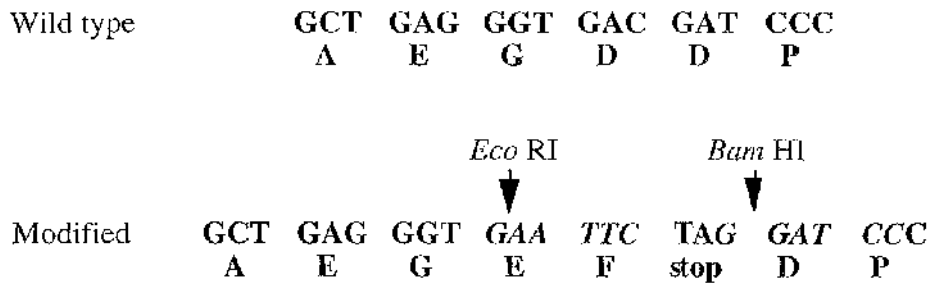


FIGURE 10 : Modification of the amino terminus of gene VIII in construction of pC89.

Restriction sites are italicised and cutting points indicated. After Felici *et al.* (1991).

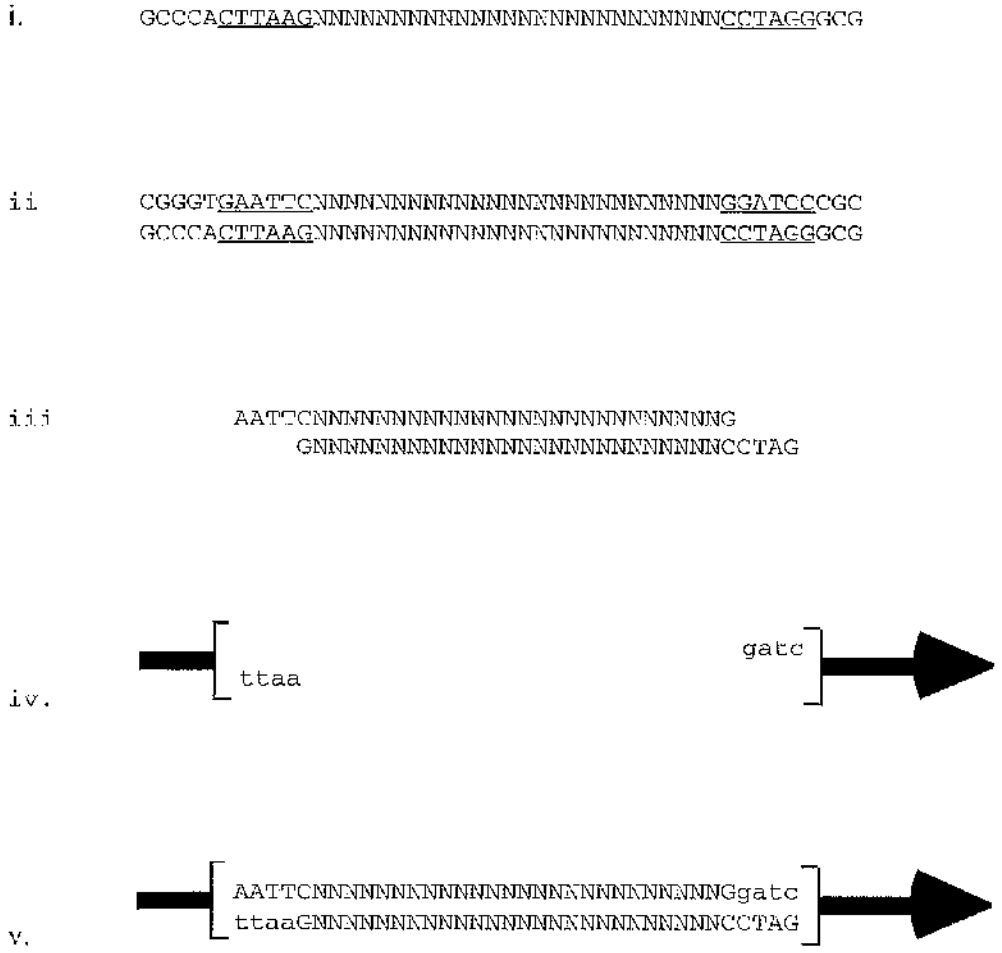


FIGURE 11 : Strategy for construction of random nonamer library in pC89.

The oligonucleotide (i) consisted of 46 bases including the 26 randomly generated bases (N), and *Bam*HI (underlined left) and *Eco*RI (right) restriction sites. The complementary strand was synthesised (ii) and the duplex digested with *Bam*HI and *Eco*RI (iii). The vector pC89 was digested to produce compatible ends (iv). Finally, the oligomer was ligated into the plasmid to yield the library (v). Adapted from Felici *et al.* (1991).

An oligonucleotide including a 27 base random sequence, flanked by the bases GCA (complementary to the **C** codon TGC) and including restriction sites for *Eco*RI and *Bam*HI was synthesised, along with a complementary primer. After annealing, the primer strand was extended and the resulting duplex was digested with the appropriate restrictionendonucleases. The fragment was then inserted into pC89, to form the Luzzago nonamer library. This process is summarised in Figure 12.

The amino acid sequence of the final insert was therefore - **E F C** $\mathbf{X_9}$ **C G** -, the additional glycine codon (GGG) being added in order that the restriction site for *Bam*HI (GGATCC) would be present in all displayed peptides. This insert therefore yielded modified protein VIII with the sequence, from the mature amino-terminus,

A E G **E F C** $\mathbf{X_9}$ **C G** D P A.....

Wild-type protein VIII has the sequence

A E G **D** D P A.....

D having been lost during the insertion.

The library consisted of 1 x 10^8 independent clones; plating on X-gal revealed that 25% carried a productive insert. The helper phage M13K07 was employed to harvest phage virions. This helper, which is based on phage M13, was developed by Vieira and Messing (1987) to package preferentially phagemid DNA ahead of phage DNA. The complete absence of **C** residues from wild-type pVIII aided analysis of the thiol content of isolated clones. Of six that were chosen at random, five were displaying some peptides in cyclised form, indicating that the peptide was held as a rigid loop structure. Luzzago *et al.* used their library to investigate the selection of phage by monoclonal antibodies (MAb) against discontinuous epitopes. Others, including Scott and Smith (1990), Cwirla *et al.* (1990) and Felici *et al.* (1991) had previously shown that the sequence of the peptides selected by MAbs against linear epitopes were highly homologous to the original epitope. However, in another paper by Luzzago's group, (Felici *et al.*, 1993) using the library described above, it had been shown that peptides that mimic a discontinuous epitope have no sequence homology with the original antigen. To investigate further the nature of the peptides selected by MAbs

(A) 5'——GCTTTTGCTG **GATCCCCGCAN$_{27}$GCAG** AATTCACCCTCAGCAG

(B) **CGTCTTAA** GTGGGAGTCGTC

Insertion yielded the following sequence -

ggt g**AA TTC TGC XXX$_9$ TGC GGG G**at ccc

G **E F C N$_9$ C G D** P

FIGURE 12 : Oligonucleotides for construction of a constrained random nonamer library in pC89

Oligonucleotide (B) was annealed to (A), extended and the duplex digested with *Eco*RI and *Bam*HI. The fragment was then ligated into pC89. The underlined codons flanking the N_{27} random segment are complementary to TGC, the codon for cysteine. After Luzzago *et al.* (1993)

against discontinuous epitopes, a MAb against human heart ferritin was panned with their cysteine-constrained library. From a first round application of 10^{10} phage, 10^{4} were eluted. At 10^{-4} %, this return was considered slightly low, although Smith suggests 10^{-5} % as a 'normal' return in the first round. Luzzago *et al.* chose not to proceed to a second round of panning, in case of loss of low affinity phage, but rather screened the first round phage directly with a plaque assay. Approximately 10^{5} plaques were lifted onto nitrocellulose and probed with the anti-ferritin MAb. More than 300 gave positive signals. Thirty-eight of these positive plaques were eluted and sequenced. They were found to have 11 different sequences, corresponding to two major groups. Analysis of two phage types, both from the first of these groups, suggested that only one carried predominantly cyclised peptides.

In work of significance to this project, the Felici and Luzzago libraries have been subsequently employed in an attempt to produce mimics of carbohydrate antigens for antibody production (Phalipon *et al.*, 1997). The outcome is discussed on page 177.

The Luzzago library is, of course, a gene VIII library, and it is worth considering what influence this may have on the work. In most reviews of the subject, gene III and gene VIII are accorded equal status as potential sites for the insertion of random DNA for display purposes, with little evidence being presented to suggest that either approach is more consistently successful than the other. In considering, therefore, why gene III has thus far predominated, the extraordinary generosity of George Smith in making freely available his bacterial and phage strains, his primers and vectors, and, of course his hexamer library, to research groups throughout the world, must be considered a major factor. Of a list of twenty-three published libraries given by Burritt *et al.* (1996), all but two featured fusion to protein III. Of these twenty-one, seven were produced using the Smith vector fUSE5, whilst a further four used fAFF1, a vector created by Cwirla *et al.* (1990) using strains and phage supplied by George Smith. The only two gene VIII libraries in the list were those of Luzzago and Felici. Smith and Petrenko (1997), in a list of fifty-nine published libraries do include sixteen different gene VIII libraries, however, it should be noted that eleven of these

came from a single paper by Bonnycastle *et al.* (1996) and consisted of variations, in terms of insert length and the relative position of constraining Cs, on a single, type 88 progenitor. The others were the Felici and Luzzago 8 + 8's, an 8 + 8 pentamer, a type 88 20-mer and an octamer library, that Smith and Petrenko themselves had created and that represents the only listed example of a type 8 fusion.

Of course, the influence of George Smith does not entirely explain the pre-eminence of gene III. Whilst it is generally agreed that the high number of copies of protein VIII on the phage capsid (2700 or more) should favour its use, especially in comparison with the very low numbers of pIII molecules (5 or less), protein III function appears to be less easily perturbed by the addition of heterologous sequences at its amino-terminus. Problems with maturation of the capsid affect the consistent production of infective phage consisting exclusively of modified gene VIII product if the fusion is in excess of five amino acids in length (Greenwood *et al.*, 1991). Whilst Felici *et al.* (1991) suggest that fusions of nine residues or more cannot be assembled at all, Iannolo *et al.* (1995) demonstrate that this assertion is broadly, but not absolutely correct. Various lengths of random insert were tested, 100 clones at each length, and the percentage that tolerated an insert of each length was tabulated. They confirmed that an insert of six amino acids was always tolerated, and that 40% of eight residue, 20% of ten residue and 1% of 16 residue inserts were also satisfactorily processed into phage. Suggested explanations for this are that the density of packing of pVIII in the capsid precludes further extension of the amino-terminus (Cesareni, 1992) that certain fusions may be unable to be translocated to the host-cell membrane (Iannolo *et al.*, 1995; Malik *et al.*, 1999) or that whilst large inserts may be tolerated by the pVIII in the body of the virion, they may prevent those at the distal or proximal ends interacting with pVII and pIX, or pVI respectively (Makowski, 1994).

Despite the suggestion of Iannolo *et al.* that a hexapeptide insert would be tolerated in a fully polyvalent display, Burritt *et al.* (1996) could find no record of such a library. However, as mentioned above, Smith and Petrenko do report on their own library that features octapeptide extensions to every copy of pVIII (Petrenko *et al.*, 1996). To

create this library, the authors simply accepted that, as Iannolo *et al.* had predicted, only around 40% of their initial clones would tolerate the insert, and sought to use that portion in panning work.

By contrast, the amino-terminus of the pIII molecule, although pivotal in the infection process, suffers no such restriction on the size of insert.

Allowing that the advent of the two-gene systems (type 88 or 8+8) has overcome this limitation on the use of gene VIII, one might have anticipated that its natural advantage, *ie.* high copy number, would have promoted its wider use. However, this apparently advantageous position is undermined by reports that the 33 or 3+3 types of display can be used to produce monovalent pIII fusions, in which only one of the pIII molecules carries the fusion. This arrangement is reportedly more favourable than polyvalent display on pIII or pVIII because a reduction in the overall avidity of the phage apparently results in greater selectivity by the library, with binding to low- or moderate-affinity sites being greatly reduced (Lowman *et al.*, 1991; Hoess, 1993). Cwirla *et al.* (1990), in one of the earliest papers on phage display, reported the difficulty of selecting fusion peptides of anything other than moderate affinity, as such peptides will outnumber the high-affinity binders. The authors conceded that the selection of a wide range of phage in the first round of panning could eliminate non-specific, 'background' phage, and that reduction of the concentration of the target molecule in subsequent rounds of panning could serve to eliminate the weakest binders, but concluded that some form of monovalent display would be ultimately desirable. McConnell *et al.* (1994) drew similar conclusions regarding monovalent display, but did not exclude the idea that the use of polyvalent display in the first round could provide potential leads. Barrett *et al.* (1992) using phage displaying fused peptides known to represent high and low affinities, demonstrated that the concentration of free peptide required to completely block binding of these phage was far in excess of the K_d of the peptide itself. They reasoned that the only explanation for this was that polyvalent display on the phage surface resulted in multivalent attachment, and therefore binding with higher affinity. Thus, very low affinity

peptides, if polyvalently displayed, could remain attached during washing and then be released in the acidic eluate. Bass *et al.* (1990), were the first to produce a monovalent pIII display system by creating a phagemid that, in conjunction with a helper phage, gave capsids that generally carried no more than one fusion peptide alongside four wild-type pIII molecules.

Another suggested advantage in pIII fusion is the location of this protein. Burritt *et al.* (1996) speculate that the terminal location of pIII is better suited to anchoring the phage to the affinity matrix, but it would seem plausible that interaction between the matrix and several points on the capsid surface, a possibility with polyvalent pVIII display, could achieve binding of comparable stability. In any case, Burritt *et al.* suggest that this fact merely compensates for the low copy number of pIII, not that pIII is an inherently more suitable site for the insertion of the fusion peptide.

Finally, an application for which polyvalent display on protein VIII is ideal, as highlighted by the work of de la Cruz *et al.* (1988) and Greenwood *et al.* (1991), is immunisation of animals, with the high number of fusion peptides ensuring maximal antibody responses. Despite the presence of wild-type pVIII, the latter group found that the antibody response to the fusion was far greater. They proposed two possible explanations for this: that the fusion was more immunogenic; that its size shielded the wild-type protein from the immune system.

Cysteine constrained libraries

The amino-terminus of mature pVIII, like that of pIII, is not fixed into a defined conformation and short peptides fused there are equally likely to assume a range of conformations, constrained only by their thermodynamic stability. As a result, the concentration of any one conformation may be low (Folgori *et al.*, 1994). In order to overcome this, constraint can be introduced into libraries by flanking the fused peptide with **C** residues. By spontaneously forming cyclic disulphides during phage assembly, these **C**'s serve to stabilise and restrict the conformations adopted.

Koivunen *et al.* (1994) report that a constrained peptide showed greatly diminished activity when the disulphide bond was reduced.

The first deliberate attempt to include C constraint in a phage display library was by O'Neil *et al.* (1992). Their gene III hexamer library was created to study the glycoprotein integrin IIb/IIIa, which was known to recognise the sequence motif **RGD** with greater affinity when it was borne within a cyclised peptide. The O'Neil *et al.* constrained library showed a 500 times greater amplification of selected clones, in the second round of panning, than a similar, but unconstrained, library.

The opposite effect was observed by McConnell *et al.* (1994) who again tested parallel constrained and unconstrained libraries with a model system (a MAb against the peptide angiotensin AII). In their case, however, it was the unconstrained library that yielded sequences known to feature within AII. The constrained library, however, identified 'mimotopes' - novel sequences bearing no relationship to the original epitope. The conclusion drawn from this result was that constraint rendered the natural binding motif unrecognisable, and it therefore was not selected from the library. To test this theory, an oligonucleotide was synthesised that coded for the consensus sequence of the peptides selected from the unconstrained library flanked by two **C** residues, and was cloned into a phage. The addition of the **C** residues reduced phage binding by 100-fold in comparison with the unconstrained phage. In light of this result, McConnell *et al.* conclude that there is a possibility that, even if a library does contain the correct motif, panning may nonetheless fail to select it if constraint renders it unsuitable as a ligand for the material being tested. McConnell *et al.* go on to conclude that if the natural epitope is constrained, then an unconstrained library may either fail to recognise it at all or may reveal unrelated mimotopes.

In a review on the subject, Ladner (1995), highlights the usefulness of constraint for discontinuous epitopes and for receptor binding sites. He suggests that a constrained ligand will have a higher affinity for the site than a similar unconstrained one because, if the conformation of the unconstrained peptide can change on dissociation

from the site, there will be a tendency for dissociation to predominate over reassociation that is absent in the constrained version.

An interesting final point on this subject is provided by Kay *et al.* (1993), who, in testing a 38 amino acid unconstrained library, discovered strong evidence of selection against **C** residues in selected peptide sequences. More interestingly, of those that did contain **C**, the number with an odd number was less than expected, whilst those with even numbers were at the expected frequency. The explanation given for this phenomenon in the original paper and expanded by Kay and Hoess (1996), is that single **C**s may form disulphide bridges with **C** residues fused to other pIII molecules, or with those present in pIII itself (there are eight). Such bonds would, most probably disrupt the structure of pIII, thereby impairing infectivity by that phage.

PART 6 : SMITH PENTADECAMER LIBRARY

The third library tested was the Smith pentadecamer (15 amino acid insert) library. The library was obtained from Dr. Russell Thompson, Division of Virology, University of Glasgow, but was originally constructed by George Smith of the University of Missouri. As in the Luzzago library, inserts are fused to protein VIII (see page 68 for a discussion of the advantages and disadvantages of fusions into this protein), but unlike the libraries tested previously, it is a type 88 fusion.

Type 88 fusion phage

The type 88 strategy represents an alternative way of overcoming the limitation on insert size that is encountered when incorporating fusions into protein VIII. Originally postulated by George Smith (1993), it has apparently failed to attract significant use; reviews that report type 88 (or its gene III equivalent, type 33) as an option always refer back to the Smith paper, and Burritt *et al.* (1996) survey twenty-three different libraries, of which none are type 88. For this reason, little information about the efficacy of this type of library is available, other than unpublished data from Russell Thompson who made it available for this project.

One paper reports the use of type 33 and 88 systems: Corey *et al.* (1993) fused trypsin to either gene III or gene VIII that was carried by M13mp18, a vector similar to the pEMBL phagemid vector used in the production of the Luzzago library (see page 60). The gene III fusions were tested by biopanning on a series of ligands. Although successful, the principle was not extended as far as the production of a library.

Smith pentadecamer library

The procedure for creating the pentadecamer library used in this work is reported in 'Cloning of fUSE vectors (Edition of February 10th, 1992)', the unpublished manual that was supplied by George Smith and was extensively used in the hexamer library work.

This type 88 phage, f88-4, carries the wild-type fd-tet genome, with all genes intact. In addition however, *Xho*I and *Nhe*I sites were created at either end of a 330-base pair non-coding region of fd-tet (bases 5649 to 5978 inclusive), and a second, synthetic gene VIII was inserted at these sites. This insert is shown in Figure 13. Although the wild-type pVIII amino acid sequence is encoded, alternative codons are used with high frequency. Being able to discriminate between the two forms of gene VIII in the genome was of particular value for sequencing - without the variation, any primer against the recombinant gene VIII would also anneal to the wild-type sequence, which would have led to the appearance of double signals on the autoradiograph. Figure 14 compares the wild-type and recombinant base sequences.

Modification of the wild-type sequence CTG TCT TTC GCT to CTA AGC TTT GCC, which still codes for **LSFA**, served to introduce the *Hin*dIII restriction site (underlined). To introduce the *Pst*I site, an additional codon (CCT coding for **P**) was introduced. Along with the modification of GCT to GCA (both code for **A**), the restriction site (CTGCAG) was created. The introduction of these sites allowed insertion of the randomised 45 bases (*ie.* 15 amino acids). The insert therefore had the sequence AGC TTT GCC $(NNN)_{15}$ CCT GCA . . . coding for **S F A** $\mathbf{X_{15}}$ **P A ..** and, after cleaving by signal peptidase, the mature protein had the sequence

base 5648
of fd-tet
| *XhoI*
GC TCGAGCTTAC TCCCCA

tac promoter
TCCC CCTGTTGACA ATTAATCATC GGCTCGTATA AGTTGTGGAA TTGTGAGCGG

ATAACAATTT C

gene VIII translation initiation region
TTAATGGAA ACTTCC

gene VIII

```
ATG AAA AAG TCT TTA GTT CTT AAA GCA TCT GTT GCT GTT GCG ACT
 M   K   K   S   L   V   L   K   A   S   V   A   V   A   T

CTT GTT CCT ATG CTA·AGC TTT GCC|AAC GTC|CCT GCA·GAA GGT GAT
 L   V   P   M   L   S   F   A ^ N   V   P   A   E   G   D

GAC CCG GCT AAA GCT GCT TTT GAC TCT CTT CAG GCT TCT GCT ACT
 D   P   A   K   A   A   F   D   S   L   Q   A   S   A   T

GAA TAC ATC GGC TAC GCT TGG GCT ATG GTG GTT GTT ATC GTT GGT
 E   Y   I   G   Y   A   W   A   M   V   V   V   I   V   G

GCT ACT ATT GGC ATC AAA CTT TTC AAA AAA TTC ACT TCT AAA GCG
 A   T   I   G   I   K   L   F   K   K   F   T   S   K   A

TCT
 S
```

terminator region *NheI*
TAAT GAACTCAGAT ACCCAGCCCG CCTAATGAGCGGGCTTTTTT TAAGCTAGCTT

FIGURE 13 : Sequence of the second copy of gene VIII in the type 88 vector, f88-4.

Gene VIII itself is shown as codons, with the amino acids marked underneath. The random pentadecamer was inserted between the *Hin*dIII and *Pst*I sites, replacing the amino acids N and V, the first two bases of the mature pVIII. The site for cleavage by signal peptidase is indicated by ^. The *tac* promoter and the terminator region are also shown.

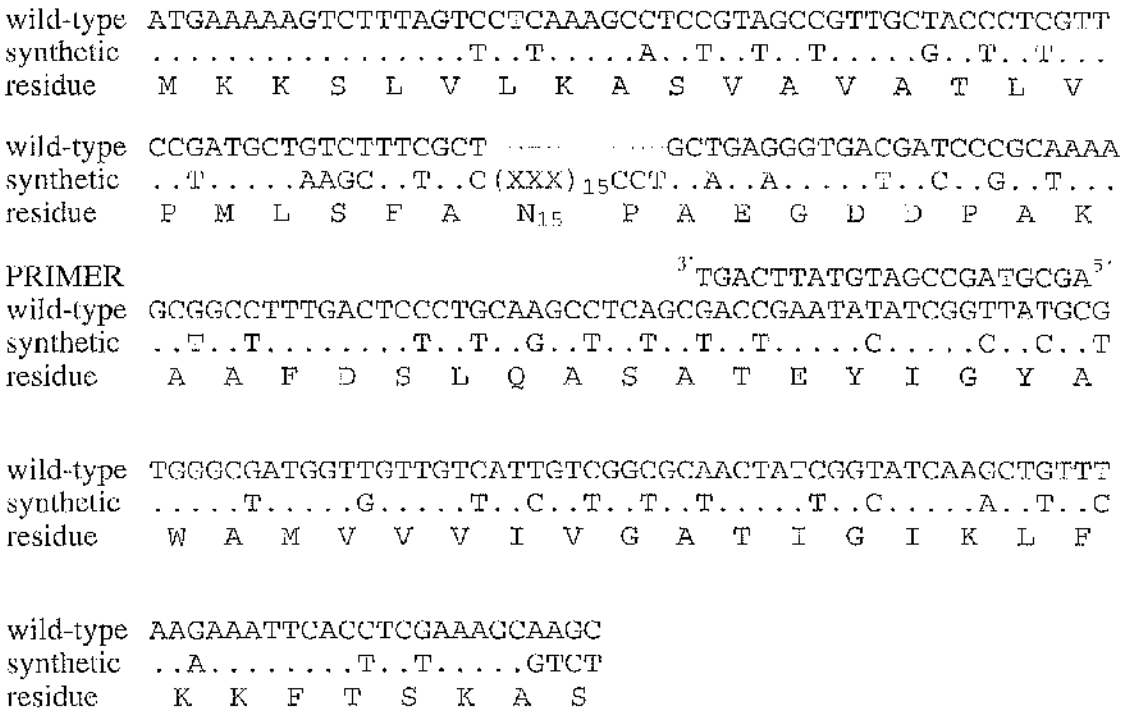


FIGURE 14 : Alignment of nucleic and amino acid sequences of wild-type and synthetic gene VIII from vector f88-4.

The 21-base primer synthesised for sequencing is also shown.

X_{15} P A E G D ... The synthetic gene VIII was placed under the control of the *tac* promoter. This strong promoter is a hybrid between the *lac* promoter that featured in the Luzzago library vector and the *trp* promoter that in *E. coli* drives transcription of enzymes involved in the biosynthesis of tryptophan. The *tac* promoter is stronger than either of these (*ie.* it can sustain a higher level of transcription) and, like *lac*, is induced by the addition of IPTG. The original library consisted of 2 x 10^9 clones and an amplification of a part of it was supplied for this work .

The result of creating a phage with two genes for protein VIII is that around 10% of the outer coat protein will carry the insert (*ie.* around 300 per virion) (George Smith in 'Cloning of fUSE vectors (Edition of February 10th, 1992)'). The presence of wild-type pVIII ensures that phage assembly and infectivity can proceed without hindrance by the large insert in the recombinant pVIII, whereas, in the absence of wild-type pVIII, inserts of greater than five to six amino acids cannot be easily packaged (see page 69).

MATERIALS AND METHODS

CULTURE AND MAINTENANCE OF STRAINS

Vibrio 60 pMMB68

This strain, which excretes LT-B, was kindly supplied by Professor T. Hirst, then at the University of Kent. The organism was cultured in Luria Bertani High Salt (LBHS) broth or in LBHS agar (see Appendix 1, page 182), both containing 100mg ml^{-1} ampicillin. Incubation was at 30°C, the broth cultures with shaking at 200 revolutions $minute^{-1}$. Stocks of the strain were stored at -70°C in LBHS broth containing 12.5 % glycerol.

Escherichia coli K91kan

This strain, which was used to amplify phage from both Smith libraries, was routinely cultured at 37°C in Luria Bertani (LB) broth or agar (see Appendix 1, page 182) containing 100µg ml^{-1} kanamycin. Broths were aerated as above. Stocks of the strain were stored at -70°C in LB broth containing 12.5 % glycerol.

Escherichia coli XL-1

This strain, which was used to amplify Luzzago library phage, was routinely cultured in LB broth or agar (see Appendix 1, page 182). Broths were aerated and stocks stored in the same manner as *E. coli* K91kan.

Escherichia coli DH5F'IQ

This strain, which is *rec*A deficient (see page 150), was routinely cultured as described for *E. coli* K91kan (above). It was obtained from Life Technologies, Paisley.

EXPRESSION, PURIFICATION AND CHARACTERISATION OF LT-B

Correlation of *Vibrio* 60 pMMB68 growth cycle with LT-B production

A flask of 50ml of LBHS broth was inoculated with 1ml of an overnight culture, and 1ml samples were withdrawn at hourly intervals. After 2h, the absorbance at 600 nm (A_{600}) having reached 0.1, IPTG (Sigma; Poole) was added to a final concentration of 100μM. After measuring the A_{600} of the 1ml samples, the supernate was clarified by centrifugation and tested for LT-B content by G_{M1} ELISA (see page 82), using an initial dilution of 1 in 64 .

Preparation of *Vibrio* 60 pMMB68 supernate

A 2 litre conical flask containing 500ml LBHS broth was inoculated with 10ml of an overnight culture of *Vibrio* 60 pMMB68 of A_{600} 1.0 - 1.2 and incubated with shaking at 30°C until the A_{600} was approximately 0.2 (around 2 to 4h post-inoculation). IPTG (as above) was added and incubation continued until 10h post-inoculation. The cells were separated from the supernate by centrifugation (8000 *g* for 20 minutes at 4°C). The supernate was cleared by further centrifugation (31000 *g* for 20 minutes at 4°C), followed by passage through a 0.45μm Acrodisc filter (Gelman Sciences; Northampton) and was stored at -20°C in 100ml batches until required.

Purification of LT-B from *Vibrio* 60 pMMB68 supernate

This was achieved using the immobilised galactose affinity chromatography method of Uesaka *et al.* (1994) discussed on page 48. A Pharmacia C16 column containing 15ml of Immobilised Galactose beads (Pierce and Warriner; Chester) was equilibrated with 30ml of 0.1M sodium phosphate buffer, pH 7.2. Individual 100ml batches of the cleared *Vibrio* 60 pMMB68 supernate, prepared as above, were run onto the column by gravity from a 30 cm head. The column was washed with a further 30ml of the phosphate buffer. All eluates thus far were run to waste. Elution of the LT-B was

achieved by running 30ml of 0.5M D-galactose in phosphate buffer through the column and collecting fractions of approximately 1ml.

Concentration of LT-B samples

Following analysis (see below), fractions from the affinity purification column containing pure LT-B were pooled, then concentrated by pouring into dialysis tubing, sealing and covering in dry polyethylene glycol with average relative mass of 10000 (PEG 10000; Sigma; Poole). This was left at 4°C until the volume was sufficiently reduced. The tube was then rapidly rinsed with water to remove PEG from the surface.

ANALYSIS OF PURIFIED LT-B

Gel electrophoresis

The electrophoretic properties of fractions from the immobilised galactose column were firstly established using SDS PAGE, adapting the method of Laemmli (1970). Concentrated, purified material, (100µl) was mixed with an equal volume of solubilising buffer (recipes and mixtures used are listed in Appendix 2, page 183). The sample was then divided into aliquots, one aliquot being boiled for 5 minutes and the other being left at room temperature. Then, 30µl of boiled and unboiled sample were applied to separate lanes of a 12.5 % mini-gel (10 cm x 6 cm) that was run for approximately 75 minutes at 150 Volts using an Atto gel tank and a Pharmacia ECPS 3000/150 power pack.

Gel staining

The gel was stained by the Coomassie Blue staining method (Appendix 3, page 185).

Western blotting

The cross reaction between the protein eluted from the affinity column and an anti-LT-B monoclonal antibody 118.87 (kindly supplied by Prof. T. Hirst, then of the University of Kent) was tested by Western blotting. Two lanes containing unboiled sample were run on an SDS-PAGE gel (as described above). The gel was then

divided between the lanes, and one half was stained by the Coomassie method (See Appendix 3, page 185 for method). The other portion was soaked, along with a similarly-sized piece of Hybond C nitrocellulose membrane (Amersham Life Science, Amersham, Bucks.) and two sheets of Whatman 3MM chromatography paper, in transfer buffer (for recipe see Appendix 4 page 186) for 30 minutes at room temperature. Then, the gel was sandwiched alongside the nitrocellulose membrane between the sheets of 3MM on the negative plate of a BioRad Transblot Semi-dry Electrophoretic Transfer Cell (Biorad, Hemel Hempstead), with the gel towards the plate. After blotting for 75 minutes at 15 volts, the membrane was probed with 5 ml of anti - LT-B antibody at a dilution of 1 in 400 in PBS-TWEEN. After 6 hours incubation at room temperature and a further three washes with PBS-TWEEN, 1 ml of HRP - conjugated goat anti-mouse IgG, at a dilution of 1 in 500 in PBS-TWEEN, was added. After 2 hours incubation at room temperature and three 10 minute washes in PBS, the membrane was developed. The developer consisted of 9.8 ml of a filtered (Whatman no. 1) solution of 0.05 % (w/v) diamino benzidine (Sigma, Poole) in PBS. Just before use, 0.2 ml of 1 % (w/v) cobalt chloride and 10 µl of hydrogen peroxide were added, and 5 ml was applied to the membrane. After 2 - 3 minutes the developer was poured off and the membrane was allowed to dry.

G_{M1} ELISA

The G_{M1} ELISA method of Svennerholm and Holmgren (1978) was adapted for use with microtitre plates. Each well of an Immunolon 96 well microtiter plate (Life Technologies; Paisley) was coated by the addition of 250µl of 1.5µM ganglioside G_{M1} (Sigma; Poole) in PBS and incubated, at room temperature, overnight. Coated plates were stored for no more than one month at 4°C.

Before being used in an assay, plates were washed three times, for three minutes each time, with PBS. Non-specific binding sites were blocked by 30 minutes incubation, at 37°C, with 200µl of PBS - BSA (for recipes see Appendix 4, page 186). Fresh PBS - BSA (316µl) was then added to each well. Then 84µl (40 µg) of purified LT-B was

added to the first well and a two-fold dilution series was generated by transfer of 200µl. The plate was incubated for four hours at room temperature after which it was washed three times, for three minutes each time, in PBS - TWEEN.

An anti - LT-B mouse monoclonal antibody 118.87, supplied by Professor T. Hirst, was diluted to 1 in 800 in PBS-TWEEN containing 0.1 % (w/v) BSA and 150µl was added to each well. After 90 minutes incubation at 37°C and a further three washes with PBS - TWEEN, 150µl of HRP - conjugated goat anti-mouse-IgG (SAPU, Carluke), at a dilution of 1 in 200 in PBS-TWEEN plus BSA, was added and incubated at room temperature for 90 minutes. The plates were washed three times, for three minutes each time, with PBS and then developed. The developing solution consisted of 20mg of the peroxidase substrate O-phenylene diamine (OPD: Sigma; Poole) in 20ml 0.1M sodium citrate, pH 4.2, to which 8µl hydrogen peroxide was added prior to use. A positive result with this substrate is a yellow coloration, the absorbance of which can be measured spectrophotometrically. Each well received 100µl of this reagent. After 15 minutes incubation at room temperature A_{450} was recorded using an ELISA plate reader.

Protein estimation of purified LT-B

Markwell *et al.* (1978) method

This method of protein estimation is based on copper labelling of the protein, as discussed on 50, and is reported to be unaffected by high sugar concentrations.

Galactose stability

Stock solutions were Reagent A: 2 % Na_2CO_3, 0.4 % NaOH, 0.16 % Na tartrate, 1 % SDS, and Reagent B: 4 % $CuSO_4$. Reagent C was freshly prepared, and consisted of 100 parts Reagent A plus 1 part Reagent B. Bovine serum albumin (BSA: Sigma; Poole) was prepared to a concentration of 120µg ml^{-1}. Aliquots of 0.5ml of this were mixed with an equal volume of galactose solution at concentrations ranging from 200 to 1000mM (to give final concentrations of 100 to 500mM galactose) and with distilled water (i.e. 0mM galactose). To these, 3ml of reagent C was added. After 10

minutes incubation at room temperature, 0.3ml of a 1 in 2 dilution of Folin and Ciocalteau Phenol reagent in distilled water (Sigma; Poole) was added. After vortexing, these were allowed to stand at room temperature for 45 minutes, after which absorbance was read at 660nm against a water (0mM galactose) control.

Bradford (1976) method

This method of protein estimation is based on recording the shift in the absorbance of Coomassie blue when bound to protein, as discussed on page 50.

Galactose stability

The reagent used was prepared by dissolving 50mg of Coomassie Brilliant Blue G-250 in 50ml of 95 % ethanol and then adding 100ml of 85 % phosphoric acid. Distilled water was added to yield a final volume of 1 litre. BSA was prepared to a final concentration of 60µg ml^{-1} in the same concentrations of galactose that were used with the Markwell method. Each 100µl aliquot of BSA/galactose solution received 5ml of reagent. After 2 minutes incubation at room temperature A_{595} was read against a reagent blank containing no protein.

Protein estimations

Bradford reagent, prepared as above, was used to estimate the protein content of all biotinylated LT-B ($LT\text{-}B^{biotin}$) samples. In each case, 5ml of reagent was added to 100µl of sample, and to 100µl solutions of BSA containing 0 to 50µg of protein. After 2 minutes incubation at room temperature A_{595} was read against a reagent blank. Protein content was estimated with reference to the BSA solutions.

Alternative source of purified LT-B

A sample of LT-B was kindly supplied by Professor T. Hirst, then at the University of Kent. Like that described above, it was produced by *Vibrio*60, however, it was purified by the method of Amin and Hirst (1994). This sample (hereafter called 'Hirst LT-B') was used in the second and third biopanning experiments with the Smith pentadecamer library, described on page 96.

BIOTINYLATION OF PURIFIED LT-B

Both the LT-B produced for this project and the Hirst LT-B were biotinylated to facilitate immobilisation *via* the bioitin - streptavidin bridge described on page 50. The protocol supplied with the Immunoprobe Biotinylation Kit (Kit no BK-101: Sigma; Poole) was followed exactly, except that in both cases, 200µl of BAC -SulfoNHS solution was added per ml of concentrated LT-B. Separation of the LT-B^{biotin} from unreacted biotin was performed as recommended in the kit instructions using the Sephadex G-25M column provided. The LT-B^{biotin} was stored at -20°C in 200µl aliquots.

Degree of biotinylation

This was calculated using the avidin-HABA assay, performed exactly as detailed in, and using the reagents supplied with, the Immunoprobe Biotinylation Kit described above. In this assay, the LT-B^{biotin} was degraded by pronase and exposed to avidin-coupled 4'-hydroxyazobenzene-2-carboxylic acid (HABA). Displacement of the avidin by biotin results in a colour change, detectable by spectrophotometry at 500nm, which is proportional to the amount of biotin present. Subtraction of the value given by LT-B^{biotin} from that obtained with a HABA control gave a corrected A_{500} value from which the biotin concentration could be calculated. The molar extinction coefficient of biotin is 34000.

The manufacturer recommends that the molar ratio of biotin to protein should be in the range three to five. The LT-B prepared for this project and biotinylated shall hereafter be referred to as 'stock LT-B^{biotin}', thus distinguishing it from 'Hirst LT-B^{biotin}'.

G_{M1} ELISA

G_{M1} ELISA was used to confirm biotinylation of the stock LT-B and to establish that biotinylation had not diminished its G_{M1}-binding capability . Stock LT-B^{biotin} was added to the plate at an initial concentration of 1 in 20 (10µl added to 190µl of PBS-

BSA) and doubling dilution effected by the transfer of 100µl. Two such series were prepared.

The first was probed with the anti-LT-B monoclonal antibody 118.87. The method followed was exactly as described on page 82, except that 100 µl volumes were added throughout. A dilution series to which no LT-B had been added served as a control.

The second series was probed with 100µl of the HRP-labelled streptavidin, 'Extravidin' (Sigma; Poole), at a dilution of 1 in 500 in PBS plus BSA. After 4h incubation at room temperature, these wells were washed three times, for three minutes each time, with PBS and then developed using O-phenylene diamine (OPD: Sigma; Poole) as for the anti-LT-B antibody. Again, a series to which no LT-B had been added served as a control.

SOURCE OF PHAGE LIBRARIES AND PROTOCOLS

Smith hexamer library

As well as the hexamer library itself, George Smith of the University of Missouri was kind enough to supply the bacterial strain used for propagation of library phage and a manual, 'Cloning of fUSE vectors (Edition of February 10th, 1992)' which served as a valuable reference for methods associated with biopanning.

Luzzago library

This was kindly supplied by Alfredo Nicosia, Instituto di Ricerche di Biologia Molecolare P. Angelletti, Rome. Methods used with the nonamer library were adapted from Luzzago et al. (1993) and Felici et al., (1991; 1993), with reference to the methods used for the Smith library (see above).

Smith pentadecamer library

Russell Thompson, of the Division of Virology, University of Glasgow kindly supplied this library, along with a biopanning protocol that is based on, but not identical to, George Smith's.

AMPLIFICATION OF LIBRARIES

Smith hexamer library

The sample provided was of insufficient volume for the work planned and therefore required amplification. *Escherichia coli* K91kan was grown overnight in LB broth, 1ml added to 100ml of Terrific Broth and the flask shaken at 250 revolutions minute^{-1} at 37°C until the A_{600} was 2.0. Shaking was then slowed to 30 revolutions minute^{-1} for 5 minutes. All of the Smith hexamer library (approximately 6µl) was added, and the slow shaking maintained for 15 minutes. The entire content of this flask was then added to 1 litre of pre-warmed LB broth containing 0.2µg ml^{-1} tetracycline, and incubation at 37°C, with 250 revolutions minute^{-1} shaking, was continued for 35 minutes. After the addition of tetracycline to a final concentration of 20µg ml^{-1}, the incubation continued for 16h.

Cells were removed by two 10 minute spins at 20000*g* and solid polyethylene glycol, 10000 Da (Sigma, Poole) was added to the supernate to 15% (w/v). Precipitation was left to proceed, on ice, for 16h. The precipitate was removed from suspension by centrifugation at 20000*g* for 40 minutes at 2°C.

The phage pellet was dispersed with gentle agitation for 30 minutes, in 30ml of TBS and then spun for 10 minutes at 20000*g* to remove insoluble material. The supernate was mixed by 100 inversions with 4.5ml (*ie.* 0.15 volumes) of PEG / NaCl (see Appendix 4, page 186) then left on ice for 1 hour. Precipitated phage were harvested by centrifugation for 15 minutes at 20000*g*, then re-dispersed by gentle agitation for 1 hour (to soften the pellet) followed by vortexing for 2 minutes. Insoluble material was removed by centrifugation at 20000*g* for a further 10 minutes. The supernate was mixed with 0.15 volumes (1.5ml) of PEG / NaCl then incubated on ice overnight. The precipitated phage were harvested as before and re-suspended in 1ml of TBS, by 1h gentle agitation and 2 minutes vortexing, and were then stored at -20°C as the 'Amplified Library Stock'.

Luzzago and Smith pentadecamer libraries

Sufficient library having been provided, amplification was unnecessary.

PHAGE INFECTIVITY

Smith hexamer library

The infectivity cycle of *E. coli* K91kan with Smith hexamer library phage was investigated thus : from an overnight culture of K91kan in LB broth, 200µl was used to inoculate 20ml of Terrific Broth (see Appendix 1, page 182). The flask was incubated with shaking at 250 revolutions minute^{-1}, at 37°C. At 30 minute intervals, from 6h onwards, aliquots were removed and the A_{600} was measured. At 30 minute intervals from 6.5h onwards, a second aliquot was removed, left static for 5 minutes to allow for regeneration of pili, then mixed with 100µl of such a dilution of the 'Amplified Library Stock', that 1 x 10^6 phage were added (phage numbers having been established by the method reported on page 97). After incubating at 37°C for 30 minutes to allow infection of the K91kan cells, 1ml of LB broth containing 0.2µgml^{-1} tetracycline was added. After shaking at 37°C for one hour to allow expression of the tetracycline resistance genes, 100µl was spread onto LB agar containing 100µg ml^{-1} kanamycin and 40µg ml^{-1} tetracycline. Plates were incubated overnight at 37°C and colonies were counted.

Luzzago library

The phage library was serially diluted ten-fold to 10^{-11} and 10µl of each dilution added to 100µl of a four hour culture of *E. coli* XL-1 in LB broth. After incubation at 37°C for one hour, all 110µl was spread on LB agar containing 100µg ml^{-1} ampicillin.

As a control, uninfected cells were plated directly onto LB agar plus ampicillin.

BIOPANNING

Smith hexamer library

First round of biopanning

Prior to commencement of the work, the streptavidin concentration and diluting buffer, as well as the brand of plastic tubes used had been optimised in an ELISA-style assay.

Streptavidin coating of tubes

Streptavidin (1ml of a 0.8mg ml^{-1} solution in 0.05M sodium phosphate buffer, pH 8.0), was added to 3 Maxisorp tubes (Life Technologies; Paisley), left for 6h at 4°C then discarded and replaced with 2ml of a blocking solution consisting of 5mg ml^{-1} BSA, 0.1mg ml^{-1} streptavidin and 0.02 % NaN_3 in 0.1M $NaHCO_3$. After reaction for 1 hour at 4°C, the tubes were washed 6 times, for 1 minute each time in TBS-TWEEN, then upturned and 'slapped' down onto dry tissue to remove all residual fluid.

LT-B coating of tubes

Immediately after washing, 1ml of a 1 in 500 dilution of the stock $LT\text{-}B^{biotin}$ solution in TBS-TWEEN, was added to the tubes, which were left overnight at 4°C. Then, 10µl of 10mM biotin solution was added and incubation at 4°C continued for a further hour before washing 6 times as before.

Panning and elution

Each tube received 1ml of TBS-TWEEN and 100µl 'Amplified Library Stock'. After 4h at 4°C, the tubes were emptied, drained and washed ten times as before. Bound phage were eluted from the first tube with 1ml of acid elution buffer (1mg ml^{-1} bovine serum albumin in 0.1N glycine-HCl, pH 2.2). This was removed after 30 minutes and neutralised with 187.5µl of 1M Tris, pH 9.1. Elution from the second tube was with 7µl of 1500µM G_{M1}, diluted in 1ml of TBS-TWEEN, whilst the third tube received 1ml of TBS alone. In both cases, the elution agent was left in the tube for 1 hour.

Portions of all eluates were immediately amplified (see below) with the remainder being frozen for further use.

Amplification and purification of phage

A culture of *E. coli* K91kan was prepared by inoculating 20ml of LB broth and incubating, without shaking, for 16h at 37°C. A 20ml flask of pre-warmed Terrific Broth (see Appendix 1, page 182) was then inoculated with 100µl of this culture, and incubated with shaking (250 revolutions minute^{-1}) at 37°C until A_{600} was between 1.25 and 2.5. Shaking was slowed to 30 revolutions minute^{-1} for at least 5 minutes and the culture was used for phage amplification within 1 hour.

To amplify, 100µl of eluted phage were added to 100µl of the Terrific Broth culture, mixed gently, and allowed to stand at room temperature for 30 minutes. This was added to 20ml of pre-warmed LB broth containing 0.2µg ml^{-1} tetracycline, and shaken at 37°C and 250 revolutions minute^{-1} for 60 minutes. Tetracycline was added to a final concentration of 20µg ml^{-1} and the flask left to incubate overnight. Then, the culture was poured into a 50ml Oak Ridge tube and the supernate clarified by centrifugation at 20000*g* for firstly 5, and then 10 minutes before being added to 3ml of PEG / NaCl. After mixing by 100 inversions, this was left on ice overnight to precipitate the phage. These were pelleted by centrifugation at 20000*g* for 15 minutes, then resuspended in 1ml of TBS, vortexed, and centrifuged for 1 minute at 25000*g*. The supernate was added to 150µl of PEG / NaCl and left on ice for 1 hour. After microfuging for 10 minutes, pelleted phage were resuspended in 200µl of TBS, clarified once by brief centrifugation and the supernate (i.e. first round amplified phage) stored at -20°C until required.

Second and third rounds of biopanning

Reaction of phage with LT-B

100µl of amplified phage, from the previous round, was mixed with 2µl of stock LT-B^{biotin} in microfuge tubes and left to react overnight at 4°C

Panning on streptavidin and elution of bound phage

Three Maxisorp tubes were coated with streptavidin and blocked in exactly the same manner as for the first round of biopanning. The overnight phage/LT-B mixes were then added to blocking solution in the appropriate tube and left at room temperature for one hour. This protocol, termed PL + S (phage and ligate plus streptavidin) contrasts with the P + LS strategy used in Round one. Elution was carried out exactly as described for round one, as was amplification and purification of eluted phage.

Fourth round of biopanning

This was carried out in an identical manner to rounds two and three through to the elution stage. However, eluted phage were not amplified; they were used to infect cells which were plated out to single colonies, each colony comprising identical cells arising from infection with a single phage. To obtain single colonies, 100µl of 10^{-1} and 10^{-2} dilutions of each eluate were added to 100µl of *E. coli* K91kan Terrific Broth culture and left for 30 minutes at room temperature. This mixture was added to 1ml of pre-warmed LB broth with 0.2µg of tetracycline, and shaken at 37°C and 250 revolutions minute^{-1} for 60 minutes. Then, 100µl was spread onto LB agar containing 100µg ml^{-1} kanamycin and 40µg ml^{-1} tetracycline. Plates were incubated at 37°C overnight and were stored at 4°C until required for sequencing.

Purification of phage

Single colonies arising from round four eluates were picked at random, inoculated into 1.7ml of LB broth and incubated overnight at 37°C , with shaking at 250 revolutions minute^{-1}. Culture supernates were then clarified by microfuging for 5 minutes at 25000*g* and 1ml was added to 150µl of PEG / NaCl. After mixing by 100 inversions the tubes were left on ice overnight. Precipitated phage were spun down by 15 minutes microfuging at 25000*g* at 4°C and the pellets dispersed in 500µl of TBS and stored at -20°C .

Luzzago library

First round of biopanning

Three Maxisorb tubes were coated with streptavidin and then with stock LT-B^{biotin} exactly as described for the Smith hexamer library.

Panning and elution

To each tube was added 400µl TBS-TWEEN, 4µl 10mM biotin and 2µl phage from the Luzzago library. The tubes were incubated at 4°C for 4h then washed 10 times with 2ml of TBS-TWEEN and eluted using 500µl of Acid Elution buffer, G_{M1} and TBS respectively (see page 89). Elution proceeded in each case, for 30 minutes at room temperature, after which, 93.5µl of 1M Tris, pH 9.1 was added to neutralise the acid eluate.

Amplification and purification of phage

To 450µl of each eluate was added 450µl of an overnight culture of *E. coli* XL-1. After mixing, the tubes were left at 37°C for 1 hour. Then, 100µl was spread onto LB agar containing 50µg ml^{-1} ampicillin, nine plates per eluate, and the 27 plates were incubated overnight at 37°C.

Following incubation, 1ml of LB broth containing 50µg ml^{-1} ampicillin was added to each plate and the cells suspended. The suspensions from each elution set were pooled and added to 250ml flasks containing 100ml of LB broth with 50µg ml^{-1} ampicillin, to an A_{600} of 0.05. The flasks were shaken at 37°C until A_{600} had risen to 0.2 then 1.6 x 10^{11} M13K07 Helper Phage (Promega; Southampton) were added and incubation was continued overnight.

Supernates were clarified by centrifugation at 20000*g* for 20 minutes and phage were precipitated overnight with 0.15 volumes of PEG / NaCl (see Appendix 4, page 186) at 4°C. Phage were harvested by 40 minutes centrifugation as above, and resuspended in 10ml of TBS. A further 1.5ml of PEG/ NaCl was added and left on ice at 4°C for 1 hour. The now-purified phage were harvested by centrifugation as above for 40

minutes, and resuspended in 500µl TBS. This final suspension constituted the First Round Amplified Eluate.

Second round of biopanning

Reaction, panning and elution

In this case, 200µl of each First Round Amplified Eluate (acid, G_{M1} and TBS eluted) were mixed with 2µl of stock LT-B^{biotin}, and were left at 4°C overnight, before being increased to 1ml by the addition of TBS. This was added to tubes coated with streptavidin (as before), which were left 1 hour at room temperature, then washed 10 times with 1ml each time of TBS-TWEEN. Elution was carried out as in round 1 (see above), except that 1ml of eluting agent was added. The acid eluate was thus neutralised with 187.5µl of 1M Tris, pH 9.1.

Amplification and purification of phage

The processes of plating, harvesting, amplification and helper phage addition were performed exactly as described for round 1 (above), as was the PEG/NaCl purification of the resultant Second Round Amplified phage.

Third Round of biopanning

Reaction, panning and elution

These stages were performed as in round 2, with only minor exceptions.

To increase the selectivity of the process, 200µl of each second round amplified phage was added to only 0.5µl of stock LT-B^{biotin}. After overnight incubation at 4°C, this was increased to a volume of 500µl by the addition of TBS and was added to tubes coated with streptavidin in the previously described manner. Incubation, washing and elution were carried out as before, with 500µl of elutant being added.

Isolation of single clones

To 100µl of each of the three eluates was added 100µl of an overnight culture of *E. coli* XL-1. After gentle mixing and 1h incubation at 37°C, ten 20µl aliquots were spread onto LB agar containing 50µg ml^{-1} ampicillin (i.e. 30 plates in all) and incubated at 37°C overnight. Each single colony was a clone of a single phage-infected cell.

From each set of plates (i.e. acid-, G_{M1}- and TBS-eluted) twenty randomly selected colonies were picked onto fresh agar (for archiving purposes) and were inoculated into 2ml LB broths that were incubated, with shaking, at 37°C overnight.

Smith pentadecamer library

The Smith pentadecamer library was the most extensively studied in this project, being used in three separate three-round panning experiments.

First panning experiment

Biopanning was carried out in Nunclon Delta plates (Life Technologies; Paisley) that consist of six, 50mm wells of 1cm depth, results obtained from a series of experiments with a variety of proteins in the Division of Virology having indicated that binding was satisfactory for the purpose of biopanning.

First round of biopanning

Streptavidin coating of wells

Streptavidin (20µg in 2ml of 0.1M unbuffered $NaHCO_3$) was added to one well of a Delta plate and was agitated gently, overnight at 4°C. The solution was then discarded and replaced with 5ml of a blocking solution consisting of 5mg ml^{-1} BSA, and 0.1mg ml^{-1} streptavidin in 0.1M $NaHCO_3$. This solution was gently agitated for two hours at room temperature. The well was then washed 3 times, for 30 seconds each time, with 5ml TBS-TWEEN, being 'slapped' onto dry tissue between washes to remove residual liquid.

Reaction, panning and elution

Library phage (30µl) were mixed with 10µl of stock LT-B^{biotin} in a microfuge tube and left overnight at 4°C. The reaction mix was diluted by the addition of 1ml of TBS-TWEEN and added to the streptavidin-coated well, which was agitated at room temperature for 20 minutes. Unreacted phage were discarded and the plate washed ten times, for 1 minute each time, with 1ml of TBS-TWEEN. Elution, which was by acid only, was effected using 500µl of Smith acid elution buffer (see page 89) containing 0.1mg ml^{-1} phenol red, which acted as an indicator. No more than 2 minutes after the

addition of the elution buffer, 90 to 100µl of 1M Tris/HCl, pH 9.1 was added, a change in the colour of the indicator from yellow to rosé, indicating neutralisation. Eluates were retained at -20°C.

Amplification and purification of phage

This was carried out exactly as for the Smith hexamer library, with one important variation. Having infected an *E. coli* K91kan culture with Round One phage, and sub-cultured into LB containing tetracycline as before, IPTG was added to a final concentration of 1mM. This served to induce, *via* the *tac* promoter, production of the synthetic pVIII that carries the fusion. Subsequent incubation and processing then followed an identical protocol to the Smith hexamer work.

Second round of biopanning

Reaction, panning and elution

In this round, 30µl of amplified phage from the previous round was mixed with only 8µl of LT-B^{biotin} in a microfuge tube which was then left overnight at 4°C. The overnight phage/LT-B mix was processed as in the first round (see above). The phage were also titred as described below in 'Enumeration of eluted phage'.

Third round of biopanning

Once again the amount of LT-B^{biotin} reacted with the phage was reduced, to 6µl that was added to 30µl of the second round phage. Although the eluted phage were not to be amplified, it was still necessary to mix them with K91kan cells, and to add this mixture to a 5ml volume of LB broth plus tetracycline, in order that the titre could be assayed by the same method as that previously employed.

Purification of single phage

When titering the third round eluates, an LB (plus tetracycline) agar plate containing single clones was produced. Colonies, picked from this plate, were processed as described on page 98.

Second panning experiment

A further three parallel three-round biopanning experiments were carried out using the Smith pentadecamer library and Hirst LT-B^{biotin}. In this case, phage-infected bacteria were assayed by a 'colony-lift' assay (for method, see page 102).

In the first round, 20 µl of the phage library in 5 ml TBS were added to wells of a Delta plate coated with streptavidin bound LT-B (prepared by adding 10 µl of LT-B^{biotin} to a plate coated with streptavidin and blocked exactly as described on page 94); with blocked streptavidin only (again prepared exactly as described on page 94); and with blocked plastic (5 ml of 25% BSA in TBS were added to an empty well, left at room temperature for 6 hours and then washed out with TBS).

Elution with acid elution buffer, amplification and purification of phage was carried out as before (see page 94). Enumeration was by direct counting of colonies as described below (page 98). Two further rounds of biopanning were carried out using an identical method, with phage eluted from each surface in one round being exposed to that same surface in the subsequent round. The plates by which the third round phage were enumerated were retained for the 'colony-lift' binding assay described below.

Third panning experiment

In order to eliminate the possibility of streptavidin-binders predominating in eluted phage, three panning rounds were carried out in which Hirst LT-B was immobilised using Protein G and an anti-cholera toxin antibody.

To one well of a 96-well microtitre plate was added 200µl of 5µg per ml Protein G (Sigma, Poole) in unbuffered 0.1M sodium bicarbonate. After overnight incubation, the well was blocked by 6 hours exposure to 200µl of a 1% solution of BSA in TBS. Then, the well was washed thoroughly with TBS-TWEEN and 200µl of a 1 in 100 dilution of anti-cholera toxin antibody (Sigma, Poole) in TBS was added. This was allowed to react for 1 hour, before the well was again washed with TBS-TWEEN.

At this point a mixture consisting of 2µl of 10µg per ml LTB and 30µl of library phage that had been allowed to react for 5 hours before dilution with 168µl of TBS

was introduced to the well. After 90 minutes incubation at room temperature, the well was emptied and was washed with TBS-TWEEN.

Elution, amplification and purification of binding phage was accomplished exactly as described above.

At this stage, in an attempt to reduce the number of phage that may specifically bind to the immobilising agent, 200µl of the amplified first-round phage were introduced to a well coated, as above, with Protein G and anti-cholera antibody. After 1 hours incubation at room temperature, the contents of the well were removed and retained as 'backchecked 1st round phage'. Those phage that had bound during incubation were eluted as before and were retained for testing as '1st round eluted phage'.

A further round of biopanning was then carried out using the backchecked phage from round 1. The protocol followed was exactly as described above, with the amplified selected phage again being backchecked and the phage eluted at backchecking again being retained for testing. A third round of panning then followed, producing further backchecked and eluted phage.

Numbers of phage produced at each stage were determined by the enumeration method described below.

ENUMERATION OF PHAGE

Smith hexamer library

The number of phage in the original library was estimated by preparing a series of ten-fold dilutions in TBS-G and mixing 20µl with a similar volume of a Terrific Broth culture of K91kan cells with an absorbance at 600nm of between 1.25 and 2.5. After 30 minutes incubation at room temperature, 1ml of LB containing 0.2µg of tetracycline was added and incubation, this time with shaking at 37°C, was continued for a further 45 minutes. After this, 100µl was spread onto LB agar containing 30µg ml^{-1} tetracycline and the plate incubated overnight at 37°C.

Luzzago library

The plates used to select phagemid-infected XL-1 cells prior to infection with helper phage (see above, page 92) also served the purpose of enumeration.

Smith pentadecamer library

First panning experiment

The titre of the phage in each round was determined at the amplification stage by removing 100µl of culture from the flask immediately after the addition of the 6.2µl of tetracycline but prior to overnight incubation. This was spread on to LB agar containing 30µg ml^{-1} tetracycline. The plate was then incubated overnight at 37°C, and the number of colonies was counted.

Second and third panning experiments

The number of eluted and amplified phage from all stages was estimated by following the method described above for the Smith Hexamer library.

PREPARATION OF SEQUENCING TEMPLATES

Smith hexamer and pentadecamer libraries

Purified DNA was produced from frozen phage stock by phenol / chloroform extraction.

To 200µl of phage was added 200µl of water-saturated phenol. After vortexing, layers were separated by microfuging at 25000*g* for 1 minute, the aqueous layer being withdrawn and added to 200µl of chloroform containing 4.16 % amyl alcohol. Vortexing and centrifugation were repeated as above and 150µl of the aqueous layer added to 250µl of TE buffer (See Appendix 4, page 186). Then, 40µl of 3M sodium acetate, pH 6.0, and 1ml of ethanol, were added. After vortexing, samples were incubated for at least 30 minutes at -20°C then microfuged for 30 minutes at 25000*g* at 2°C. Supernates were aspirated and the tubes re-spun briefly to collect supernate for further aspiration. Pellets were washed twice, for 8 minutes each time, with 1ml of 70

% ethanol and allowed to air dry before being resuspended in 6µl of deionised water, and stored in sealed tubes at -20°C.

Luzzago library

Phagemid DNA was prepared from the overnight cultures using the Wizard Minipreps DNA Purification System (Promega; Southampton), following the protocol supplied with the system. This process produced 50µl of double-stranded DNA, that was denatured to single-stranded form by alkaline precipitation, as detailed in the protocol leaflet accompanying the Sequenase 2.0 sequencing system. The final pellet was allowed to dry at room temperature for 10 minutes, then redissolved in 7µl deionised water.

SEQUENCING PRIMERS

Smith hexamer library

OLIGO 2303 (see Appendix 7, page 190 for base sequence) was prepared in the Division of Biochemistry and Molecular Biology at the University of Glasgow. For use, 360µl of stock was added to 40µl of 3M sodium acetate and 1.2ml of ice-cold ethanol added. This was vortexed, then left at -70°C for 15 minutes. Precipitated primer was harvested by centrifugation in a microfuge for 10 minutes at 4°C , the pellet washed with 1ml of ice-cold ethanol, dried at room temperature and finally resuspended in 100µl of water to a concentration of 61 pM. This was diluted to a working stock of 3 pM.

Luzzago library

The -40 primer provided with the USB Sequenase Version 2.0 DNA Sequencing Kit (Amersham Life Science; Amersham, Bucks.) has the sequence, GTTTTCCCAGTCACGAC, and recognises a site in the pUC19 portion of pEMBL19. It was provided at a concentration suitable for immediate use.

Smith pentadecamer library

The sequencing primer for this library, shown in Figure 14, page 76, was synthesised and supplied by Russell Thompson and did not need dilution for use.

SEQUENCING PROTOCOL

The sequencing reaction was carried out according to the protocol, and using the reagents provided with the USB Sequenase Version 2.0 DNA Sequencing Kit (Amersham Life Science; Amersham, Bucks.). The labelled dATP Redivue [^{35}S] dATP, was also supplied by Amersham.

Sequencing gels were run on a BRL Model S2 sequencing system (Life Technologies; Paisley) using 30 cm by 40 cm glass plates separated by a 0.2mm spacer. The acrylamide/bis solution (Scotlabs; Coatbridge) was 6% (w/v) acrylamide, 0.3% (w/v) bis/acrylamide, 7M urea, 1x TBE, 70ml being polymerised by the addition of 140µl of ammonium persulphate (25 %, w/v) and 70µl of tetramethyl ethylene diamine (TEMED).

The gel was placed on the sequencing apparatus and 1 litre of single strength TBE running buffer added (see Appendix 4, page 186). As a pre-run to test the integrity of the wells and to warm the gel, 2µl of stop solution was added to each well and the gel run at 85 Watts for 30 minutes.

Samples to be sequenced were heated for 2 minutes at 75°C and 2µl added per well. Gels were run, using a Gibco BRL 2500 power supply (Life Technologies; Paisley), at 85 Watts constant power, with a surface temperature of approximately 65°C, for 135 minutes. The gel was then transferred to Whatman 3MM chromatography paper and dried at 85°C for 2h on a Bio Rad 483 slab gel dryer. Dried gels were scanned with a Geiger counter to determine the level of radioactivity and then exposed to Fuji Medical X-ray film, type RX, for 2 to 10 days, depending on the signal obtained (typically, a gel giving around 80 to 100 counts per minute was left for 2 days, a gel giving less than 5 counts per minute for 10 days). Autoradiographs were developed in

Kodak LX24 X-ray developer for 4 min, rinsed with running tap water, then fixed in Kodak Unifix for 10 min.

BINDING ASSAY

Smith hexamer library

The extent to which the phage isolated by biopanning bound to LT-B^{biotin} was established in an ELISA-style assay.

Phage stocks used in sequencing were also used here with 100µl being added to one well, and 100µl of TBS being added to a second well, of a 96-well, flat-bottomed microtitre plate, which was left overnight at 4°C The phage were removed and reclaimed for further use, but the TBS was discarded. Wells were blocked at room temperature for 2h with 360µl of TBS-G (See Appendix 4, page 186), washed five times with TBS-TWEEN, then 80µl of a 1 in 500 dilution of stock LT-B^{biotin} in TBS was added. After overnight incubation at 4°C, both wells were washed five times with TBS-TWEEN, and 70µl of a 1 in 800 dilution of the monoclonal antibody 118.87 was added. Plates were washed and processed as described for G_{M1} ELISA, (page 82).

To confirm binding of phage to the plastic, selected wells were probed with a 1 in 250 dilution of a rabbit anti-fd phage antibody (Sigma, Poole) and processed as above with HRP-conjugated anti-rabbit reagent.

Smith pentadecamer library

Microtitre plate assay for phage from first panning experiment

Recognition of LT-B^{biotin} and streptavidin by phage selected in the first panning experiment was tested in a binding assay carried out in 'Immunosorb' 96-well microtitre plates (Life Technologies; Paisley).

The phage produced for DNA sequencing could not be used in the binding assays as they carried no recombinant pVIII. Thus, a second stock of around 10^{12} phage was produced from the archived cultures prepared above. Again, the culture was inoculated into 10ml of LB broth and shaken at 37°C, but this time, when the A_{600}

reached 0.2, IPTG was added to 1mM. Incubation continued overnight, and phage stocks were produced by PEG/NaCl precipitation as before.

To test for recognition of LT-B^{biotin}, 200µl of 0.8mg ml^{-1} streptavidin was added to eleven sets of four wells on the Immunosorb plate. After overnight incubation at 4°C, the streptavidin was removed and 400µl of the BSA/streptavidin blocker, as used in biopanning (see above), was added to all four wells of each set, and to a fifth well. After 90 minutes incubation at room temperature, this too was discarded. Then, a 1 in 500 dilution of stock LT-B^{biotin}, in TBS-TWEEN, was added to wells 2 and 4 of each set. The plate was then stored at 4°C for two hours, to allow the LT-B^{biotin} to bind. Unbound LT-B^{biotin} was washed out with 400µl of TBS-TWEEN by five washes, each of thirty seconds duration. Phage from a single clone were now added to wells 1, 2 and 5 of each set, 30µl per well. After overnight incubation at 4°C, and further washing, all five wells in each set were probed for binding phage. Firstly, 30µl of a 1 in 500 dilution, in TBS-TWEEN, of rabbit anti-fd phage antibody was added to each well. The plate was then incubated at 37°C for one hour. After washing with TBS-TWEEN (as before), 30µl of a 1 in 500 dilution, in TBS-TWEEN, of HRP-conjugated anti-rabbit antibody (Sigma; Poole) was added. Finally, after five more TBS-TWEEN washes, the plate was developed and read as described on page 82.

'Colony lift' assay for phage from second panning experiment

Colonies produced when enumerating the final round of biopanning against LT-B and blocked streptavidin in the second panning experiment were sub-cultured onto LB agar plates containing 30µg $ml.^{-1}$ tetracycline and 1mM IPTG that were incubated at 37°C overnight. The plates were then overlaid with 82 mm diameter, 0.45 µm cellulose nitrate membranes that were left in contact with the colonies on the agar for 5 minutes. The alignment of these test membranes on the plates was marked. This process was repeated using a second membrane for each plate, these second 'prints' serving as controls.

All membranes were soaked for 6 hours in 5 ml of a 25% BSA solution in PBS to block non-specific attachment and were then washed three times, for one minute each time, with TBS-TWEEN.

A fresh test membrane that had been exposed to the colonies infected with LT-B - panned phage was then transferred into 5 ml of TBS containing 10µl of 5 mg $ml.^{-1}$ non-biotinylated Hirst LT-B and was left for 2 hours at room temperature. After 3 washes with 10 ml TBS-TWEEN, 5 ml of a 1 in 500 dilution of anti-LT-B monoclonal antibody 118.87 was added. After a further two hours at room temperature, 5 ml of a 1 in 500 dilution of HRP-conjugated anti-mouse antibody in TBS-TWEEN was added. After two hours incubation, the membrane was developed (see below).

A second such test membrane was probed with 5 ml of 10 µg $ml.^{-1}$ streptavidin in 0.1M $NaHCO_3$. After 2 hours at room temperature and 3 washes with 10 ml TBS-TWEEN, 5 ml of a 1 in 500 dilution of biotinylated peroxidase (Sigma, Poole) was added. After a further two hours at room temperature, the membrane was developed (see below).

A test membrane from the colonies infected with streptavidin - panned phage was also exposed to streptavidin, according to the protocol described above.

All control membranes were transferred into 5 ml volumes of a 1 in 500 dilution of anti-fd phage antibody in TBS and after two hours at room temperature and washing with TBS-TWEEN, 5 ml volumes of a 1 in 500 dilution of HRP-conjugated anti-rabbit antibody were added. After a further two hours incubation, the membranes were developed.

To confirm that sufficient ligand was being added, samples of membrane spotted with 5 µl of 1.5 µM G_{M1} (for LT-B) or neat Hirst LT-B^{biotin} (for streptavidin) were processed along with the test membranes.

All test and control membranes were visualised by the addition of 5 ml of a developer that consisted of 3 ml of a 3 mg ml^{-1} solution (in methanol) of the HRP substrate 4-

chloronaphthol to which 47 ml of PBS was added. The visualisation reaction was stopped by washing the membrane with PBS.

Microtitre plate assay for phage from third panning experiment

'Backchecking', exposing the amplified, eluted phage from each round to a surface of Protein G immobilised anti-cholera toxin antibody, had been introduced to the panning process to remove phage that bore peptides that reacted with this surface from the phage pool, thus allowing phage bearing LT-B-reactive peptides to predominate. Ten examples of backchecked final round phage and ten of those eluted from the well used for backchecking, were tested in an ELISA-based binding assay.

Protein G could not be used in the binding assay ELISA because of its cross-reactivity with immunoglobulin, therefore, plates were coated with LT-B alone and with anti-cholera toxin antibody alone as follows.

Each phage was tested in eight wells of an 'Immunosorb' 96-well microtitre plate. Hirst LT-B was added to three wells at concentrations of 1, 10 and 100μg ml^{-1} of TBS respectively. Anti-cholera toxin antibody was added to a further three wells at concentrations of 1/10, 1/100 and 1/1000 in TBS respectively. The plate was then incubated at 4°C overnight. The content of the wells was discarded and 1% BSA in TBS was added to all six, and to one previously unused well and was left for 6 hours at room temperature to block non-specific sites. After thoroughly washing all used wells with TBS-TWEEN (which was used for washing throughout this procedure), 100μl of phage at a 1/100 dilution in TBS (approximately 10^8 phage) was added to all seven wells and to one more that was previously unused. One test set to which no phage were added served as a control. The plate was then left for one hour at room temperature and then washed.

As the anti-cholera toxin antibody was raised in rabbit, the enzyme system previously employed (rabbit derived anti-fd phage antibody followed by HRP-conjugated anti rabbit antibody) could not be used. Rather, 100μl of a biotinylated anti-fd phage antibody at a dilution of 1/1000 in TBS was added to all wells and allowed to react for one hour. After washing, the HRP-labelled streptavidin, 'Extravidin' was added,

again at 1/1000 in TBS and again reaction was for 1 hour at room temperature. Finally, the plate was developed and read as described on page 82.

In attempt to further characterise the phage under examination they were also tested by ELISA against a range of protein and detergent blockers.

The proteins tested were BSA, ovalbumin, gelatin and skimmed milk, each at a concentration of 1% in TBS. The detergents were Tween 20, Triton X-100 and Nonidet P40, each at a concentration of 0.5% in TBS.

Twenty-one series of test wells were prepared, each well in a series receiving 100µl of protein or detergent. After overnight incubation at room temperature, the wells that contained protein were washed three times with TBS-TWEEN.

The same twenty phage used above were tested, 100µl of each one being added to each well in a series. In addition, phage were also added to one previously unused well adjacent to the protein and the detergent wells. The twenty-first series received no phage and served as a control.

After 8 hours exposure, phage were removed from all wells. The protein and adjacent wells were washed with TBS-TWEEN, the detergent wells with the respective detergent.

Biotinylated anti-fd phage antibody, Extravidin and developer were then used as above except that washing was with the appropriate detergent. Finally the result was read at 492nm.

RESULTS

EXPRESSION, PURIFICATION AND CHARACTERISATION OF LT-B

Expression of LT-B

Correlation of *Vibrio* 60 pMMB68 growth cycle with LT-B production

The point in the *Vibrio* 60 pMMB68 growth cycle at which expression of LT-B was optimal was determined by performing G_{M1} ELISA (see page 82) on supernates of samples drawn at hourly intervals from a flask of growing cells. The titre was arbitrarily taken as the $\log_2$ dilution which produced an absorbance of 0.2. Figure 15 shows these titres superimposed onto the *Vibrio* 60 pMMB68 growth curve as determined by absorbance readings at 600nm. A rapid burst of LT-B production followed the addition of IPTG at 2h, the concentration of which increased exponentially until 4h, after which it levelled off, becoming stationary, along with cell concentration, at 8h. As there was no subsequent decrease in LT-B concentration, culture supernate was routinely harvested in LT-B production runs at 10h post-inoculation for processing by affinity chromatography by the method of Uesaka *et al.* (1994) (see page 80).

Purification of LT-B

Five cycles of adsorption of *Vibrio* 60 pMMB68 10 hour culture supernate to the immobilised galactose affinity column and elution with galactose, provided material for initial analysis of the purified product.

A further 15 full production cycles were then completed to provide the material for biopanning work. The peak fractions from the fifteen runs, as determined by SDS-PAGE and G_{M1} ELISA, were pooled and then concentrated with PEG to 7ml.

Characterisation of purified LT-B

The presence of LT-B in the concentrated product of affinity-column purification was demonstrated by SDS-PAGE, Western blotting and G_{M1} ELISA.

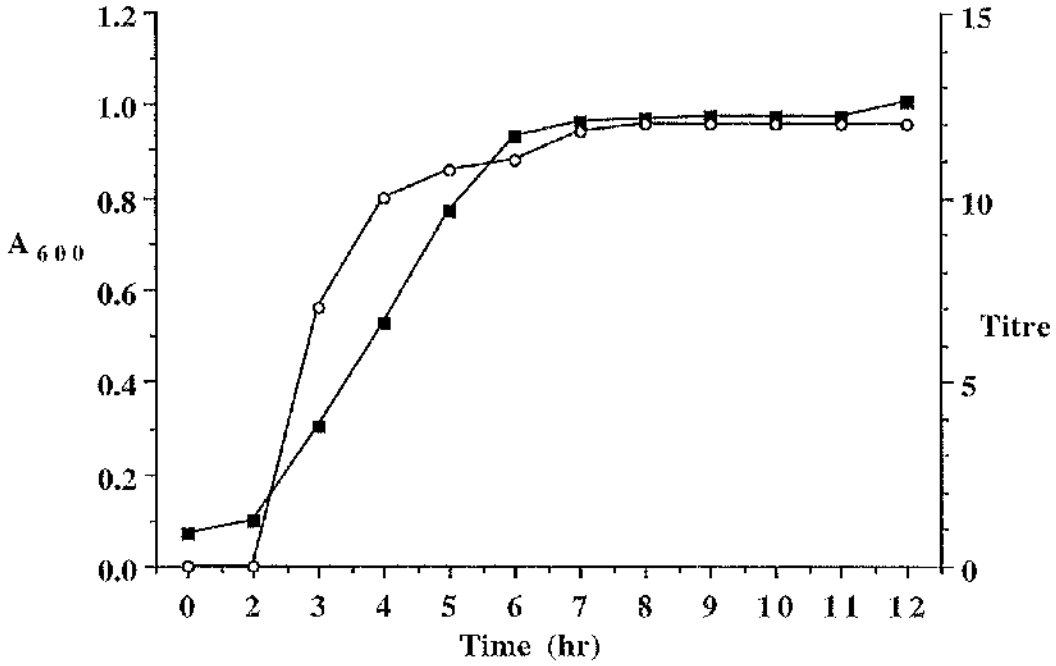


FIGURE 15 : Production of LT-B during growth of *Vibrio* 60 pMMB68

The A_{600} of samples from the growing culture and the relative concentration of LT-B, determined by G_{M1} ELISA, were measured. Titre indicates the $\log_2$ dilution at which the ELISA absorbance was 0.2. IPTG was added to the culture 2 hours post-inoculation.

—■— A_{600} ; —○— titre.

SDS PAGE

Figure 16 shows the result of SDS-PAGE with heated (100°C) and unheated samples of the column eluate. Both lanes showed only one band; in the unheated sample the band had an apparent M_r of 41500, but in the heated sample the M_r was 14000. Thus it appeared that the protein eluted from the affinity column disassociated when heated in the presence of SDS and mercaptoethanol, but did not do so at room temperature. The observed M_r and pattern of dissociation were those reported for LT-B, as discussed on page 142.

Western blotting

The protein eluted from the affinity column was transferred from an SDS-PAGE gel to nitrocellulose membrane by Western Blotting. After probing with the anti-LT-B monoclonal antibody 118.87, and adding peroxidase substrate, the membrane showed a single band of high intensity that corresponded to the $M_r = 41500$ band revealed by Coomassie blue staining. The gel and membrane are shown in Figure 17.

G_{M1} ELISA

G_{M1} ELISA was performed on the affinity column eluate, using the anti-LT-B monoclonal antibody as a probe. The result (Figure 18), demonstrated that the protein interacted strongly with G_{M1} such that at a dilution 1/4096, the greatest dilution tested, the ELISA A_{450} was still greater than double the background signal. The reduced A_{450} occurring at dilutions up to 1/256 may have resulted from steric hinderance (discussed on page 145).

Given that the eluted protein displayed electrophoretic properties characteristic of LT-B, a strong affinity with the anti-LT-B monoclonal antibody 118.87 as demonstrated by both Western blotting and G_{M1} ELISA, and the ability to bind to ganglioside G_{M1},

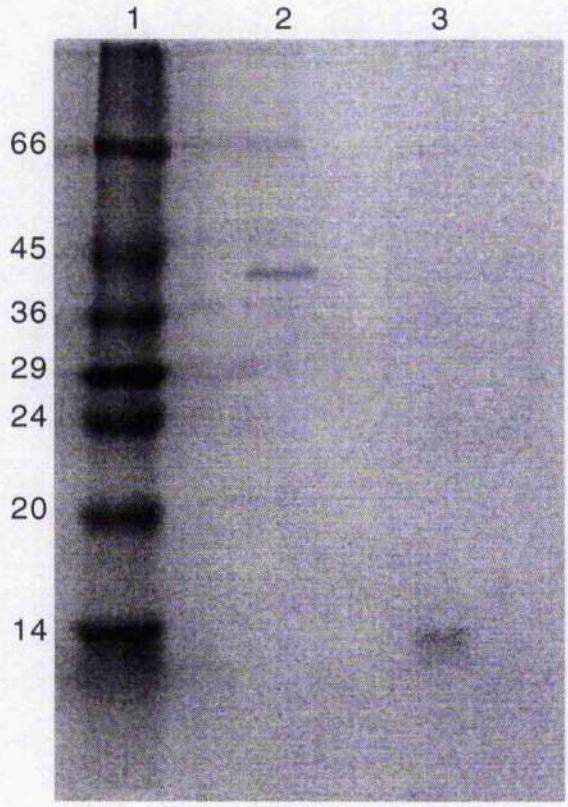


FIGURE 16 : SDS-PAGE analysis of samples produced by galactose affinity purification from *Vibrio* 60 pMMB68 culture supernate.

Lane 1 contained SDS7 molecular weight markers (molecular weights in kDa shown to left); Lanes 2 and 3 contained protein eluted from the affinity chromatography column and added to solubilising buffer. The sample in lane 3 was heated prior to loading on the gel, that in lane 2 was left at room temperature.

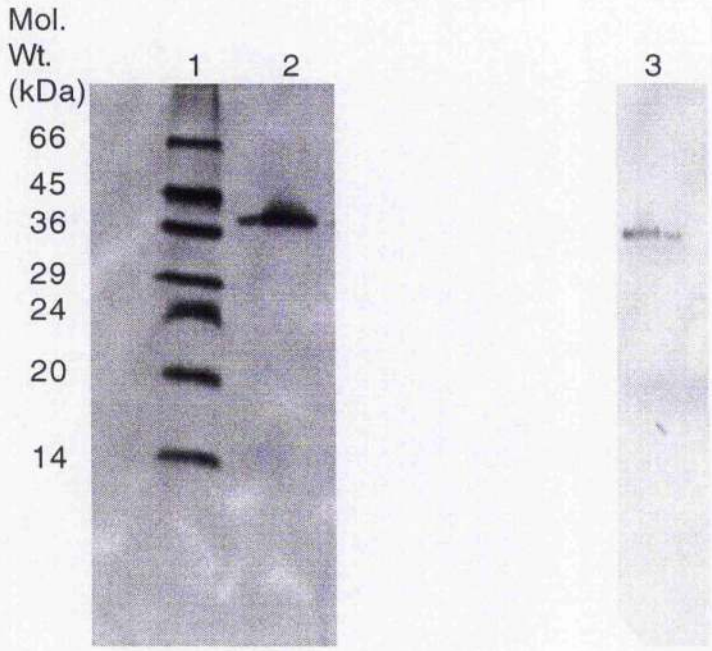


FIGURE 17 : Analysis of affinity column purified sample by SDS-PAGE and Western blotting.

Lanes 1 and 2 show the Coomassie blue stained gel of, respectively, SDS7 molecular weight markers (molecular weights shown to left) and the affinity column fraction (unheated). Lane 3 shows the Western blot of this gel, probed with anti-LT-B antibody and HRP-conjugated second antibody, and developed by addition of an HRP substrate.

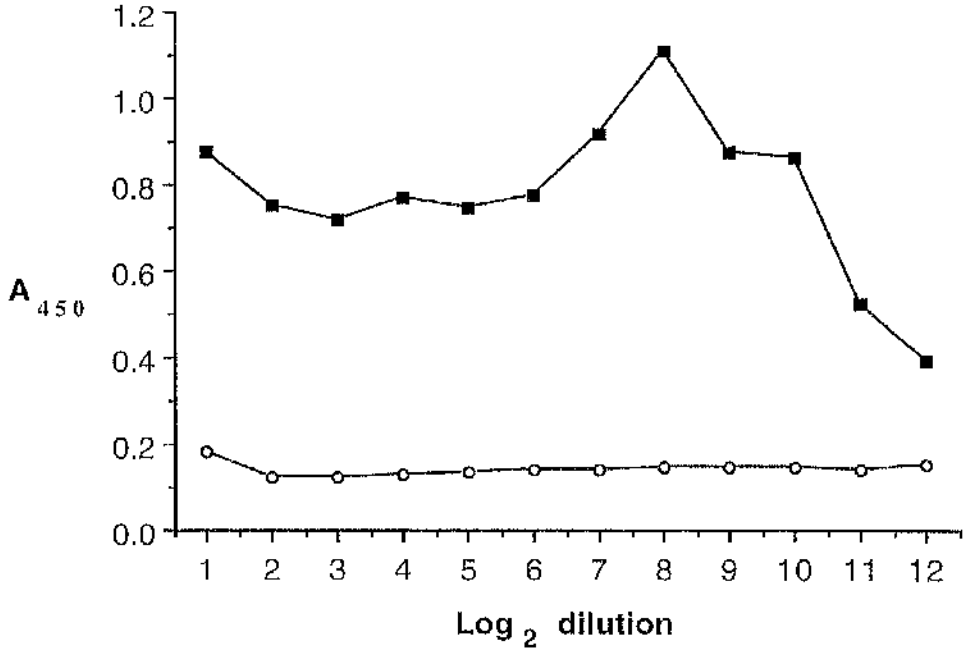


FIGURE 18 : Characterisation of purified column fraction by G_{M1} ELISA.

Doubling dilutions of the protein purified from the affinity column (—■—) and of PBS (—○—) were made in the wells of a microtitre plate coated with G_{M1}. After absorption and appropriate washing cycles (see page 82) wells were probed with anti-LT-B antibody and HRP-conjugated second antibody before addition of an HRP substrate. The absorbance at 450 nm was measured.

it was considered that the protein purified from the immobilised galactose affinity column was indeed LT-B.

Protein concentration of purified LT-B samples

Before analysing eluted material for protein content, it was necessary to ensure that the method employed would not be affected by the presence of varying concentrations of galactose in the eluted fractions.

Markwell *et al.* method

Figure 19 (a) shows that aliquots of BSA, prepared to a constant concentration of 60µg ml^{-1} in increasing concentrations of galactose, failed to give a consistent reading with this adaptation of the Lowry assay. Read against the 0mM galactose sample, the absorbance was a full unit higher when 500mM galactose was present.

Bradford method

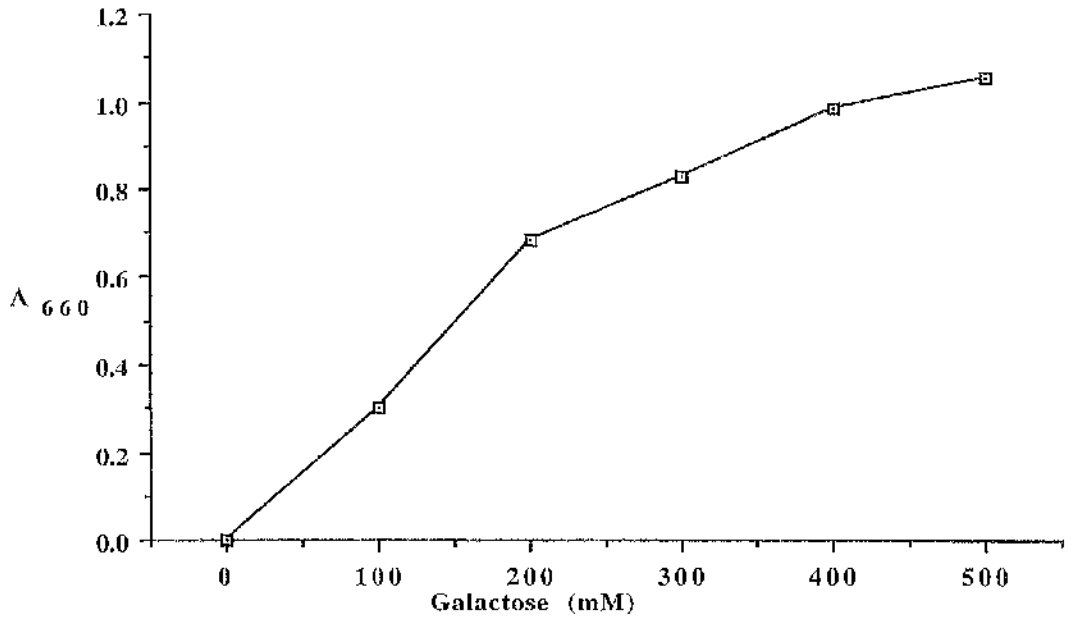
Figure 19 (b) demonstrates that the absorbances, when read against a reagent blank, remained constant regardless of increasing galactose concentration. All further protein estimations were performed using this method.

LT-B yields as determined by Bradford Assay

The fractions eluted from five purification runs with the immobilised galactose column, shown by SDS PAGE to contain pure LT-B, were pooled and tested for protein content. These pools had an average volume of 10.6ml (maximum 12.5ml, minimum 7.6ml), and an average protein content of 880µg (maximum 1025µg, minimum 700µg). The pools were then individually concentrated with PEG 10000 to an average of 780µl total volume, (maximum 1050µl, minimum 600µl), each of which contained an average of 762µg of protein (maximum 960µg, minimum 670µg). The 7ml pool from the fifteen production runs contained 2.1mg protein.

The Hirst LT-B was supplied at a concentration of 10mg per ml.

(a)



(b)

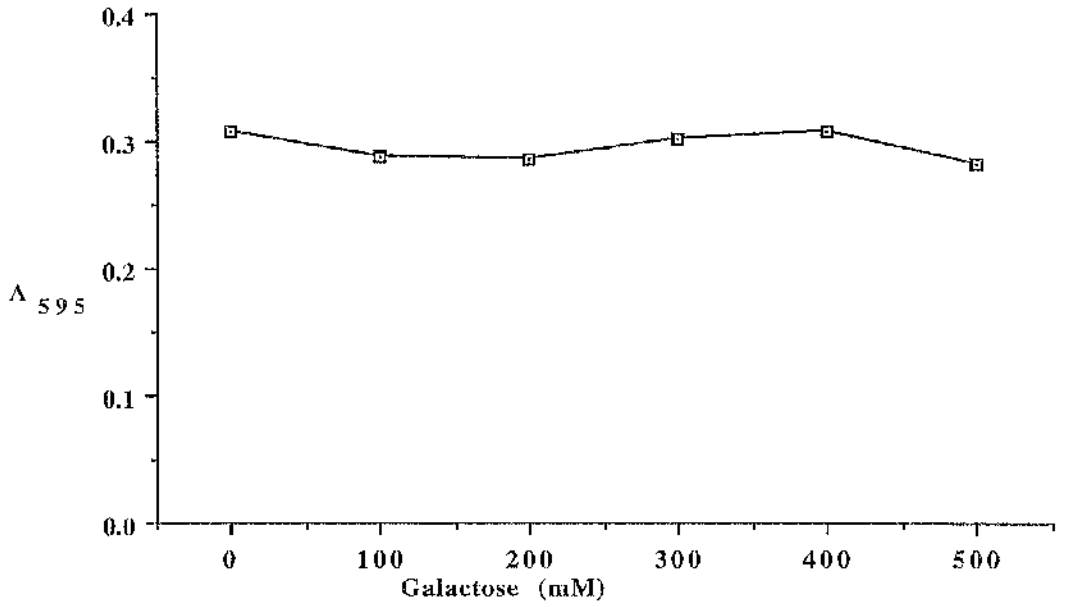


FIGURE 19 : Assay of a fixed amount of BSA in the presence of increasing concentrations of galactose.

(a) Markwell adaptation of the Lowry method; (b) Bradford assay.

Having established the protein content of the LT-B samples, they were then biotinylated.

Biotinylation of purified LT-B

Biotinylation of the stock LT-B, produced and assayed as described above, and the Hirst LT-B was performed with the Sigma Immunoprobe Biotinylation kit, as described on page 84, according to the protocol supplied.

Degree of biotinylation

The degree of biotinylation of both batches of LT-B was determined using the method outlined on page 84.

Stock LT-B

By this method a corrected A_{500} of 0.032, equivalent to a final biotin concentration of 9.41 nmoles ml^{-1}, was indicated. The biotin / protein ratio was calculated from the protein concentration (2.1 nmoles ml^{-1}), yielding 4.48 moles of biotin per mole of LT-B.

Hirst LT-B

Using the method described above, the biotin / protein ratio of the Hirst LT-B^{biotin} was demonstrated to be 2.60 moles of biotin per mole of LT-B.

Analysis of biotinylated LT-B

G_{M1} ELISA was used to confirm that biotinylation had not diminished the G_{M1}-binding capability of stock LT-B.

In Figure 20 (a) the ELISA was carried out using the anti-LT-B monoclonal antibody 118.87. Starting with a 1 / 20 dilution, a further 7 doubling-dilutions were required to reduce the signal to control level. Figure 20 (b) shows the result obtained when an identical dilution series was analysed with HRP-labelled streptavidin (Extravidin).

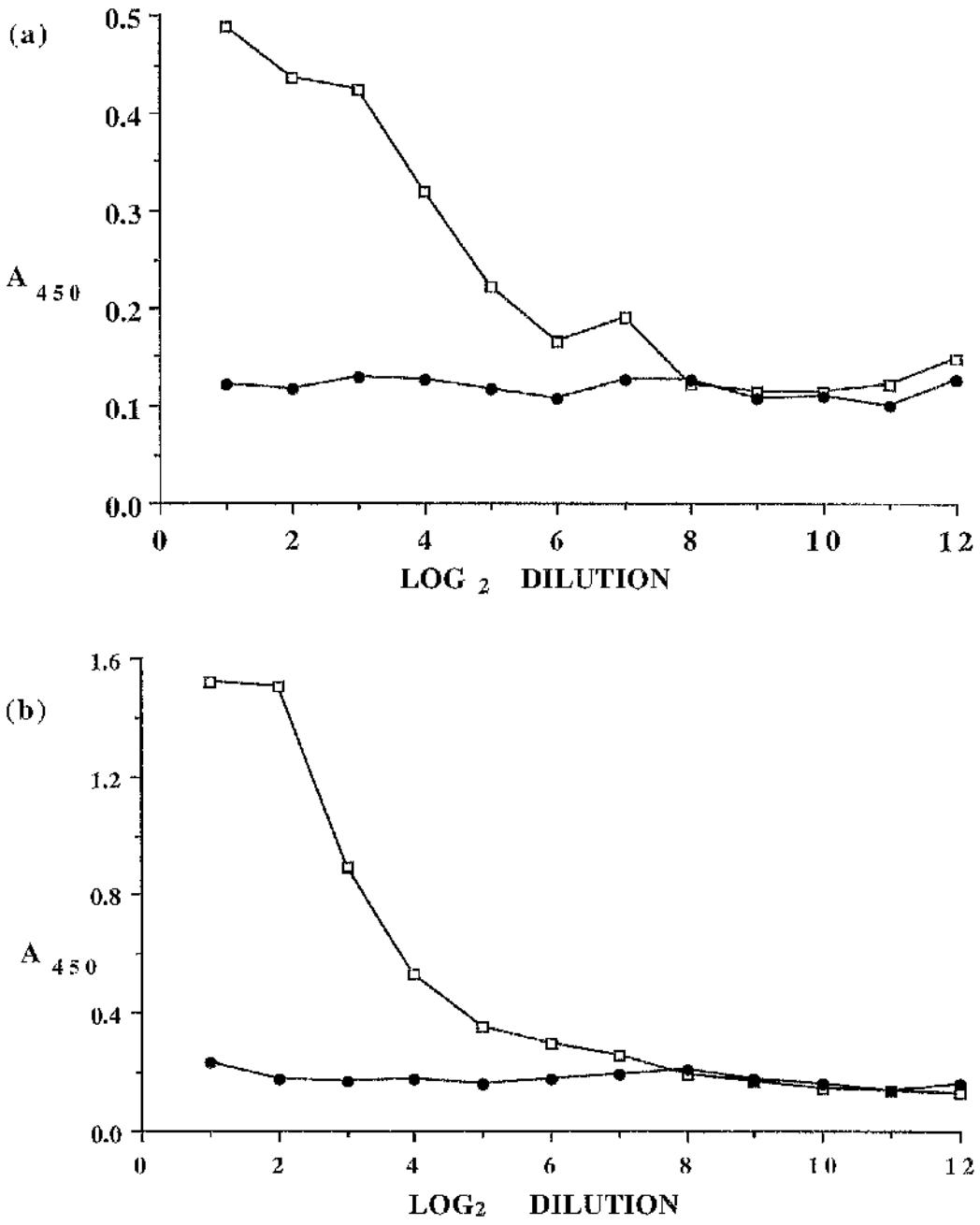


FIGURE 20 : G_{M1} ELISA of biotinylated LT-B.

(a) Probed with anti-LT-B monoclonal antibody 118.87. (b) probed with Extravidin.

—●— A_{450} of PBS-only control ; —□— A_{450} values for biotinylated LT-B.

Although in this case the initial absorbance was higher, it again took eight dilutions to reach background level.

These results confirmed that the biotinylated LT-B was capable of binding to G_{M1}.

Finally, biotinylation of LT-B did not affect the ability of the protein to interact with a monoclonal antibody specific for the native protein or to interact with ganglioside G_{M1}. The biotinylated LT-B ($LT\text{-}B^{biotin}$) was therefore considered suitable for use with the phage libraries.

SMITH HEXAMER LIBRARY

Infectivity of *E. coli* K91kan

The optimum time in the *E. coli* growth cycle at which the maximum number of transformants could be produced following infection by phage was determined using cells removed from the culture at half-hourly intervals from 6 hours post-inoculation onwards. Figure 21 shows the variation in colony numbers (representing T.U.) over the test period and clearly confirms that the ability to infect K91kan is growth-cycle dependent, with peak infectivity occurring when the culture had an A_{600} in the range 1.5 to 2.5.

Selection of binding phage by biopanning

Enumeration of eluted phage

The initial 'Amplified Library Stock' contained 1.1×10^{13} phage ml^{-1}, and 100µl of this suspension (i.e. 1.1×10^{12} phage) was panned against streptavidin-bound $LT\text{-}B^{biotin}$. Three such experiments were conducted, being eluted with acid elution buffer, G_{M1} and TBS respectively. Phage eluted by each agent were amplified and retested in a further three identical panning rounds.

The numbers of phage applied to and eluted from the tubes in each round were determined and the percentage of applied phage which was recovered was measured to calculate the degree of enrichment achieved.

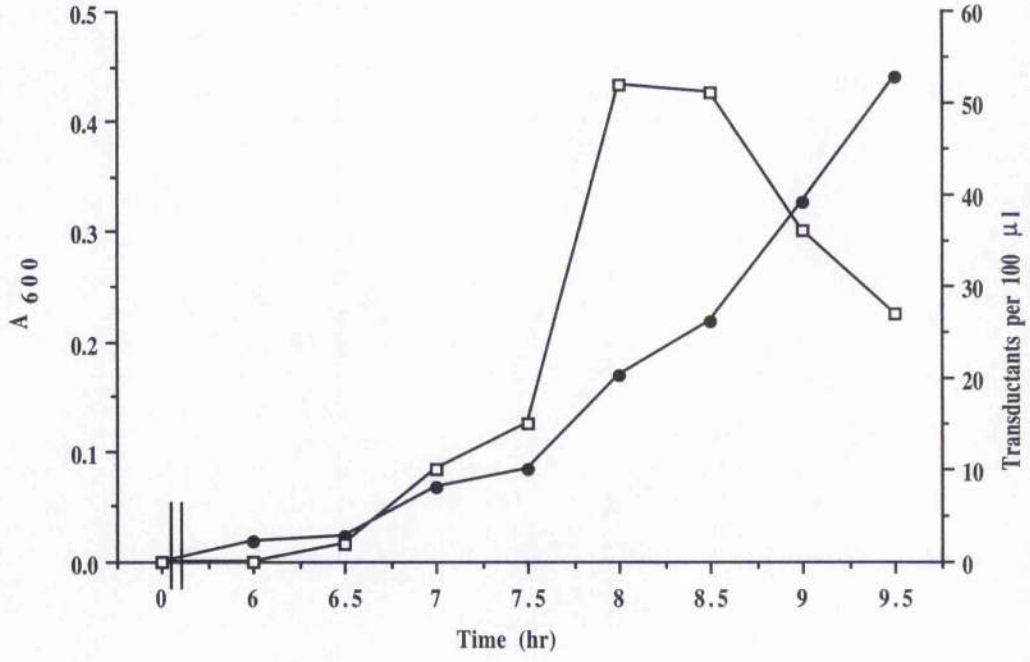


FIGURE 21 : Number of transductants produced on infection of *Escherichia coli* K91 kan at different times in the growth cycle.

Escherichia coli K91 kan was grown in Terrific broth at 37°C . At 30 minute intervals from 6 h, A_{600} was recorded and an aliquot of the culture was infected with stock Smith hexamer library phage. Infected cells were plated on LB (plus tetracycline and kanamycin) agar that were incubated at 37° overnight. Each tetracycline-resistant colony represented 1 transducing unit.

—●— A_{600} of the culture; —□— number of transductants arising from infection of the culture.

Table 6 shows the number of phage eluted in the first round, the numbers applied and eluted in subsequent rounds and the overall enrichment, defined as the increase in percentage of phage recovered from the first round to the last.

Sequencing of selected phage

Sequencing reaction

To determine whether the phage selected by biopanning carried consensus amino acid sequences, DNA inserts were determined for representative recovered phage.

However, in attempting to read the sequences from the autoradiographs a major problem arose because a second sequence, identified as that of wild-type pIII gene (i.e. unmodified by insertion), was present in the same lanes of the gel as that of the selected phage. A sample autoradiograph, highlighting the problem, is shown in Figure 22. The wild-type sequence did not appear with consistent intensity, being in some cases close to background levels but in other samples it was the predominant signal. In the latter cases, the slightly offset nature of the wild-type sequence meant that it could not simply be disregarded; in these cases the insert sequences could not be determined. A fuller explanation of the problem, and of the attempts made to overcome it, are discussed later (see page 148).

The sequences from 20 acid-eluted, 8 G_{M1}-eluted and 11 TBS-eluted clones are presented in Tables 7 (a), (b) and (c).

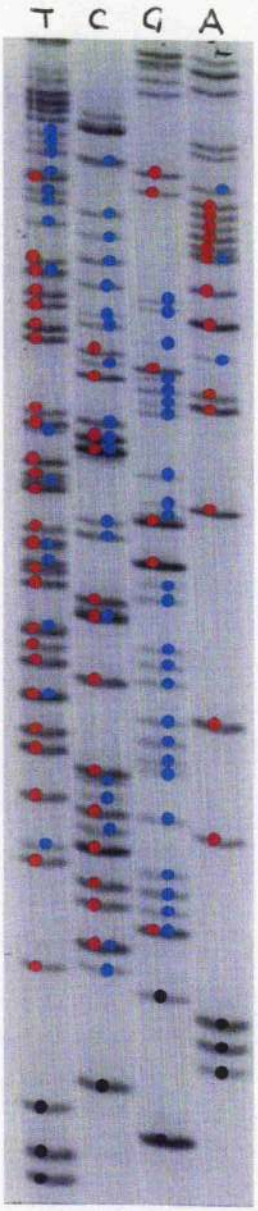
Acid eluted phage carried inserts which could be viewed as belonging to three distinct groups. Group 1 consisted of ten clones which carried three different fusions. Seven had the sequence **PWPWLG**, two of which (type ii) were translated from a different base sequence to the other five (type i); *ie.* they featured the codon CCG, as opposed to CCT, for the **P** residue in position 3. Two (type iii) had a sequence two amino acids different to this, **PWPCAG**. One (type iv) had a sequence, **PLPLAG**, that showed three amino acid differences from type i, but sufficiently strong similarities to justify inclusion in this group : *ie.* **P** in positions 1 and 3; repeated, hydrophobic amino acids in positions 2 and 4 (although aliphatic **L** in this case, as opposed to aromatic **W** in

	ACID ELUTED	% OF APPLIED	TBS ELUTED	% OF APPLIED	GM1 ELUTED	% OF APPLIED
Round 1 applied	1.1×10^{12}		1.1×10^{12}		1.1×10^{12}	
Round 1 eluted	4.5×10^{5}	4.1×10^{-5}	6.7×10^{5}	6.1×10^{-5}	5.9×10^{5}	5.4×10^{-5}
Round 2 applied	9.7×10^{11}		4.7×10^{11}		3.0×10^{11}	
Round 2 eluted	4.2×10^{6}	4.3×10^{-4}	1.8×10^{6}	3.8×10^{-4}	1.2×10^{6}	4.0×10^{-4}
Round 3 applied	3.3×10^{11}		3.5×10^{11}		2.5×10^{11}	
Round 3 eluted	7.1×10^{5}	2.1×10^{-3}	3.0×10^{6}	8.5×10^{-4}	5.1×10^{6}	2.0×10^{-3}
Round 4 applied	2.7×10^{10}		3.5×10^{10}		1.1×10^{10}	
Round 4 eluted	9.1×10^{6}	3.3×10^{-2}	3.3×10^{6}	9.4×10^{-3}	4.4×10^{6}	1.9×10^{-2}
ENRICHMENT		**805**		**154**		**351**

TABLE 6 : Biopanning with the Smith hexamer library

Values show the number of phage applied to, and eluted from, the plate in each round; the number eluted as a percentage of the number applied, and the overall enrichment of the phage (i.e. percentage eluted in Round 4 divided by percentage eluted in Round 1) are also shown.

(a)



(b)

Signal peptide

```
5' TTTCTATTCTCACTCG-GCCGACGGGGCTCCGTGGCCTTGGCTGGGTGGGGCCGCTGGGGCC---
   ||||||||||||||||
5' TTTCTATTCTCACTC-C----------------------------------------------GCT
```

GAAACTGTT 3'

FIGURE 22: Dual sequencing signal obtained from Smith hexamer library phage.

(a) An autoradiograph showing a mixture of wild-type (red dots) and recombinant (blue dots) signals. Prior to the insert the sequence is common (black dots).

(b) The sequences read directly from the autoradiograph are presented with the common sequence (black), wild type (red) and recombinant (insert) sequence (blue). The following key applies :

Underlined sequence = hexameric insert.
"Overlined" sequence = restriction sites engineered to facilitate insertion of random hexamer in the phage library.
Vertical lines = nucleotide identity
Dashes = spaces introduced to aid alignment

WILD TYPE

```
5' TCTCACTCCGCT----------GAAACTGTT 3'
   S H S A          G T V
```

GENERIC RECOMBINANT

```
5' TCTCACTCGGCCGACGGGGCTNNKNNKNNKNNKNNKNNKGGCGCCGCTGGGGCCGAAACTGTT 3'
   S H S A D G A X X X X X X G A A G A G T V
```

(a)

GROUP 1

```
(i)   .....CCGTGGCCTTGGCTGGGT.....
      S H S A D G A P W P W L G G A A G A G T V

(ii)  .....CCGTGGCCGTGGCTGGGT.....
      S H S A D G A P W P W L G G A A G A G T V

(iii) .....CCGTGGCCTTGTGCGGGT.....
      S H S A D G A P W P C A G G A A G A G T V

(iv)  .....CCGTTGCCTTTGGCGGGT.....
      S H S A D G A P L P L A G G A A G A G T V
```

GROUP 2

```
(i)   .....TGGTTTAATGAGCCGTTT.....
      S H S A D G A W F N E P F G A A G A G T V

(ii)  .....TGGTTTAATATTGTTTTT.....
      S H S A D G A W F N I V F G A A G A G T V

(iii) .....TGGTTGAAGATGCTGTTT.....
      S H S A D G A W F K M L F G A A G A G T V

(iv)  .....TGGGTTAATAGGTCGTTT.....
      S H S A D G A W V N R S F G A A G A G T V

(v)   .....TGTTTTAATATGCCGTTT.....
      S H S A D G A C F N M P F G A A G A G T V
```

GROUP 3

```
(i)   .....TTGTTGAGTTGCTGGTTT.....
      S H S A D G A L V E L L F G A A G A G T V

(ii)  .....ACTTCGCTGGTCCCGGCT.....
      S H S A D G A T S L V P A G A A G A G T V

(iii) .....CGTACGCCTAAGAATTCT.....
      S H S A D G A R T P K N S G A A G A G T V
```

TABLE 7 : Nucleotide and amino acid sequences of inserts carried by clones from the Smith hexamer library.

The generic recombinant sequence is shown at the top of each page, followed by the nucleotide sequence and amino acid translation of each eluted phage (a - Acid eluted ; b - G_{M1} eluted; c - TBS eluted). The wild-type sequence is shown at the top of this page for comparison.

Dots show identity to the generic recombinant sequence; dashes, indicating absent nucleotides, were introduced to maintain alignment.

Note that c (v) carries a tripeptide insert

GENERIC RECOMBINANT SEQUENCE

5' TCTCACTCGGCCGACGGGGCTNNKNNKNNKNNKNNKNNKGGGGCCGCTGGGGCCGAAACTGTT 3'

S H S A D G A X X X X X X C A A G A G T V

(b)

(i) .....TTTTTTATGTCTCGTTTT.....

S H S A D G A F F M S R F G A A G A G T V

(ii) .....AGTAATTCGGTGAATGGG.....

S H S A D G A S N S V N G G A A G A G T V

(iii) .....TTTCTTAATTGCATCTT.....

S H S A D G A F L N L H L G A A G A C T V

(iv) .....ATTGCGTTGTTGCATTTT.....

S H S A D G A I A L L H F G A A G A G T V

(v) .....GGGGGGCGTCAGAAGCTT.....

S H S A D G A G G R Q K L C A A G A G T V

(vi) .....GGTTTGCGTATTAGTTAT.....

S H S A D C A G L R I S Y G A A G A G T V

(vii) .....AATACGAGTAAGGTGTTT.....

S H S A D G A N T S K V F G A A G A G T V

(viii) .....TCGTCTAGGATTTCGAAT.....

S H S A D G A S S R I S N G A A G A G T V

(c)

(i) .....CATCTTCTGTATCCGAAT.....

S H S A D G A H L L Y P N C A A G A G T V

(ii) .....TTGCAGGGTTTGATGTTT.....

S H S A D G A L Q G L M F G A A G A G T V

(iii) .....TTGGGGCTGAGGGGGAAT.....

S H S A D G A L G L R G N G A A G A C T V

(iv) .....CCTGCTGTTGTTCTTAAG.....

S H S A D G A F A V V L K G A A G A G T V

(v) .....TCGCTTGTT---------.....

S H S A D G A S L V G A A G A G T V

(vi) .....CGTCCTATGCATGCGATT.....

S H S A D G A R P M H A I G A A G A G T V

(vii) .....TTTGTGCCGGTGCGTAAT.....

S H S A D G A F V R V R N G A A G A G T V

(viii) .....TTTCCGATGGTTACTCGG.....

S H S A D G A F P M V T R G A A G A G T V

(ix) .....CGGGTGGTTGCGGCTCCT.....

S H S A D G A R V V A A P G A A G A G T V

(x) .....TCTGCGTCGGGGCTTTGG.....

S H S A D G A S A S G L W G A A G A G T V

(xi) .....ACTACTACTTCTACTCAG.....

S H S A D G A T T T S S E G A A G A C T V

types i and ii) and **G** in position 6. The **A** in position 5 of type iii, like the **L** in the same position in types i and ii, is aliphatic.

Group 2 consisted of seven clones, of five sequence types, all of which featured **F** as the last amino acid in the random insert. Four types (i, ii, iii and iv) had **W** as their first residue, four (i, ii, iii and v) had **F** as their second, four (i, ii, iv and v) had **N** as their third. In addition, two (iii and v) had **M** in the fourth position and two (i and v) had **P** in the fifth.

No alternate codons were observed for any of the amino acids in group 2, but of all the amino acids featuring more than once in the group, only **P** can be translated from an alternative codon.

Group 3 consisted of three sequences which bear no apparent relationship, with the possible exception that Group 3(i) may be related to Group 2(iii) via the residues **L** and **F** in positions 5 and 6.

G_{M1} elution yielded eight phage with inserts that showed only tenuous similarities. The most noteworthy are types i, iv and vii, each of which terminate with **F**, as did the Group 2 acid eluted phage, although in this case the other amino acids in the sequence bear no relationship either to these or to each other. A low degree of similarity is also observed between type iii (**FLNLHL**) and iv (**IALLHF**), in that both feature **LH** in positions 4 and 5, as well as **F** and an additional **L**, but again, the dissimilarities outweigh the similarities.

TBS elution gave eleven unrelated phage, including one (v) that had a tripeptide insert.

Binding assays

To determine whether the phage selected by biopanning had been enriched because of their affinity for LT-B rather than streptavidin or plastic, phage prepared from each of the acid, G_{M1} and TBS-eluted clones were assayed. The phage were used to coat microtitre plate wells whilst control wells received TBS alone. Stock LT-B^{biotin} was added to all wells and the extent to which it was 'captured' in the well was

determined by measuring A_{450} after the addition of anti-LT-B monoclonal antibody, HRP-conjugated anti-mouse antibody and HRP-substrate.

Acid-eluted phage gave an average A_{450} of 0.053, G_{M1} eluted phage an average of 0.062 and TBS eluted phage an average of 0.066. No well containing phage of any type gave a reading of greater than 0.08. When compared with the average for the control wells (0.052), it was apparent that no phage showed any evidence of ability to capture LT-B.

To demonstrate that this result was not due to inadequate phage-coating of the wells, bound G_{M1}-eluted phage were tested by the addition of anti-phage antibody, HRP-conjugated anti-rabbit antibody and HRP-substrate. Control wells, which did not contain phage, were treated similarly. In this case, phage-coated wells gave A_{450} readings in excess of 2.0, some 7.7 times the average of the control well readings, thus confirming that the plastic surface was efficiently coated with the phage under investigation.

LUZZAGO NONAMER LIBRARY

The Luzzago library expresses a constrained nonameric peptide on pVIII (see page 63 for details), and a transformation assay with *E. coli* XL-1 cells indicated that the stock library contained approximately 2.5×10^{14} phage ml^{-1}, therefore amplification was unnecessary. Uninfected *E. coli* XL-1 failed to grow on the ampicillin agar used in the assay.

Selection of binding phage by biopanning

First round of biopanning

Three 2µl aliquots (i.e. 5×10^{11} phage) from the phage library were individually panned against streptavidin-bound stock LT-B^{biotin}. Bound phage were eluted with acid elution buffer, G_{M1} and TBS respectively.

Each eluate was mixed with XL-1 cells in LB broth and was plated onto nine LB agar plates containing ampicillin. After incubation, all twenty-seven plates were covered with around 1500 small colonies, which were scraped from the plate, resuspended in

broth, infected with M13K07, and used to produce phage for the second round of panning.

Subsequent rounds of biopanning

Amplified eluates were mixed with LT-B^{biotin} and panned in streptavidin-coated Maxisorp tubes. Bound phage were eluted with the same agent used in the previous round.

The second and third rounds of biopanning produced eluates that, when mixed with XL-1 cells and plated out, all yielded approximately 1500 small colonies per plate. Those produced from second round eluates were used to prepare phage virions for use in the third round. Some of the third round colonies were picked at random and sequenced.

Sequencing of phagemid DNA

The nature of the peptides expressed by 20 acid-eluted, 10 G_{M1}-eluted and 10 TBS-eluted clones were determined by DNA sequencing. That there was a serious problem with the construction of the library was immediately apparent. In every case there were only 26 bases between the two TGC (**C**) codons that flanked the random insert, and a base was missing from the codon for $\mathbf{P}^{19}$ of mature pVIII (CC instead of CCC). In addition, there was a base substitution in the codon for $\mathbf{S}^{-14}$ in the signal sequence (TCT for TCG). The expected and actual sequences are shown in Figure 23.

Also shown in this figure is a sequence characteristic of three of the TBS-eluted phage. Downstream of the codon for $\mathbf{F}^{5}$, these were as described above (*ie.* they lacked a C in the codon for $\mathbf{P}^{19}$ and carried a random insert of 26 bases). Upstream of the codon for $\mathbf{F}^{5}$ the nucleotide sequence bore no relationship to that predicted.

Since the location of the missing base in the characterised inserts could not be identified, no inference could be made of the protein sequence carried on phage eluted in the final round of biopanning. Taking all sequences into account, C was the most prominent base in the random insert (approximately 35.5%); then G (24.6%); T (23.4) and A (16.1%).

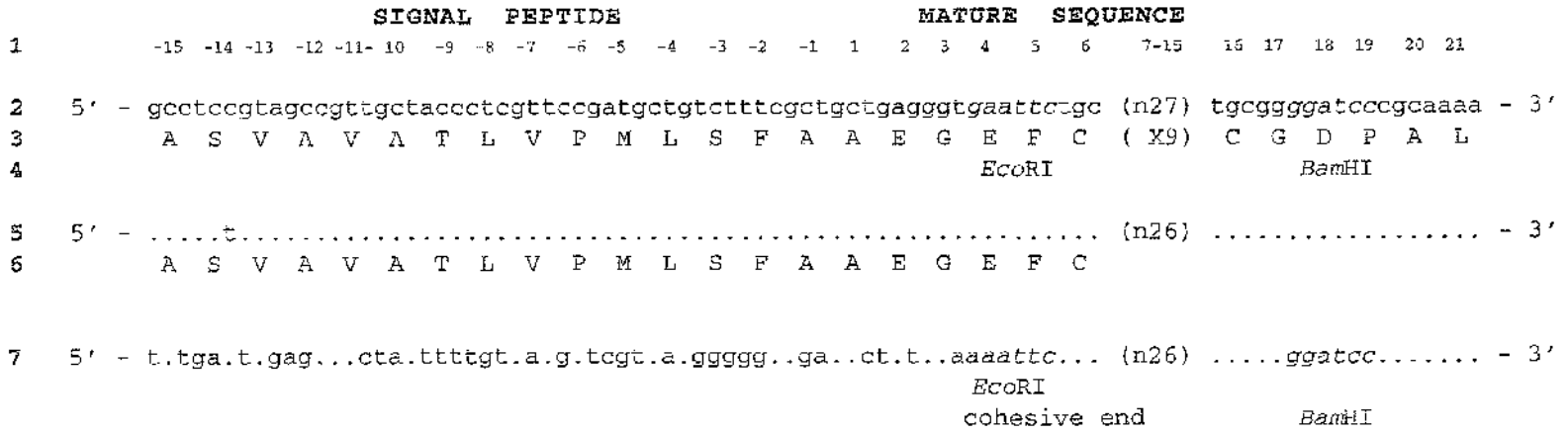


FIGURE 23 : Sequencing of clones from the Luzzago library.

Row 1 shows the position relative to the signal peptide and gene VIII itself, whilst Rows 2 and 3 respectively show the predicted codon and amino acid sequences. Restriction sites, italicised in Row 2, are named in Row 4. Rows 5 and 6, respectively, show the actual codon sequence and amino acid translations obtained from all clones selected by biopanning apart from some obtained with TBS-elution that had a more unusual sequence (as shown in row 7). Note that in this latter sequence, only the *Eco*RI cohesive end was present.

Nucleotide sequences of the 26-base pair insert from the three unusual TBS-eluted phage described above are shown in part (a) of Table 8, whilst those of five representative phage eluted by acid are shown in part (b).

SMITH PENTADECAMER LIBRARY

The Smith pentadecamer was originally constructed by George Smith of the University of Missouri who provided the hexamer library also used in this work. The viral capsid displays both wild-type and synthetic pVIII, the 15 base-pair insert being fused to the latter. This form of fusion, Type 88, is discussed on page 10.

Selection of binding phage by biopanning

First biopanning experiment

The number of phage in the original library was ascertained to be 5 x 10^{12} T.U. ml^{-1}. In the first biopanning round, 30 µl (1.5 x 10^{11} phage) was mixed with 10µl stock LT-B^{biotin} prior to exposure to a streptavidin-coated plastic surface. Bound phage were eluted from the plastic with acid buffer. Amplified first round phage were used in a second identical round of panning (although 30 µl of phage were in this case mixed with 8 µl stock LT-B^{biotin}). A third panning round, using 30 µl amplified second round phage and 6 µl LT-B^{biotin} was then conducted.

The number of phage eluted at each round was as follows :

First round : => 1.2 x 10^3 T.U.

Second round : => 4.1 x 10^4 T.U.

Third round : => 1.4 x 10^5 T.U.

Sequencing of selected phage

Ten clones from the final round were chosen at random for further characterisation. The sequences of the inserts carried by each are shown in Table 9. In all but one of the clones the second codon is CAT, coding for **H**. In the other, clone number 8, CAT is the first codon. This represents the only occurrence of **H** in all sequences bar that of

a)

5'	CTGC	CCTA	CCCT	CGTC	AAAG	TTCA	GC	3'
	TACG	TAGT	ACGA	ATCT	GACC	CTCC	GA	
	TGGG	ATTG	ACCC	GTGT	CCAC	CTCA	CC	

b)

5'	CCTG	CCTG	ACGC	CCCT	GCTG	GTGA	CT	3'
	CGTC	GCGT	CCAC	TCTC	TCGC	TAGC	AC	
	CTAA	TGAT	GTCC	GGCC	CGCC	GGTC	GG	
	CCTG	CCTG	ACGC	CCCT	GCTG	GTGA	TC	
	CATA	CGAG	AAGT	CCTC	CGGC	CCTT	CT	

TABLE 8 : Sequences of random inserts carried by clones from the Luzzago library.

(a) Data from TBS-eluted clones with unusual sequences upstream of the codon for $\mathbf{F}^{5}$; (b) Sequences from representative acid-eluted clones.

Placement of the reading frame is unknown.

1	TCA	CAT	AGG	ACG	CGC	CGC	CAA	GGC	CTT	AGT	TCG	CTG	GGC	CGC	CTA
	S	H	R	T	R	R	Q	G	L	S	S	L	G	R	L
2	AAT	CAT	CGT	CCT	TTT	TCG	TCG	ACG	CGG	TCT	CTT	CAG	GCT	TTT	GCT
	N	H	R	P	F	S	S	T	R	S	L	Q	A	F	A
3	CAG	CAT	CGT	TCT	CTG	CGG	GGT	GAT	TGT	CGG	GGT	AGG	ATG	TGT	TCG
	Q	H	R	S	L	R	G	D	C	R	G	R	M	C	S
4	AGG	CAT	CTG	CCT	GCT	GGT	CAG	GGG	TTT	ATG	GTT	CCG	TTT	CTT	TCG
	R	H	L	P	A	G	Q	G	F	M	V	P	F	L	S
5	GCG	CAT	AGG	CAG	TCT	CCG	GCT	GCT	CCG	ATG	ACT	AAT	TTT	ATG	ACT
	A	H	R	Q	S	P	A	A	P	M	T	N	F	M	T
6	AGG	CAT	GGT	TTT	ACT	TTG	ATG	ACC	CTA	AAG	CGG	TTG	CGC	CTA	CTG
	R	H	G	F	T	L	M	T	L	K	R	L	R	L	L
7	CAG	CAT	AGT	AGT	TCG	CGT	CCG	AAT	CGG	CCT	CCG	TCT	ACT	CCG	CCT
	Q	H	S	S	S	R	P	N	R	P	P	S	T	P	P
8	CAT	ACG	ATC	ATT	CTG	CCC	AGG	TTG	GGG	ACG	AGG	CCG	CTT	ACT	ACT
	H	T	I	I	L	P	R	L	G	T	R	P	L	T	T
9	AAT	CAT	TGG	GCT	GCT	TAC	GAG	GGT	TAC	TGC	GAA	TCC	GTC	TCC	TTC
	N	H	W	A	A	V	E	G	V	C	E	S	V	S	F
10	AGG	CAT	AAA	TTC	GCG	CGT	GCG	GTT	TCG	ACT	CCT	CAT	CAG	CAG	ATA
	R	H	K	F	A	R	A	V	S	T	P	H	Q	Q	I

TABLE 9 : Nucleotide and amino acid sequences of inserts carried by acid-eluted clones from the Smith pentadecamer library after biopanning with LT-B^{biotin}.

The first row in each case shows the nucleotide sequence, the second shows the amino acid translation.

clone 10, where it also appears as residue 12. In seven cases, the **H** was immediately adjacent to **R** (preceding on three occasions, succeeding on four). The residues **N** and **Q** also each preceded **H** in two cases, with **S** and **A** preceding once each. Table 10 presents a table of expected and actual occurrence of amino acids. The most frequently occurring was **R**, constituting 21 of the 150 codons that were characterised; it occurred four times in two of the sequences (1 and 3) and failed to occur only in clone 9. The next most common residue was **S**, which featured 17 times, including 3 successively in number 7, which also was the clone that bore the least varied complement of amino acids, with five **P**, four **S**, and two **R** among its fifteen residues. The only amino acid that failed to feature at all was **Y**, with **W** and **D** appearing only once each.

Binding assay

An assay was devised to ascertain whether the phage isolated and sequenced as described above bound specifically to LT-B^{biotin}. Preliminary tests having indicated that they adhered to streptavidin rather than to LT-B^{biotin}, a spot-check assay was developed. In this assay, final-round phage were exposed in microtitre-plate wells to either streptavidin-bound LT-B^{biotin} or streptavidin alone. Those that bound were identified by successive additions of rabbit anti-phage antibody, HRP-conjugated anti-rabbit antibody and HRP substrate. The coloured product of the reaction was read at A_{450}. The result of this assay is shown in Figure 24.

The anti-phage antibodies did not react with either streptavidin alone or with streptavidin-bound LT-B (wells 3 and 4 respectively in the assay format described on page 101), both of which gave background readings not recorded in Figure 24. Background-level results were also obtained from the binding of phage to blocker alone (well 5).

AMINO ACID	NOMINAL FREQUENCY	NOMINAL OCCURRENCE	OBSERVED OCCURRENCE
A	0.065	9.75	10
C	0.032	4.8	3
D	0.032	4.8	1
E	0.032	4.8	2
F	0.032	4.8	8
G	0.065	9.75	9
H	0.032	4.8	11
I	0.049	7.5	3
K	0.032	4.8	2
L	0.098	14.7	15
M	0.016	2.4	5
N	0.032	4.8	4
P	0.065	9.75	13
Q	0.032	4.8	8
R	0.098	14.7	21
S	0.098	14.7	17
T	0.065	9.75	12
V	0.065	9.75	5
W	0.016	2.4	1
Y	0.032	4.8	0

TABLE 10 : Expected and actual occurrence of amino acids in the clones from the Smith pentadecamer library.

The nominal frequency is obtained by dividing the number of codons for any amino acid by 31, the number of coding triplets in the NNK genetic code. Nominal occurrence is derived by multiplying the nominal frequency by 150, the number of amino acids in the 10 pentadecamer inserts.

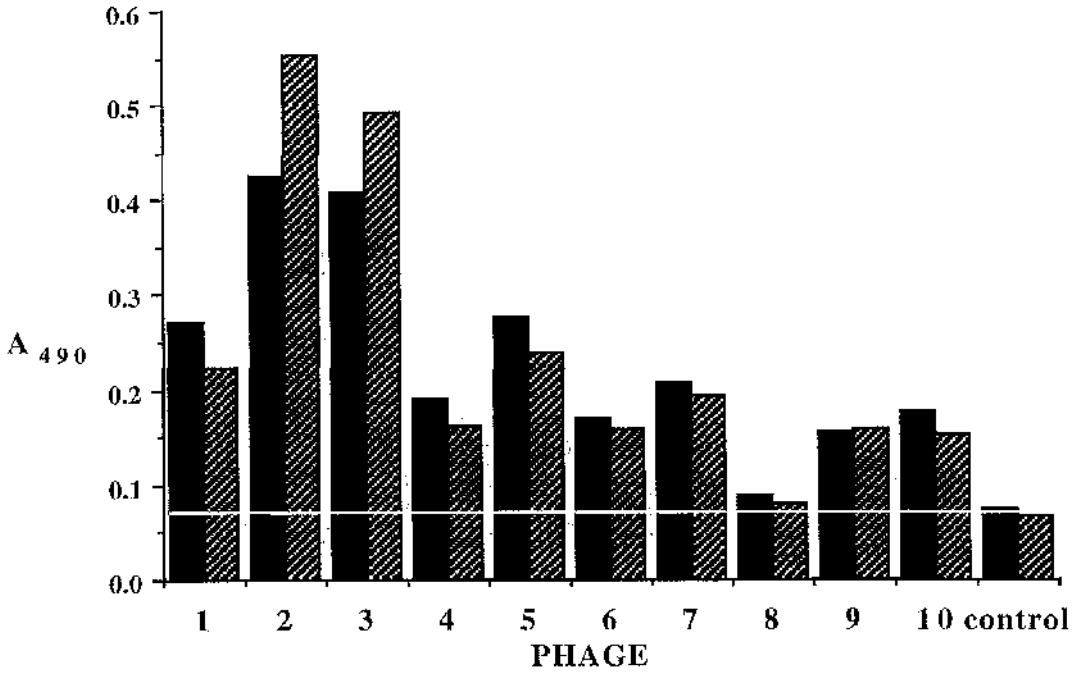


FIGURE 24 : Binding of phage selected from the Smith pentadecamer library to streptavidin and streptavidin-bound LT-B.

Purified phage clones were exposed to microtitre-plate wells coated with streptavidin-bound LT-B^{biotin} (■) or streptavidin alone (▨). Bound phage were identified by the addition of anti-phage antibody, HRP-conjugated second antibody and HRP substrate that yielded a chromagen read at A_{450}. Clone numbers are designated on the x-axis, with 'control' indicating results obtained with an aliquot of the original library. The horizontal line shows the background signal determined from this control.

Figure 24 shows the binding of each of the ten selected phage, and an aliquot of the original library, to streptavidin alone and to streptavidin-bound LT-B^{biotin}. Phage from the original library did not bind to streptavidin or LT-B^{biotin} at levels higher than background. Binding of the selected phage can therefore be measured against this figure, which, in Figure 24, is indicated by a horizontal line.

Clones 2 and 3 bound to the highest levels, (readings of 0.4 to 0.55) with readings in wells coated only with streptavidin some 25% higher than those with streptavidin-bound LT-B^{biotin}. Clones 1, 5 and 7 gave moderate readings (0.2 to 0.3) - in each case values were around 15% higher in streptavidin-bound LT-B^{biotin} wells than when the wells were coated only with streptavidin.

Clones 4, 6 and 10 gave low readings (0.1 to 0.2), which were 16%, 6% and 14% higher respectively in wells coated with streptavidin-bound LT-B^{biotin}. Values for clones 8 and 9 were barely higher than background and showed little difference between coatings.

Second biopanning experiment

In view of the findings from the binding assay described above a second panning experiment, using an alternative source of purified LT-B (Hirst LT-B^{biotin}), was conducted as previously described for the Smith pentadecamer library.

To assess the possible effect of binding to the medium supporting the LT-B during the panning process, 20µl phage were also panned, in parallel series, against streptavidin only and BSA-blocked plastic respectively. Elution was with acid buffer in each case and two further identical panning rounds followed.

The number of phage in the original library was ascertained to be 3×10^{11} T.U. ml^{-1}.

The number of phage eluted in each round, from each surface, was as follows :

		PANNING SURFACE		
		LT-B	STREPTAVIDIN	PLASTIC
First round	=>	1.0×10^4	1.0×10^4	1.0×10^5
Second round	=>	6.0×10^5	1.4×10^6	1.0×10^6
Third round	=>	1.6×10^7	2.3×10^6	2.7×10^6

'Colony lift' assay

The agar plates containing phage-infected colonies that were the by-product of the enumeration of final round eluates in this panning experiment were used in an alternative form of binding assay. They were overlaid with nitrocellulose membranes so that phage associated with each colony would be transferred to the membrane. Membrane-bound phage from the LT-B^{biotin}-panning experiment were exposed to non-biotinylated LT-B and streptavidin; phage from the streptavidin-panning experiment were exposed to streptavidin alone. Captured ligand was visualised by the addition of an antibody against the ligand and HRP-conjugated second antibody (in the case of LT-B) or biotinylated peroxidase (for captured streptavidin). To ensure that sufficient ligand was added, membranes impregnated with G_{M1} (for LT-B) and LT-B^{biotin} (for streptavidin) were simultaneously processed.

A second membrane was also overlaid on the plates and was used to confirm transfer of phage to the membranes by the addition of rabbit-anti-phage antibody, HRP-conjugated anti-rabbit antibody and HRP substrate.

LT-B - panned phage failed to capture any LT-B, the membranes remaining clear after substrate had been added. The G_{M1}-impregnated membrane darkened intensely. However, these phage did demonstrate a degree of affinity for streptavidin, as very faint spots appeared at positions that corresponded with colonies on the original agar plate. Streptavidin -panned phage demonstrated a similar, weak affinity. By contrast, the membrane spotted with LT-B^{biotin} reacted strongly with streptavidin as expected.

The control membranes that were probed with rabbit-anti-phage antibody and HRP-conjugated anti-rabbit antibody, gave a strong, positive reaction, the position of each colony being easily discerned. One such control is shown in Figure 25.

The 'colony lift' binding assay therefore identified colonies that displayed phage capable of binding streptavidin but not LT-B. That there were sufficient phage bound to the membranes was confirmed by the result obtained when anti-phage antibody was used, whilst the success of the assay principle was confirmed by the strongly-positive reactions obtained with membranes impregnated with G_{M1} and LT-B^{biotin}. This result confirmed that obtained with the microtitre plate assay. In that case, the A_{450} produced when phage were added to LT-B^{biotin} never significantly exceeded that obtained when streptavidin alone was used. The inference is therefore that it was the streptavidin to which the phage were binding. This implications of this result are discussed on page 165.

Third biopanning experiment

The previous experiments having apparently yielded phage with an affinity for streptavidin, it was decided to conduct one further panning experiment in which measures aimed at eliminating such undesired binding were employed. Firstly, an alternative form of immobilisation of Hirst LT-B, via Protein G-bound anti-cholera toxin, was used. Secondly, as described on page 97, amplified phage from each round were 'backchecked' by exposure to a well containing Protein G-bound anti-cholera toxin prior to their use in each subsequent panning round.

The number of phage eluted at each round was as follows :

First round : => 1.64×10^5 T.U.

Second round : => 1.94×10^5 T.U.

Third round : => 2.7×10^5 T.U.

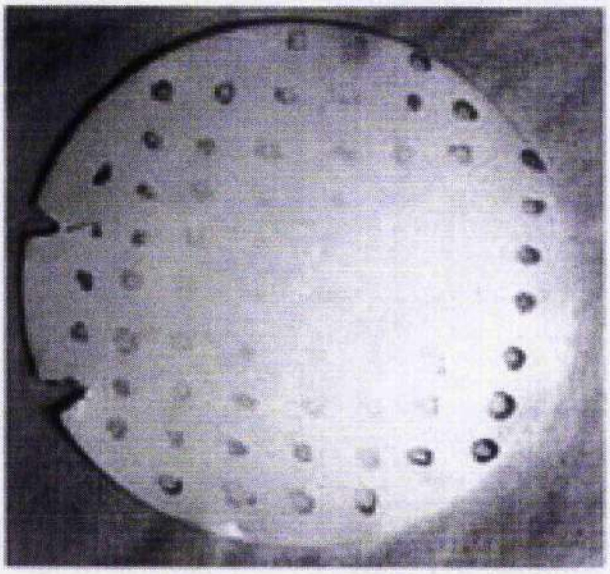


FIGURE 25 : Control membrane used in 'colony lift' binding assay.
A membrane filter was placed against colonies infected with phage selected by biopanning on an LT-B coated surface. Adsorbed phage were probed with rabbit anti-fd antibody and HRP conjugated anti-rabbit antibody. Result was visualised using the HRP substrate 4-chloro naphthol. The position of individual colonies is clearly apparent.

The wells of Protein G-bound anti-cholera toxin antibody used in backchecking were afterwards treated with acid elution buffer to release phage that had been removed by the backchecking process for enumeration purposes. The numbers eluted from these wells at each round were :

First round : => 5.82×10^3 T.U.

Second round : => 8.03×10^3 T.U.

Third round : => 3.63×10^3 T.U.

Microtitre plate binding assay

Two groups of third-round phage from this experiment were tested in a binding assay - ten of those that did not bind at 'backchecking' ('B-type' phage) and ten of those that did bind and were subsequently eluted ('E-type').

Phage were exposed to three dilutions of LT-B, to three dilutions of anti-cholera toxin antibody, to blocked plastic and 'naked' plastic.

On exposure to LT-B or anti-cholera toxin antibody, none of the phage tested showed differential binding by increasing dilution, with all three dilutions returning similar signals. However, four distinct phage types did emerge based on the strength of these signals and those from the wells containing blocker.

GROUP 1 : Phage B-2, B-4, B-5, B-6, B-8, B-9, E-5, E-9

These gave a high absorbance (ie. greater than 0.6) in wells that contained LT-B and blocker and moderate absorbance (between 0.3 and 0.6) in those coated with anti-cholera toxin antibody.

GROUP 2 : Phage B-1, B-7

Both of these phage gave a low absorbance reading (less than 0.3) in the wells containing LT-B, blocker and anti-cholera toxin antibody.

GROUP 3 : Phage B-3, E-1, E-2, E-3, E-4, E-6, E-7, E-8, E-10,

These gave high readings with all 3 coatings

GROUP 4 : Phage B-10

This phage gave moderate readings with all 3 of these surfaces.

The otherwise empty wells to which phage were added gave a high signals in every case. The groups are presented graphically in Figure 26.

These same phage were further tested in wells coated with ovalbumin, gelatin, skimmed milk, Tween 20, Triton X-100 and Nonidet P40.

In this case, all Group 1 phage behaved in an identical fashion, giving moderate readings against ovalbumin, low against skimmed milk, and high , as did all in Group 3.

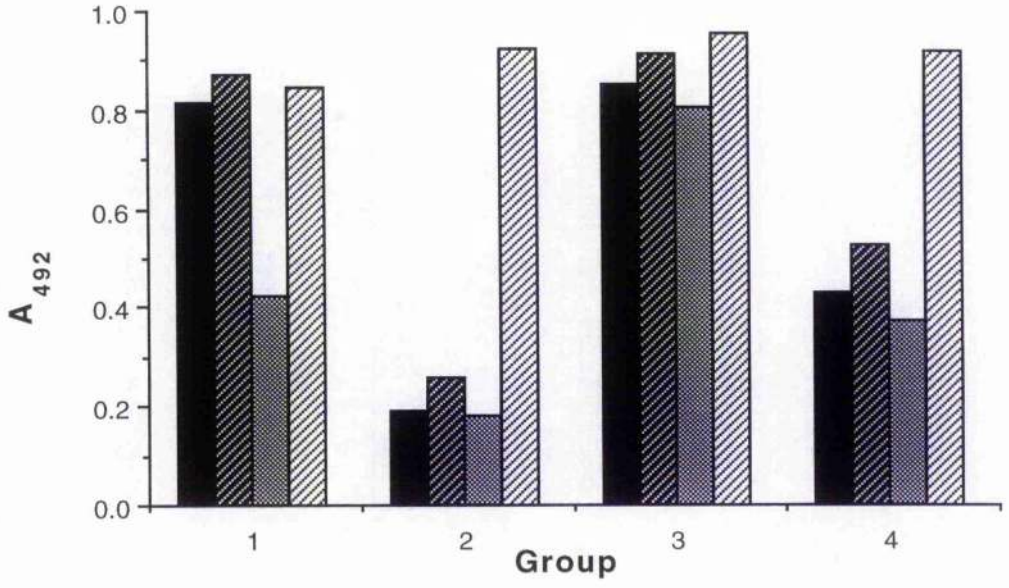


FIGURE 26 : Four groups of phage selected in the third Smith Pentadecamer panning experiment as revealed by binding ELISA.

The average absorbances obtained from exposure of each phage to three dilutions of LT-B, to three dilutions of anti-cholera toxin antibody, to one blocked plastic well and to one 'naked' plastic well were determined. Phage were then assigned to groups according to those averages. Each graph above shows the average of readings obtained from all phage in the group when bound to LT-B ■ ; blocker ▨ ; anti-cholera toxin antibody ▩ and naked plastic ▨ .

DISCUSSION

EXPRESSION, PURIFICATION AND CHARACTERISATION OF LT-B

This section of the project focused on the need to produce quantities of highly purified material, to confirm it to be LT-B, to biotinylate it, and to ensure that this process had not perturbed the properties of the protein.

Purification of LT-B was simplified by using a recombinant expression system rather than a wild-type ETEC culture. The advantage of expression controlled by an inducible strong promoter such as *tac*, which allows a high rate of transcription, is the consequently high output of gene product. Osek *et al.* (1995), using such a system, have over-expressed CT to yield a staggering 800 mg litre^{-1} in fermenter-grown cultures. The expression system used in this project, the marine strain, *Vibrio* 60 pMMB68 (Leece and Hirst, 1992), is a well-tested system for expression of LT-B itself (Amin and Hirst, 1994) and LT-B fusions (Marcello *et al.*, 1994; Loregian *et al.*, 1996).

The output of LT-B over time from this strain was monitored to determine the optimum time for harvesting the supernate. Production of LT-B started immediately upon addition of IPTG and closely mirrored cell density so that prolonged incubation of cultures was not necessary; in fact, the LT-B concentration in the supernate stabilised before the cells fully entered stationary phase (see Figure 15, page 108). This down-turn in the rate of LT-B production, despite continuing increase in cell number, may be due to the IPTG concentration in the medium becoming limiting. As IPTG induces the activity of *lac* and related promoters by binding to the repressor product of *lacI*, thereby preventing it from binding to the operator site, free concentration of IPTG would therefore decline to zero once every molecule was bound to a repressor, meaning that cells generated after this point would remain repressed, and therefore unable to produce LT-B.

Given the apparent stability of the LT-B in the medium, and the possibility of slower growth rates in some cultures, cultures were routinely harvested at 10h post-inoculation.

That the protein eluted with galactose from the affinity column was LT-B was demonstrated by SDS-PAGE, Western blotting and G_{M1} ELISA. The result of SDS-PAGE (see Figure 16, page 110) shows that the protein displays the unusual electrophoretic properties characteristic of LT-B. Firstly, as initially reported by Sandkvist *et al.* (1990), LT-B retains its pentameric structure in the presence of SDS and 2-mercaptoethanol at room temperature, disassociating to monomeric form only when boiled. Secondly, although the B subunit pentamer of LT has a theoretical Mr of 58000, it consistently appears otherwise on SDS-PAGE. First isolated by Hirst *et al.* (1983), when the apparent M_r was given as 49000, it has since been reported as "approximately 45 kDa" (Sandkvist *et al.*, 1990) and is seen in Amin and Hirst (1994) to run at an Mr of 41600. These same three sources show that the monomeric form has an M_r in the region of 12500.

The protein isolated from the affinity column in this project demonstrated a similar stability in the presence of SDS and 2-mercaptoethanol, with the gel lanes containing boiled and unboiled sample each containing a single protein band (see Figure 16, page 110). The M_r of these bands, 41500 (unboiled sample) and 13000 (boiled sample) were again as expected for the pentameric and monomeric forms of LT-B.

Further confirmation that the eluted protein was LT-B was provided by Western blotting and G_{M1} ELISA (Figures 17, page 111, and 18, page 112, respectively). In both cases, an anti-LT-B monoclonal antibody was used and in both cases a strong reaction with the eluted protein was observed. In the Western blot, this reaction was with the 41500 kDa band (the antibody does not react with monomeric LT-B).

Thus, the protein purified by affinity chromatography from *Vibrio* 60 pMMB68 culture supernate gave, by SDS-PAGE, bands characteristic of LT-B, reacted after Western blotting with a monoclonal antibody to LT-B and was bound by G_{M1}, the

ganglioside associated with LT-B binding. Thus biochemical and immunological properties unambiguously confirm the identity of the purified protein to be LT-B.

The LT-B was then biotinylated to facilitate immobilisation for biopanning. The biotinylation kit instructions required the addition of a known amount of protein therefore a protein concentration estimation step was necessary. It was possible that the galactose used in eluting LT-B from the affinity column could affect this assay (as discussed on page 50) and a variable amount of galactose could be present. Therefore, two protein estimation methods were tested, the Markwell *et al.* (1978) modification of the Lowry assay, and the method of Bradford (1976). The test used was based on the one that Markwell *et al.* themselves used to demonstrate the superiority of their technique over the original Lowry method. In their test, the BSA standard was diluted in 0 to 200mM sucrose; the readings obtained in their adapted method remaining stable whilst the Lowry method gave values that decreased. Using galactose, however, the Markwell *et al.* method performed poorly (see Figure 19, page 114), demonstrating a rise of 1 absorbance unit over the range 0 to 500mM, in contrast with the Bradford method, which gave stable readings and which was therefore used to estimate the protein content of the purified LT-B eluted from the column.

Initially, each 100ml of culture supernate applied to the column yielded between 700 and 1000µg of purified LT-B, or 47 to 70% of the 14.8mg litre^{-1} produced by Amin and Hirst (1994), using *Vibrio* 60 pMMB68 culture supernate and their own purification method (see page 48). This was considered an acceptable yield. However, after purifying 1500ml of supernate, the total produced was just over 2mg, or just 10% of the yield of Amin and Hirst (1994).

SDS-PAGE analysis of every fraction eluted from the column showed the LT-B band intensity to have remained constant throughout the whole process, whilst the intensity of the LT-B band in the material which failed to absorb to the column did not increase. The possibility that the column eventually became saturated with bound material is thus ruled out. In practice, the column was regenerated with 1M galactose after each run, and was routinely purged with 6M guanidine.

The poor yield of LT-B may have arisen through proteolytic degradation. As discussed earlier, protease is exported by *Vibrio* 60 pMMB68. Loregian *et al.* (1996) reported attempts to produce protease-deficient mutants of *Vibrio* 60 pMMB68, such is the interference of the enzyme in production of fusion proteins. Although the protease should be removed from the LT-B by the galactose affinity column, it would have been present in the initial 2 litre batch of *Vibrio* 60 pMMB68 supernate. This was stored as 100ml aliquots, at -20°C, until required. These aliquots were each processed by thawing and clarification by passage through 0.45µm nitrocellulose filters before being run by gravity onto the affinity column. Each of these processes was carried out at room temperature and in all took around 2h to complete. By comparison, the initial tests were performed on 100ml batches specifically produced for the purpose. It is therefore likely that proteolytic degradation occurred at this point, giving the smaller than expected final batch of LT-B.

Another possibility, that the LT-B spontaneously bound to the dialysis tubing at the concentration stage, might also be considered. That proteins adhere to surfaces is well accepted but this is usually a problem at low concentrations and in the short exposure time, the removal of 80% of the total protein seems unlikely.

The stock LT-B was successfully biotinylated yielding 4.48 moles of biotin per mole of LT-B, which is within the range recommended by the manufacturers of the biotinylation kit. As this form of biotinylation preferentially attached the biotin to lysine residues, of which there are nine in the LT-B molecule (see Appendix 9, page 193), including $\mathbf{K}^{91}$ in a critical position in the binding pocket (see Figure 5, page 40) it was then essential to check that biotinylation had in no way corrupted binding ability. G_{M1} ELISA was therefore performed using, as probes, an anti-LT-B monoclonal antibody, to demonstrate that LT-B^{biotin} could bind to G_{M1}, and the biotin-binding, peroxidase-conjugated antibody, 'Extravidin', to demonstrate the presence of biotin in the bound material (See Figure 20, page 116). The finding that biotinylation did not affect the ganglioside-binding property of LT-B can be contrasted with that of Khine and Lingwood (1994). In their investigation into the

internalisation of verotoxin (shiga-like toxin), they found that biotinylation of the B-pentamer, by sulfo-NHS, reduced binding to its ganglioside receptor (G_{b3}) and prevented internalisation of the toxin. The verotoxin B-subunit has 69 amino acids of which 5 are lysines, as compared with LT-B which has 103 residues including 9 lysines, and it was concluded that biotinylation of one of these lysine residues, K^{53}, was responsible for the deleterious effects. Although the differences between the B subunits of verotoxin and LT reduce the relevance of Khine and Lingwood's finding to this project, it nonetheless highlights the need for an assay to confirm that no such problem existed here. Having proved this, G_{M1} ELISA, which gave a semi-quantitative result, was thereafter used to assay LT-B^{biotin}.

Whilst it would have been expected that the highest concentrations of LT-B would have bound the greatest quantities of anti-LT-B, and that this would then have bound the most HRP-conjugated antibody to give the highest ELISA signal, the result of some ELISAs obtained in this project (e.g. that seen in Figure 18, page 112) suggested that this was not always the case. Rather, the wells containing the most concentrated LT-B gave readings which were sub-optimal in comparison with more dilute samples. Steric hindrance caused by local overcrowding at the assay surface is the most likely explanation for this phenomenon, with closely packed 118.87 antibody being less available to HRP-conjugated anti-mouse antibodies than more widely-distributed immunoglobulins.

SMITH HEXAMER LIBRARY

The aim of this section of the work was to 'pan' streptavidin-immobilised LT-B^{biotin} with the Smith hexamer library, to elute off and characterise any phage that bound and finally to assay their binding specificity.

Cultures of *E. coli* K91kan grown to A_{600} in the range 1.25 to 2.5 proved optimally susceptible to infection (see Figure 21, page 118). Smith, however, reports infectivity of fd-tet-derived virions to be "not very reproducible from one experiment to the next". Since the assay is inconsistent, titering was performed only after the four

rounds of panning were completed in an attempt to minimise this effect: all of the applied and eluted phage were enumerated in one experiment using a single batch of K91kan cells.

Use of the P + LS strategy (in which phage are added to streptavidin-bound LT-B^{biotin}) in the first round was intended to ensure that the target molecule, a peptide capable of mimicking the conformation of G_{M1}, was present at maximum density, facilitating recovery of a high number of binding phage. Subsequent rounds pursued the alternative PL + S strategy (in which phage are mixed with LT-B^{biotin} prior to exposure to streptavidin-coated plastic). Biopanning with acid elution did show an 805-fold enrichment (see Table 6, page 120), however, this is little more than 5 times greater than the 154-fold enrichment obtained when phage were eluted with TBS alone. These yields are clearly far below the 10^5-fold enrichment in a single round reported when Scott and Smith (1990) used the same library.

The enrichment of phage eluted with TBS alone is probably due to two factors. Firstly, having started each round with the addition of between 10^{10} and 10^{12} phage, it is not surprising that at least some remain behind in the tube, despite the ten washes employed. It is possible that, by increasing the stringency of the washing, all free phage could be removed, however to do so might be to risk losing some of those specifically selected but bound with low affinity. Compounding this problem are the properties of the peptides carried by the phage. If there is a limited number of phage in a library of billions capable of attaching, *via* their peptide fusions, to LT-B^{biotin}, it is probable that there are also phage which may bind to the tube for other reasons. For example, Devlin *et al.* (1990) have reported the isolation of streptavidin-binding phage and Adey *et al.* (1995) have isolated phage which bind to the plastic matrix from two separate random peptide libraries, one with a 36 and the other with a 22 amino acid insert. The latter authors suggest that skimmed milk is a more effective blocking agent for plastic surfaces than BSA but it was not used here as it is known that lactose can occupy the G_{M1}-binding pocket on LT-B (Sixma *et al.*, 1992). In addition, phage carrying peptides with an affinity for sites on LT-B^{biotin} other than the

receptor-binding pocket may be eluted. Therefore phage which are acid- or TBS-eluted from sites other than the G_{M1}-binding site will have been just as capable of being amplified, and therefore of being present in increasing numbers in subsequent rounds, as those which were sought specifically. It was for this reason that G_{M1}-elution was included, to provide a specificity lacking in acid and TBS, which will remove phage binding to the surface with very low affinity.

The binding assay, an ELISA-type test based on applying the phage directly to the plastic, was chosen in order to preserve the stocks of the phage for any further work that may have developed had the results proved more satisfactory. By adding the phage first, unbound virions could be collected and later recovered uncontaminated, an option that would not have been available had the phage been added to immobilised LT-B^{biotin}. In the event, phage eluted by acid, G_{M1} or TBS proved unable to capture LT-B^{biotin} at levels in excess of controls.

Perhaps the most compelling explanation for the disappointing results obtained in this experiment relates to the library itself. When originally produced by Scott and Smith, it consisted of 1.3×10^{14} phage. This is well in excess of the number required to guarantee that all possible hexapeptides are present, which is X^y, where X is the number of codons available and y is the length of insert. In this case, that would be 32^6, or 1.1×10^9. However, the material supplied by George Smith was an amplification of a portion of the original library. On receipt, this was amplified again. The composition of the library stock, used in this work, may well have differed from that of the original Smith library due to this processing, as it is possible that each phage may be differently affected by the fusion peptide that it carries. Any phage with a fusion which slows infection rate or which inhibits growth of its host will be under-represented in the final amplification relative to those which grow at normal, or even enhanced rates. Cwirla *et al.* (1990) examined this phenomenon. They produced their own random hexapeptide library, fused to pIII in fd-tet employing the NNK bias, using Smith's methods and cultures. However, on sequencing 52 randomly selected clones from the library, they found that there were around 50 % more G than T

residues at the third position (59 % G versus 39% T). This may have arisen from the synthesis reactions which produced the random inserts. It might, however, be that phage carrying those amino acids that are produced from codons containing a G in the third position gained some form of growth advantage over those with codons ending in a T, perhaps for reasons associated with availability of the appropriate tRNAs (as discussed on page 12).

The possibility must therefore exist that the 'Amplified Library Stock' simply contained an insufficient range of different peptides, and that peptides which could bind to the receptor-binding site of LT-B were not present in the library.

The problem posed by the 'double-signal' must also be addressed. The super-imposition of the wild-type sequence over the insert sequence made reading the sequences difficult (Figure 22, page 121 demonstrates this problem) and had an unknown effect on the whole conduct of the experiment. One explanation for this problem is infection of the K91kan stock by wild-type fd-tet prior to its use in amplifying the panned phage. This possibility can probably be dismissed, as protein V (see page 9) spontaneously coats any single stranded DNA in the cell and 'escorts' it to the membrane for assembly. Any single-stranded DNA entering the cell on super-infection would be similarly coated and packaged, thus preventing it from making replicative forms (RF) and from being further amplified. In spite of these considerations, experimental evaluation of super-infection was pursued. Firstly, an overnight culture of K91kan was mixed with an equivalent amount of 'top' agar (0.5% agar in water) and poured over an LB agar plate. After incubation, no plaques were observed. Some of the overnight culture supernate was then mixed with PEG/NaCl, in order to precipitate any phage that were present. However, when spun down there was no visible phage pellet. A colleague, Calum McCafferty continued this work, by processing and sequencing the contents of this tube. No extension / termination products could be detected on sequencing gels.

In addition, a new K91kan stock was obtained from Russell Thompson, Division of Virology, Glasgow University. Running some simple panning experiments using this

culture gave a similar result, confirming that the problem was indeed associated with the phage and not the bacterial host.

A second possible explanation is that wild-type phage were present in the library stock, dating from the time that the library was first made. This is possible but in order to appear so frequently among the sequences they would have to have been selected and amplified through the biopanning procedure, contrary to the principle of affinity-selection of adhering phage. Additionally, host cells infected with the wild-type phage would then have had to contaminate each of the single colonies chosen for amplification prior to sequencing. Again this is unlikely.

Another idea that was explored was that the wild-type sequence was the result of recombination events leading to the excision of the insert sequence from some phage during amplification in the host cell. Any wild-type phage produced by recombination, during the amplification stage between biopanning rounds, would be selected against in the next round. However, any that were produced during amplification of the fourth round single colony would be sequenced. The product of the gene *recA* is a protein of M_r 38000 that helps to modulate recombination by recognising and pairing homologous sequences between single-stranded and double-stranded DNA, the process that results in the re-arrangement of DNA (Kurumizaka and Shibata, 1996). The possibility therefore existed that, during amplification, deletion of the insert occurred with sufficient frequency through recombination between flanking homologous sequences. This theory was tested by growing some fourth round phage in *E. coli* DH5F'IQ , a mutant deficient in the *recA* gene, and in which cross-overs would occur with much reduced frequency. Amplification in DH5F'IQ was carried out exactly as normally performed in K91kan. The resulting cultures were then spread to single colonies, which were individually picked, phage amplified and sequenced. The resulting data still showed the presence of wild-type sequences. It was therefore reckoned unlikely that recombination accounted for the double signal.

Regardless of the anomolies encountered in sequencing, evidence for positive selection of phage from the library was strong; as shown in Table 7, page 122, the sequences of the peptides carried by phage which were acid eluted showed one clear consensus and a number of other common features. Given the diversity of the peptide library, it is not surprising that phage with a specificity for one or more components of the panning surface were eluted. The identification of two separate base sequences coding for **PWPWLG** rules out the possibility that they arose from a single clone which was propagated with high frequency. That there were three other clearly related clones (with two different sequences), served to reinforce this point.

Another aspect of the groups of peptide sequences relates to their chemistry : the Group 1 acid-eluted phage, for example, consist entirely of hydrophobic amino acids, without a single charged residue. Amongst clones classified as Group 2, residues 1, 2 and 6 are invariably hydrophobic and base 3 is invariably charged (i, ii, iv and v are acidic; iii is basic). Only residues 4 and 5 show variation, with a mixture of acidic, basic, hydrophobic and hydroxylated species all appearing in these positions. The Group 3 amino acids are from unrelated families.

Further evidence for positive selection can be derived from a comparison of the actual occurrence of amino acids in the acid eluted phage groups 1 and 2 versus the occurrence expected if selection was completely random (see Table 11). Although the sample number is low, it is nonetheless obvious that **F**, **P**, and **W** are unequivocally over-selected, by approximately 3-fold, 4-fold and 7-fold respectively and that **T** is underselected (no appearances when at least 6 were expected).

Despite the strength of the argument for specific selection, the binding assay gave little evidence to suggest that those phage enriched by biopanning could capture LT-B. It is also worth noting that they bore no resemblance to the **AEYHN** loop (reported by Merritt *et al.*, 1995) borne on LT-B and thought to be capable of interacting with the G_{M1}-binding site of adjacent pentamers (previously discussed on page 39). In fact, **A** and **E** are under-represented amongst the acid-eluted phage, **Y** and **H** are totally

AMINO ACID	NOMINAL FREQUENCY	NOMINAL OCCURRENCE	OBSERVED OCCURRENCE
A	0.065	6.63	3
C	0.032	3.264	3
D	0.032	3.264	0
E	0.032	3.264	1
F	0.032	3.264	13
G	0.065	6.63	10
H	0.032	3.264	0
I	0.032	3.264	1
K	0.032	3.264	3
L	0.097	9.894	12
M	0.032	3.264	4
N	0.032	3.264	4
P	0.065	6.63	22
Q	0.032	3.264	0
R	0.097	9.894	1
S	0.097	9.894	1
T	0.065	6.63	0
V	0.065	6.63	2
W	0.032	3.264	22
Y	0.032	3.264	0

TABLE 11 : Expected and actual occurrence of amino acids in the clones from the Smith hexamer library.

The nominal frequency is obtained by dividing the number of codons for any amino acid by 31, the number of coding triplets in the NNK genetic code. Nominal occurrence is derived by multiplying the nominal frequency by 102, the number of amino acids in the 17 hexamer inserts.

absent and **N** is present in no more than the expected frequency. The phage selected from this library also fail to display the sequence characteristics (the consensus **HPQ**), published by Devlin *et al.* (1990), for streptavidin binding. As well as **H**, **Q** was absent from the 17 peptide sequences characterised representing notable under-representation. It is more plausible that the acid-eluted phage are capable of direct interaction with plastic. Adey *et al.* (1995) described the defining characteristic of this property as being a high frequency of **Y** and / or **W** residues. Although **Y** is absent from the peptides selected by biopanning, **W** is over-represented by 6.7-fold (observed frequency/ expected frequency). However, Adey *et al.* comment that the presence of **W** or **Y** residues is not of itself indicative of a peptide potentially capable of binding to plastic and go on to report that the bulk of the **Y** and **W** residues that they observed were towards the amino-termini of their inserts. They suggest that this may reflect the reduced steric hindrance from pIII in this location. They also mention that they failed to observe plastic binding when hexapeptide or **C**-constrained octapeptide libraries were used, due, they supposed, to the short length of the hexapeptide and / or the artificially shortened length of the octapeptide under **C**-constraint. As further evidence for this theory, they also cite the absence of **C**s from all of their plastic-binding clones. In relation to this, it should be noted that the **W** residues carried by Group 2 acid-eluted clones are all at the extreme amino-terminus, whilst in Group 1 peptides they are flanked by **P** residues. It is possible that such alteration of the conformation of these Group 1 peptides could equally 'present' the **W** to the plastic in such a way as to overcome any steric hindrance from pIII. It should also be noted that the plastic tubes used in this work were selected specifically because of their highly absorptive characteristic, whilst Adey *et al.* used standard polystyrene or polyvinyl chloride microtitre plates. This therefore raises the possibility that a single, amino-terminally located, or two, 'proline-presented', **W** residues are able to bind the phage to a Maxisorp tube with sufficiently high affinity that acid is necessary to elute them.

Set against this, however, is the presence of the **PLPLAG** and **CFNMPF** peptides in Groups 1 and 2 respectively. The absence of **W** from these otherwise obviously related peptides makes plastic binding, as defined by Adey *et al.* rather less likely. However, given that the Group 1 phage peptides consisted entirely of hydrophobic residues, attraction to the matrix remains a possibility.

Finally, there is little evidence in these results to suggest that the sequences obtained are characteristic saccharide mimics. Both Hoess (1993) and Scott *et al.* (1992) emphasised the potential of **Y** residues as mimics of saccharide rings. These residues were completely absent from peptides borne on the eluted phage.

LUZZAGO NONAMER LIBRARY

The Luzzago library was of interest as it provided a contrast to the Smith hexamer library in several ways. The randomised DNA was inserted into gene VIII as opposed to gene III, the insert was potentially constrained through the formation of a disulphide bond by the **C** residues present at each end of the random peptide and this library was based on a phagemid-borne copy of gene VIII and thus was of type 8 + 8. Although the library phage supplied by Alfredo Nicosia were clearly infective, allowing *E. coli* XL-1 to grow on agar including ampicillin, the results obtained from sequencing the resultant clones indicated that there was a problem in the construction of the library.

Panning produced substantial numbers of eluted phage at each round, although there was little evidence of the amplification expected if phage were being specifically selected. After the first two rounds all clones were collected, resuspended in broth and co-infected with helper phage to produce the phage used in the subsequent round of panning. After the third round, the clones were individually picked and plasmid DNA isolated and sequenced.

Analysis of the sequences showed that, in most cases the gene VIII sequence was substantially as expected outwith the randomised insert (see Figure 23, page 127). The substitution of a TCT for TCC at S^{-14} is silent. It is likely that the apparent

deletion at position 19 of the mature peptide (CC instead of CCC) is a compression artefact : this is caused when areas of the DNA that are rich in a single base, particularly G or C, are not fully denatured by the electrophoretic process, resulting in the normally uniform spacing of bands on the gel being extended or compressed (Sequenase Protocols, 1994). On examining the autoradiograph in the area concerned (see Figure 27) it can be seen that the lower band is somewhat thicker than the upper thereby suggesting that a compression was present.

In analysing the randomised inserts, some of which are shown in Table 8 (page 129), it is less plausible to look to compression as an explanation for the fact that only 26 bases could be seen in every clone examined. Although G and C are the most numerous bases, constituting 60% of the inserts, they lie in a variety of positions in the sequence; it would be expected that at least one of them would resolve correctly, particularly as runs of up to four C's, and six G's, both above and below the random insert, were consistently and satisfactorily denatured.

There were also three TBS-eluted clones that had a sequence, running from position 4 of the mature protein back into the signal peptide, that bore no relationship to that predicted (see Figure 23, page 127). The actual departure from predicted sequence begins at the first base of the codon for residue four of the mature protein (AAA rather than GAA). This forms the 5' end of the inserted oligonucleotide, raising the possibility that these clones are derived from aberrant restriction / ligation events. A small part of the base sequence immediately adjacent to this site, covering codons -4 to 4 on Figure 23, (AAAAGGTATCCGAGGCGGGGGG) was analysed using the database search tool 'BLAST' (Basic Local Alignment Search Tool), at the National Center for Biotechnology Information (reached through website address http:/www.ncbi.nlm.nih.gov/BLAST/). A search for the sequence and its complement revealed similarity to the origins of replication from various plasmids

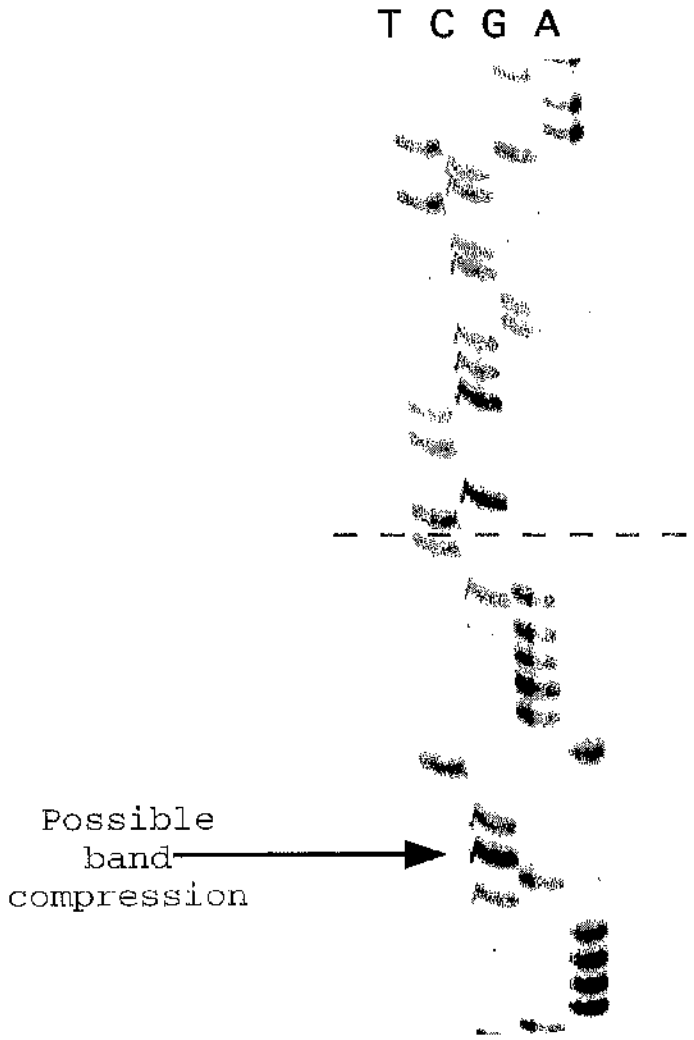


FIGURE 27 : Autoradiograph of sequencing gel prepared with clones from the Luzzago library.

A possible compression is labelled. Nucleotides are shown above the relevant lanes. The sequence of the random insert runs upwards from the dashed line.

including ColE1. The Luzzago phagemid is a ColE1 derivative. A comparison of some 70 bases of the ColE1 sequence with that gathered from clones of the Luzzago library (see Figure 28) clearly shows complementarity.

The following conclusions about events during production of the library can be drawn from the sequencing and BLAST search results.

a) The randomised insert was correctly prepared, as the cohesive ends for *Eco*RI and *Bam*HI were present.

b) *Bam*HI digestion of PC89 and ligation of the insert at that site proceeded as intended, as evidenced by the presence of the full recognition sequence for this enzyme (see Figure 23, page 127).

c) *Eco*RI digestion of PC89 did not occur as anticipated. The *Eco*RI cohesive end of the insert is not adjacent to the base G, indicative of insertion at the *Eco*RI restriction site, but rather lies within the ColE1 ori. The actual position at which the ColE1 ori was cut was between bases 417 and 418. The bases surrounding this cut were :

411 CAGGCGT // TTTTCCA 424

however, they bear little resemblance to the restriction site of *Eco*RI suggesting that the cut was non-specific. The net effect was the elimination of a section of DNA that included three amino-terminal amino acids and the signal sequence of gene VIII and the *lac* promoter.

The mechanism of this error is summarised in Figure 29. That three of the thirty clones sequenced should carry this error indicates that it occurs commonly.

A report by Beekwilder *et al* (1999) is worthy of consideration at this point. They had manufactured a library of phage carrying variants of a protease inhibitor. On panning against trypsin, they failed to recover a single phage with any binding activity. On analysis, the phage in their amplified library proved to be carrying either deletions or amber stop codons. It seems that those variants that carried nonsense mutations in the gene encoding the peptides attached to pIII, were entirely favoured over all variants that carried an extended pIII. The authors attribute this phenomenon to differential growth rates; the nonsense-peptide-carrying phage being able to outgrow the others to

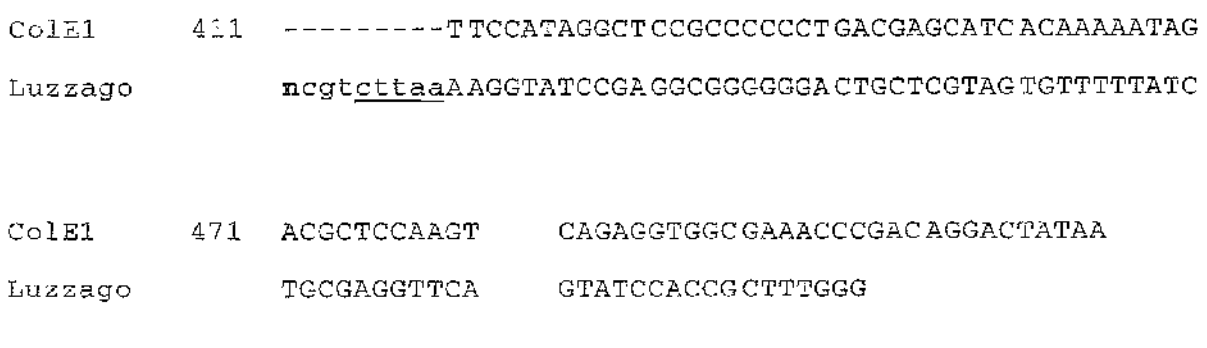


FIGURE 28 : Comparison of unusual Luzzago sequence with the ColE1 origin of replication

The last few bases of the expected Luzzago sequence, are shown in lowercase including, underlined, the *Eco*RI cohesive end.

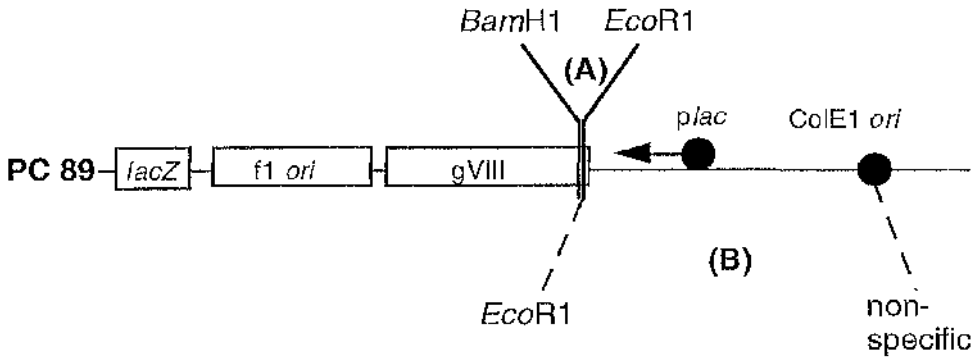


FIGURE 29 : Origin of the unusual sequences identified in clones from the Luzzago library.

The random nonapeptide and its flanking residues should have been inserted at (A), between the *Eco*RI and *Bam*HI sites. It appears that the actual insertion was between *Bam*HI and a non-specific site in the ColE1 *ori*. As a result the DNA designated (B), between the *Eco*RI-cut end of the insert and the non-specific site, was lost.

such an extent that their numbers became insignificantly small in the amplified library. de Bruin *et al* (1999), when panning with a library of synthetic heavy and light chain variable regions, observed that incomplete inserts predominated after 3 rounds. They suggest that these did not arise by deletion of full length inserts but rather that they were present in small numbers in the original library and were subject to a sufficiently favoured growth rate to allow them to overwhelm slower growing phage .

It could well be that a similar occurrence took place during the production of the Luzzago library, or during the amplification of the portion that I received, with phage carrying mutant inserts overwhelming correctly-formed phage. The fact that phage were produced despite this, and could be precipitated out in the absence of pVIII is easily explicable: employing the phagemid system, a 'wild-type' capsid could still be formed consisting entirely of helper phage proteins. Ampicillin resistance could be contributed by the phagemid regardless of a frame shift in gene VIII. Although attempts were made to address this issue, the data produced was ambivalent and the problem remains unresolved. In view of the deficiencies of this library, the question whether constrained peptides can effectively mimic carbohydrates unfortunately remains unresolved. It is worth noting, however, that Hoess (1993), in the section of his review of phage display that dealt with this topic, comments that none of the reported successes involved constrained libraries. On the other hand, Ladner (1995), suggests that constrained peptides are more suited as mimics of discontinuous epitopes than continuous : the complexity of the interaction between G_{M1} and LT-B is such that the carbohydrate moieties resemble a discontinuous epitope. It might be therefore imagined, in this context, that a constrained peptide could have some advantages over one in which conformation was not restricted.

SMITH PENTADECAMER LIBRARY

This Smith pentadecamer library represented a further variation on those already tested. Although, like the Luzzago nonamer library, it carried a fusion to protein VIII,

in this case it was a type 88 phage, with two copies of gene VIII, one wild-type and one synthetic carried in the viral genome. Addition of IPTG was required to induce production of the recombinant gene VIII, as it was under the control of the *tac* promoter. The library, as supplied, was shown to contain around 5×10^{12} transducing units. This represents only 1.5×10^{-5} % of the 3.28×10^{19} possible combinations of 15 amino acids, assuming each clone to be unique, but each 15 amino acid insert will contain 10 different hexamers, or 13 different trimers, (see page 12) providing a compensating level of diversity.

The first biopanning experiment was carried out according to the methods recommended by Russell Thompson. Biopanning proceeded as expected, and amplification of the number of eluted phage was observed. Phage numbers were calculated using the rapid method recommended by Russell Thompson, which counts only the number of eluted phage, but not the number applied in biopanning. As a result, the percentage of applied phage that were eluted was not calculable. However, an increase in the number of eluted phage, from 10^3 in the first round to 10^5 in the third, was in line with the results routinely achieved with this method in the Division of Virology.

Sequencing of phage selected in this experiment presented none of the problems previously encountered with the Smith hexamer or Luzzago libraries. Aside from the **H** residues in position 2, the other most noteworthy features of the sequences produced are best seen when the characteristics of each amino acid are related to the position in which they occur in the insert. In order to do this, amino acids were grouped according to the properties of their side chains, as shown below. The value to the right shows the percentage of all peptides that each group would constitute, allowing for the differential in number of codons for each amino acid.

i	Acidic side chain	**(D, E)**	6.5%
ii	Basic side chain	**(K, H, R)**	16%
iii	Uncharged polar side chain	**(C, N, Q, S, T, Y)**	30%
iv	Non-polar side chain	**(A, F, G, I, L, M, P, V,W)**	47.5%

The actual composition of the 10 peptides analysed here was as follows :

i 1.8%

ii 22.6%

iii 29.4%

iv 46.2%

Aside from under-representation of amino acids with acidic side chains and over-representation of those with basic side chains, these figures do not provide any significant indication of a selective effect through biopanning. However, analysis of the distribution by position in the peptide reveals a quite different picture. Table 12, shows the distribution of each group of amino acids according to position. It is clear from this table that the peptides selected by biopanning are predominantly basic at their amino-termini (residues 1 to 3). This is not only due to the presence of **H**; six of them also feature **R**, whilst clone 10 features both **R** and **K**. Hydrophobic residues are substantially under-represented in this region, appearing only once in the first position, not at all in the second and four times in the third.

The occurrence of amino acids in the second and third regions (residues 4 to 9) are, aside from the lack of acidic amino acids, more consistent with random distribution, suggesting that behind the charged amino terminus, the fusions to pVIII have an uncharged, predominantly hydrophobic core. Region four is different again, with more common occurrence of residues with polar side chains, and fewer hydrophobic amino acids. The final region is the most hydrophobic, least charged of all, with only two of the inserts encoding charged residues. Given that this is the region closest to the capsid body, and most closely associated with the native pVIII, perhaps this lack of charge minimises interference with the coat protein.

The most striking individual feature of all the inserts is the **H** in the second (or in one case, first) position of each of the clones sequenced. In all cases, as with the other **H** that featured (at position 12 in sequence 10), the codon was CAT which, according to the codon usage table given in the phage display laboratory manual (Kay *et al.*, 1996),

RESIDUE			REGION		
TYPE	1	2	3	4	5
i	0	0	6	3	0
ii	60	17	10	20	6
iii	23	26	27	37	34
iv	17	57	57	40	60

TABLE 12 : Distribution of amino acids occurring in the inserts carried by clones from the Smith pentadecamer library.

The percentage ocurrence is shown.

Regions: **1**, residues 1 to 3; **2**, residues 4 to 6; **3**, residues 7 to 9; **4**, residues 10 to 12; **5**, residues 13 to 15. Residue type **i**, acidic amino acids; **ii**, basic; **iii** , uncharged polar; **iv**, non-polar.

features four times more frequently than CAC.

Obviously, the amino acid sequence of any peptide selected in this experiment will reflect its affinity for the panning surface. It is apparent from the results of the Microtitre plate binding assay that these peptides bound streptavidin rather than LT-B^{biotin} (see Figures 24, page 133, and 25, page 137). The signal from the well coated with streptavidin-bound LT-B^{biotin} was occasionally higher than that from the well coated with streptavidin alone, but in no case did the difference exceed 17% and in every case the streptavidin signal was practically double the background level. Were the phage specifically binding to LT-B, one would have expected signals from wells coated only with streptavidin to be at background levels. Given the high level of reaction to streptavidin, it must be concluded that it is to this that the phage were binding.

The results of the 'colony lift' binding assay confirmed those obtained in the microtitre plate assay. Phage selected by panning streptavidin-bound LT-B, once adsorbed onto a nitrocellulose membrane, showed no discernible affinity for the LT-B to which the membrane was then exposed. By contrast, a piece of membrane impregnated with G_{M1} and exposed to the same LT-B solution, bound anti-LT-B antibodies effectively. When exposed to streptavidin solution and probed with biotinylated peroxidase these same phage revealed a very weak affinity for the streptavidin. A similar weak affinity for streptavidin was demonstrated by the phage selected by panning streptavidin-coated plastic. Again the effect was negligible, but the positions of the faint traces that were seen did correspond with those of colonies on the agar.

Thus phage selected by panning the pentadecamer library on streptavidin-bound LT-B demonstrated, in both binding assays, a weak affinity for streptavidin, but no affinity for LT-B.

To consider the nature of the streptavidin binders that were selected, it is useful to bring together the sequencing data with that from the microtitre plate binding assay. Phage 2 and 3 gave the strongest signals in the assay and both carry an uncharged

residue (**N** and **Q** respectively) as their first amino acid, followed by **H** and **R**. The next-highest signals were obtained from clones 1 and 5, which carry **S H R** and **A H R** respectively, whilst the third moderate binder, number 7, like number 2, also commences **N H**, but in this case **S**, rather than **R**, is the third residue. Among those clones that gave weak signals, numbers 4, 6 and 10 all commence with **R H**, followed by a non-polar residue (**L** and **G** respectively) in the case of the former two, and by **K** in the case of the latter, and clone 9, like clones 2 and 7, commences **N H**, followed by **W**. Clone 8, which gave the weakest signal of all, had the sequence in which **H** was the first, rather than the second, residue.

Having established that low-affinity streptavidin-binding phage were selected, it is pertinent to compare the sequences carried by these phage with those previously identified as streptavidin binding motifs. This work, which has been briefly dealt with earlier (see page 58) includes that of Devlin *et al.* (1990), who deliberately set out to isolate streptavidin binding phage. In doing so, they provided the first evidence that phage display technology could be used to identify amino acid sequences that recognised proteins with no previously known affinity for peptides. The peculiarities of their library were discussed earlier (see page 56), but, in brief, it was created in M13 and consisted of NNS-biased pentadecamer fusions to pIII, that were extended away from the phage on a poly-proline linker. In the paper, they describe the inserts from twenty phage, carrying nine different sequences. These are reproduced in Table 13. The most significant feature of these sequences is the **HP** motif that appears in each of the nine sequences presented, and which represents the core consensus. The amino acid **Q** is appended to this core in seven sequences, followed by **N** in three instances.

Drawing these findings together with my own, the most obvious similarity is in the appearance of **H** in each sequence. The residues **P** and **Q** which also feature strongly among sequences reported by Devlin *et al.* were both over-represented in this study (see Table 10, page 132) and, indeed, five sequences featured all three of these amino acids distributed throughout the insert. In clones 2, 4 and 10 these amino acids fell in

ISOLATE	FREQUENCY	SEQUENCE
i)	3	SDDWWHD HPQN LRSS
ii)	1	MLWYSPHSFS HPQN T
iii)	1	SWWWLSW HPQN TKELG
iv)	5	ISFENTWLW HPQF SS
v)	1	LC HPQF PRCNLFRKV
vi)	2	PC HPQY RLCQRPLKQ
vii)	2	QPFL HPQG DERWYMI
viii)	1	ALCCLSSP HPNG AIF
ix)	4	LN HPMD NRLHVSTSP

TABLE 13 : Amino acid sequences carried by twenty streptavidin-binding phage.

After Devlin *et al.* (1990).

the order **HPQ**, but in no instance were they contiguous. In clone 10, the motif **PHQQ**, in positions eleven to fourteen, may be significant. In clone 2, although **H** and **P** were close together (positions two and four), **Q** was in twelfth position. In clone 4, however, the residues appeared in the second, fourth and seventh positions.

Positional effects seem to be strong. In clones that were characterised, the amino terminal featured both **H** and **Q** and the former residue also appeared nine times in position 2. The fourth residue was **P** in two sequences and **Q** in one; the sixth position twice features **P** and the seventh features **Q** twice, and **P**. Overall, of thirty-two occurrences of the three amino acids, twenty of them are in the first seven positions. Although not highlighted by Devlin *et al.*, in six of their nine sequences there is a basic residue (**R** five times and **K** once) at one, two or three residues to the carboxyl-terminal end of the **HPQx** core. A major difference is in the position of the **H**, which in my sequences was always at the amino-terminus of the insert, but which, in the sequences described by Devlin *et al.* is in any position from 3 to 11. They make particular note of the position of **H** in sequence ii), suggesting that the consensus motif may bind even when placed very close to the phage capsid. It must, however, be remembered that these inserts were succeeded (at the carboxyl terminus), by six **P** residues, thereby making **H** the eleventh amino acid from the capsid body in sequence ii), and between the twelfth and nineteenth in the others.

The appearance of **H**, **P** and **Q** in my sequences does not correlate with signals from the binding assay. Clone 3, which performed best in the assay, does not feature **P** at all, nor does clone 1, (intermediate signal strength). Clones 7 and 5 gave similar assay results to clone 1, yet caried **H** and **Q** in adjacent positions and one residue apart, respectively.

Further work on streptavidin binding was carried out by Kay *et al.* (1993), who chose this protein to determine the utility of a library consisting of random 38-amino acid fusions to gene III. On sequencing the inserts from forty-nine of their phage, they found thirteen different sequences, each of which featured the **HP** motif, followed by **Q** or **M** and then by a non-polar residue. The most extreme carboxyl-terminal location

for the **H** was, in a single case, six places from the capsid, but overall, only seven of the forty-nine carried **H** closer than twelve places from pIII. Six of these thirteen sequences, representing the inserts from twenty-nine phage, also carried the motif **PG**, in various locations to the carboxyl-terminal side of **HP**. Ten further streptavidin phage that were characterised carried two inserts that lacked **HPQ** and were otherwise dissimilar save for the **PG** motif. Finally, assays were performed to confirm that both the **HP** *etc.* and the **PG** type phage were binding to the biotin-binding site of the streptavidin. A detailed comparison of my sequences with those of Kay *et al.* will not be necessary, given the similarities with the Devlin *et al.* sequences already discussed, except in regard to the **PG** motif. These residues appear in positions four and six respectively in my sequence 4, in positions six and ten of my sequence 8, but do not feature together in any other sequence. Unlike **P**, **G** is neither over- nor under-represented among these sequences.

Streptavidin was also employed to select phage from the library of McLafferty *et al.* (1993). This gene III library (see page 56) included a 'variegated' (*ie.* not fully randomised) insert. As with Devlin *et al.* and Kay *et al.*, the motif **HPQ** was again revealed as a consensus sequence, featuring in all the fourteen binding phage that were tested, in each case followed by **F**. This motif was, in every case, placed between fixed **C** residues, on either side of which was an unrelated amino acid.

Devlin *et al.* conclude their paper by highlighting the similarity between the structures of biotin and histidine, suggesting that the shared nitrogen- and carbon-containing ring may account for the presence of this amino acid in all of their selected phage. Figure 30 demonstrates this similarity. Crystallography studies into the mechanism of the interaction of the **HPQ** peptide with streptavidin have been published (Weber *et al.*, 1992; Katz and Cass, 1997). The structure reported by Weber *et al.*, shown in Figure 31, involved the heptapeptide **FSHPQNT**. The **H** and **Q** residues bind, *via* a captured water molecule, to **S**, **D** and **T**, whilst the **P** facilitates this binding by controlling the local conformation of the peptide. No reaction with any other amino acids is shown, although both reports suggest, from electron density mapping, that

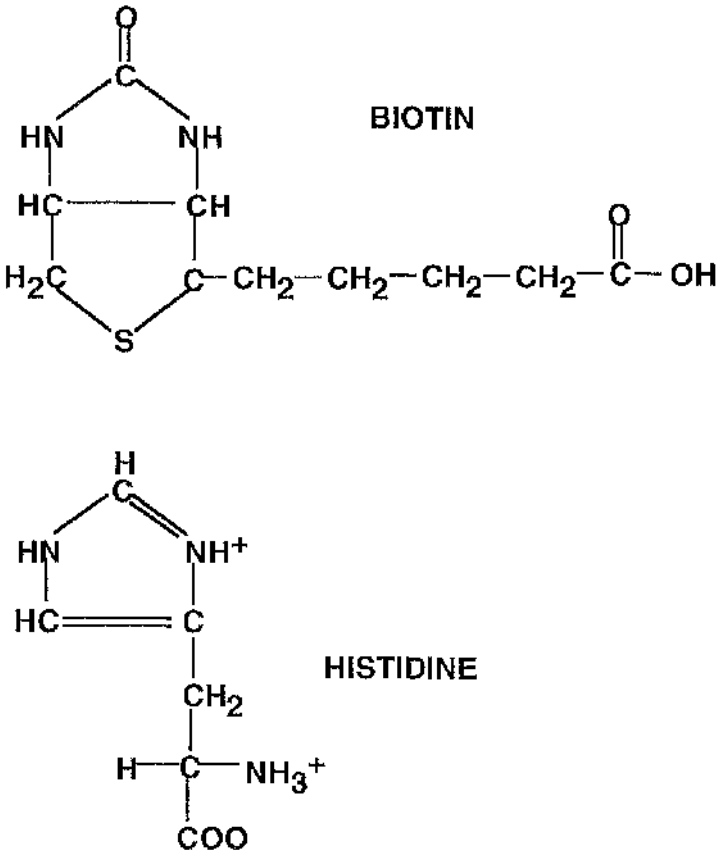


FIGURE 30 : Comparison of the structures of biotin and histidine.

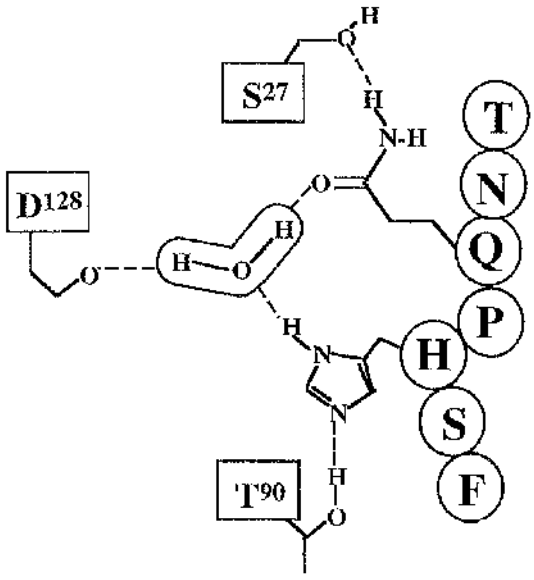


FIGURE 31 : Binding of the heptapeptide FSHPQNT to streptavidin.

After Weber *et al.*, 1992. Streptavidin residues are shown boxed and the heptapeptide residues are circled. A captured water molecule is outlined (centre).

these do also play a part in binding to streptavidin. Weber *et al.*, in 1992, had suggested that a more stringent assay could reveal an extended consensus, but the subsequent investigations of Kay *et al.* and McLafferty *et al.* failed to reveal one. The binding assays emphatically indicated that the phage selected from the Smith pentadecamer library were streptavidin binders, even if they do lack consecutive **H, P** and **Q** residues.

In considering the absence of this established consensus from my results, the first point to bear in mind is the fact that all of the work that determined **HPQ** was carried out using gene III libraries, as opposed to the gene VIII library employed in this phase of the project. The difference that this may make, in terms of monovalent versus polyvalent binding, and terminal versus central location, was discussed, starting at page 68. Another possible, if less likely explanation for the lack of the **HPQ** triplet is that the library simply did not contain inserts with the appropriate codons. There must remain a certain degree of doubt as to how many variants there were in the library as used; although initially produced from 2×10^9 different clones, it must be assumed that a sample from this original library was amplified for distribution. As Russell Thompson did not receive his aliquot directly from George Smith, it is likely that a further amplification had been carried out before he received it. Allowing that the portion removed for amplification may not contain all the possible variants, there is clearly scope for loss of particular sequences. All possible amino acid triplets can be coded by 8×10^3 base sequences, (*ie.* 20^3), and it would require a library of 2.6×10^5 unique nine-base sequences (*ie.* 64^3, for 64 codons) to ensure that all triplets were present. However, due to the sliding reading frame, each 15 amino acid insert would include 13 different triplets, thus, only 2×10^4 different pentadecamer inserts would be sufficient to code for the entire tripeptide range and, as there are 16 different codons for **HPQ**, then it would seem unlikely that the library was completely depleted of these sequences encoding this peptide.

Given the similarities and patterns revealed when viewing the sequences in conjunction with the binding assay results, it may be the case that the nature of the first three amino acids at the amino-terminus of the sequence (which did, in every example, include **H**) are more significant for streptavidin binding than **HPQ** when considering fusions to pVIII.

Given that streptavidin-binding phage were apparently selected in the foregoing experiments, another strategy was sought. An alternative means of immobilising LT-B, using Protein G-bound anti-cholera toxin antibody (see page 51), was employed whilst an additional stage, namely 'backchecking', was added to the panning procedure. After each round of panning, eluted phage were amplified as usual and then exposed in a microtitre plate to the immobilisation agent. Only phage that did not bind at this stage were used in the next panning round. Final-round phage were similarly backchecked, with those that did not bind to the plate being retained as 'B-type phage'. Those that did bind were acid-eluted and were retained as 'E-type phage'. Representatives of these final round phage were then tested in plate assays against, firstly, LT-B, anti-cholera toxin antibody, blocker and naked plastic and then against a range of protein and detergent blockers.

The first assay showed that the 20 phage tested could be assigned to 4 groups based on their binding to LT-B, anti-cholera toxin antibody and blocker (all phage bound with similar efficiency to plastic).

Group 1 contained eight phage, six of which were B-type and two of which were E-type. These all showed remarkably similar binding to all four surfaces, giving a high reading against LT-B and blocker and a moderate reading against anti-cholera toxin antibody.

Group 2 consisted of two B-type phage that bound poorly to all three surfaces.

Group 3 contained phage that demonstrated good binding to all three surfaces. Eight members of this group were E-type and one was B-type.

Finally, Group 4 was a sole B-type phage that showed moderate binding to all three surfaces.

These groupings were illustrated graphically in Figure 26 on page 140.

The lack of round-on-round amplification during panning strongly suggested that no LT-B-specific phage had been isolated, and the assay results appeared to confirm this; no phage bound strongly to LT-B alone. However, the strength of binding of phage to a variety of coated surfaces suggested that some specific interaction was occurring. The same twenty phage were added to wells coated with the proteins BSA, ovalbumin, gelatin and skimmed milk, and to wells coated with, and washed during the assay procedure with, the detergents Tween 20, Triton X-100 and Nonidet P40.

The results from these tests generally confirmed those obtained in the first ELISAs, with three groups reacting strictly to type. Group 1 phage gave strong signals in wells coated with ovalbumin or previously washed with Tween 20 and Triton X-100 and a moderate reading with Nonidet p40 and a low reading with skimmed milk. Group 3 phage gave high readings with ovalbumin, Tween 20 and Nonidet p40 and moderate readings with Triton X-100 and skimmed milk. The Group 4 phage showed moderate numbers binding to wells coated with ovalbumin, Tween 20, Triton X-100 and Nonidet p40 and a low reading with skimmed milk.

Only the Group 2 phage failed to bind in a similar manner. Both phage bound to a moderate degree to the wells blocked with ovalbumin and to a low extent to those which had been blocked with skimmed milk. Phage B-1 bound moderately to the three wells that had been washed in each detergent, however, phage B-7 bound poorly to these wells.

All phage bound poorly to the gelatin-blocked well.

Despite the lack of an increase in the number of phage eluted in each succeeding round the binding assay results clearly indicate that some selection did occur in panning, with four groups of phage identified that behaved in a similar manner. There is also evidence to suggest that backchecking did remove a specific sub-set of phage from the amplified pool, as eight of the nine phage that constituted Group 3 were eluted from the backchecking well.

However, once again the one definite conclusion that can be drawn is that this third panning experiment did not promote the amplification of LT-B-binding phage.

OVERVIEW AND CONCLUSIONS

Background to the project

Fourteen years ago the first steps were taken in the development of phage display (Smith, 1985), eleven years ago the first paper outlining the basic methods was published (Parmley and Smith, 1988) and nine years ago the production and utilisation of the first phage display libraries was reported (Cwirla *et al.*, 1990; Devlin *et al.*, 1990; Scott and Smith, 1990). George P. Smith's original vision (Smith, 1996) was of an immunological tool, and displayed peptides have indeed been used to discover epitope-mimics ('mimotopes'), whilst the insertion of antibody genes into phage has resulted in the display of active antibody fragments, leading to the isolation of human antibody fragments with specificity against protein and non-protein antigens (Winter *et al.*, 1994).

Phage display has proved equally applicable in other fields. Devlin *et al.* (1990) described fusions that bound to streptavidin, a protein with no previously reported affinity for peptides, and Oldenburg *et al.* (1992) and Scott *et al.* (1992) simultaneously reported that displayed random peptides could bind to the lectin, concanavalin A, for which the natural ligand was a carbohydrate. Hoess *et al.* (1993), in discovering peptides that bound to the carbohydrate receptor site of a monoclonal antibody, appeared to confirm the existence of peptides able to mimic carbohydrate determinants. O'Neil *et al.* (1992) showed that phage-displayed peptides could mimic fibronectin in binding to the cell adhesion glycoprotein integrin, IIb/IIIa, thus demonstrating that such peptides could specifically bind to natural receptor molecules.

It was against this background that the idea of applying phage display to the study of toxin-receptor interactions was first conceived. The toxin chosen for analysis was LT-B, which had recently been crystallised, yielding a wealth of information on the interaction with its natural receptor (Sixma *et al.*, 1991; 1992; 1993; Merritt *et al.*, 1994a; 1994b). That the gene for the binding subunit had recently been cloned into an expression system (Amin and Hirst, 1994), increased its appeal as a model system.

Overview of experimental procedure

Acid elution was employed when biopanning all the libraries tested. The resilience of bacteriophage to extreme pH makes this most stringent form of elution ideal and it has been the agent of choice in most display work since Parmley and Smith (1988) first recommended it. O'Neil *et al.* (1992) did suggest that such non-specific elution used in the first panning round could promote domination by irrelevant phage. Indeed, they found no identifiable sequence motifs when using it at this stage, and recommended that a specific ligand should be used instead, to provide an initial pool of binders that could be refined by further rounds of panning. In accordance with this, G_{M1} elution of phage from LT-B^{biotin} was attempted with the Smith hexamer library, however no clones that demonstrated LT-B- specific binding were isolated.

Recovery with TBS proved useful as a control, giving a measure of non-specific binding and elution, resulting from the use of such vast numbers of phage in the panning process. After ten washes, detectable numbers of phage remained, possibly associated with the plastic tube, or with the streptavidin or LT-B^{biotin}. These phage were not amplified to any extent from round to round, and would have been of little significance if a population of specifically eluted, amplifiable phage had been present.

Overview of results

Binding of phage-borne peptides was investigated using gene III and gene VIII libraries, representing display types 3, 8+8 and 88, with inserts of varied lengths (hexapeptide, nonapeptide and pentadecapeptide).

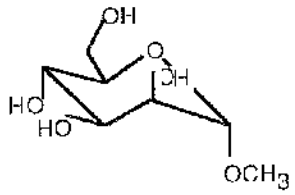
The results with the Smith hexamer library were disappointing. Despite the fact that nucleic acid sequences that could be read over the interfering wild-type signal carried related inserts, eluted phage showed little evidence of amplification over four rounds and performed poorly in the binding assays. This suggests that they were associating with something other than LT-B^{biotin}. Comparison with known streptavidin or plastic binding sequences failed to indicate relatedness.

Flaws in construction of the Luzzago nonamer library, revealed by sequencing meant that no useful data was generated by this part of the study. This was particularly unfortunate as this library had been employed to select carbohydrate epitope-mimicking peptides (Phalipon *et al.*, 1997). When used to immunise mice, anti-carbohydrate antibodies were raised against these selected peptides.

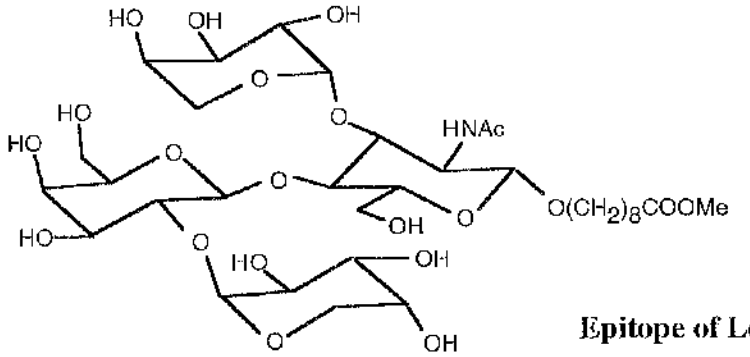
The Smith pentadecamer library panned and sequenced well, however, phage isolated at the final round failed to bind to LT-B^{biotin}. Some clones did, however, demonstrate specificity for streptavidin, and the sequences of the binding phage were internally consistent and bore elements of similarity, in respect of the residues **H**, **P** and **Q**, to those previously published (Devlin *et al.*, 1990; Kay *et al.*, 1993; McLafferty *et al.*, 1993).

Therefore, using the heat-labile toxin of *E. coli* as a model system, the phage display technique failed to generate peptide mimics of the natural carbohydrate ligand.

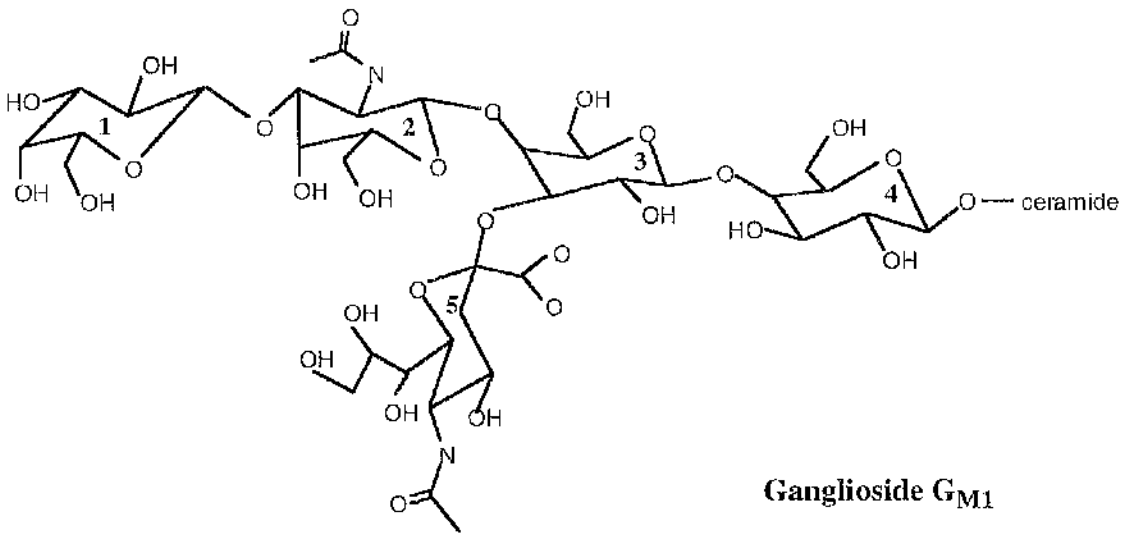
In considering why this should be, it is valuable to compare G_{M1} with α-D-methylmannoside (the saccharide ligand of concanavalin A) and the Lewis^y antigen. These structures are shown in Figure 31. Although G_{M1} is by far the most complex, the galactose (1) component (refer to Figure 31 for numbering scheme) has a superficially comparable structure to the α-D-methylmannoside and the non-acetylated rings of the Lewis^y antigen. The other critical structures for G_{M1} binding, principally the sialic acid (5), and to a lesser extent the N-acetyl glucosamine (2) are much more complex, and it may be that peptides are simply unable to mimic these residues. Galactose alone is sufficient to bind LT-B (a property exploited in the purification process), however, the LT-B^{biotin} used in the panning work possibly



α-D-methylmannoside



Epitope of Lewis^y antigen



Ganglioside G_{M1}

FIGURE 32 : Comparison of the structures of α-D-methylmannoside, the epitope of the Lewis^y antigen and G_{M1}.

Rings comprising the G_{M1} are galactose (1); N-acetyl-glucosamine (2); galactose (3); glucose (4); sialic acid (5).

carried bound galactose, and, while it was hoped that this would have helped select for peptides with higher binding affinities, it was equally a possibility that it served to prevent any of the peptides from binding. For that reason, 'Hirst LT-B' produced by the alternative purification strategy of Amin and Hirst (1994) was used in the second and third biopanning experiments carried out with the Smith pentadecamer library. The results obtained from these two pannings and the subsequent binding assays did not suggest that the stock LT-B^{biotin} used in the earlier experiments performed any differently to that produced by Professor Hirst.

A second possibility that always exists in phage display work, is that the binding peptide was simply absent from the libraries tested. The Smith hexamer library, as originally produced, was calculated to contain 1.3×10^{14} unique clones, when only 1.1×10^{9} (*ie.* 32^{6}) were required to guarantee that all possible hexamers were present. However, it must be remembered that what was used here was not the original library, but an amplification of an amplified part of that. Initially, the library had 2×10^{14} clones, representing around one million copies of each unique insert. It would thus seem unlikely that substantial loss of diversity would have occured through repeated amplification However, this view depends upon each phage possessing an equal ability to amplify and infect; this is not necessarily the case. Burritt *et al.* suggest that dominance by certain phage may be due to possession of some characteristic, unrelated to binding specificity, that promotes proliferation and that, if amplification is carried out on solid medium, this imbalance can be reduced. In their review of phage display, Burritt *et al.* list thirty proven consensus sequences of which nineteen are less than six amino acids in length. The longest consists of twenty-two residues, although this includes many variable positions interspersed among the set amino acids. The longest sequences that are defined throughout are seven amino acids in length. The streptavidin consensus is **HPQ**, and, were the sequence sought here also this length, then the number of variants in the original library would have been perfectly adequate. Set against this is the observation of Weber *et al.* (1992), that

amongst streptavidin-binding phage, the residues adjacent to **HPQ** are also involved in binding, suggesting that the consensus sequence for binding may well be much longer. Obviously, the pentadecamer library is far from complete - this would require 3.28×10^{19} independent clones - but, as discussed earlier, it will contain, within its inserts, many shorter peptides. Therefore, with a small consensus sequence there is a greater chance of its being present within the pentadecapeptide insert than within the hexapeptide.

The third possibility is that only conformationally-constrained peptides will bind to LT-B. The failure with the Luzzago library prevented investigation of this issue. Ladner (1995) suggests that discoveries of peptide mimics of protein-binding sequences are rare when unconstrained libraries are used. Set against this argument is the fact that the previous carbohydrate-mimicking peptides were discovered using unconstrained libraries - in fact, Hoess *et al.* (1993) actually tested constrained and unconstrained libraries and discontinued work with the former due to lack of success.

A fourth possibility is that peptide mimicking of carbohydrate ligands is not actually feasible. It was certainly curious, against a background of mass acceptance of phage display (there were 246 papers in the Science Citation Index of the BIDS database for the year 1998) and its application in ever more areas, that in the first four years after the Scott, Oldenburg and Hoess papers first appeared, only three other such examples were reported. Two of these involved peptide mimicking of carbohydrate epitopes, along the lines of the Hoess *et al.* / Lewis^y antigen work, and the other was by Martens *et al.* (1995), in which E-selectin, a cell adhesion molecule, was panned with the Cwirla *et al.* hexapeptide library. The isolated peptides did not compete for binding with the natural ligand and were considered to be binding outwith the site recognised by the carbohydrate ligand, possibly to a protein.

A paper by Bonnycastle *et al.* (1996), from the group that reported peptide binding to concanavalin A, presented results from testing eleven pVIII libraries against a bank of antibodies. Two of them were specific for saccharide epitopes and in both these cases binding peptides (*ie.* carbohydrate mimics) were eluted. However, the authors also

reported that their libraries failed to bind to any of a panel of ten animal and plant lectins. Further, they conceded that the peptide that had apparently mimicked the binding of α–D-methylmannoside, the concanavalin A ligand (Oldenburg *et al.*, 1992; Scott *et al.*, 1992) had not, in fact, reacted with the sugar-binding site. Drawing together these results with those of Martens *et al.*, they conclude that although phage display peptides that undoubtedly react well with anti-carbohydrate antibodies (eg. the previously-discussed papers by Hoess *et al.*, 1993, and Phalipon *et al.*, 1997) they may well be unable to bind to lectins in a manner analogous to their native ligands. In a further paper (Harris *et al.*, 1997) the same group reported further success when peptides were reacted with anti-carbohydrate antibodies but repeated their earlier view that true carbohydrate mimicking by peptides was a reasonably unlikely prospect. The reason given for this was that antigenic mimickry by peptides is dependent on factors other than the peptide demonstrating chemical or structural similarity to the native antigen, such as the nature of the binding site of the anti-carbohydrate antibody.

This same group have since begun investigations into the chemical and structural bases of antigen mimickry by phage displayed peptides (Craig *et al.*, 1998), however, at this stage they have yet to publish data relating to anti-carbohydrate antibodies.

Final conclusions

It is unfortunate that the Bonnycastle *et al.* paper was published so late in the course of the practical element of this work as, in the light of their findings, LT-I, appears to be an inappropriate model system for investigating phage display. However, based on the information available in 1994, when the project was first commenced, it was a legitimate choice.

Remembering that the whole intention was to investigate the potential use of phage display as a tool in the study of toxin-receptor interactions, then the firm conclusion that would have been drawn from these results, regardless of the Bonnycastle *et al.* paper, is that it had no role to play as far as LT-I was concerned. Now, in

consideration of the Bonnycastle *et al.* results and my own, that conclusion can be extended to cover not just LT-I, but any other toxin that specifically, or preferentially recognises carbohydrate receptors. These other toxins include LT-II and cholera toxin, *C. perfringens* δ– toxin, shiga and shiga-like (vero) toxin, botulinum toxins, tetanus toxin and toxin A of *Clostridium difficile*.

The potential use of phage display to discover previously unidentified toxin receptors is also limited. Whilst a positive result could be obtained if the receptor was a protein, a negative result could suggest that the receptor was possibly, but not definitely, a carbohydrate.

Of course, where phage display may still have a role is in conjunction with those toxins, such as diphtheria toxin, *E. coli* ST-I, and pertussis toxin that are known to have protein receptors. It would be of great interest to see if a peptide could be selected from a display library that would attach within the receptor binding sites of these toxins, either by imitating the original receptor or by establishing a novel binding strategy of its own. Such a peptide might ultimately prove to have therapeutic potential, thereby extending yet further the range of uses for George Smith's visionary idea.

APPENDICES

APPENDIX 1 : MEDIA RECIPES

Luria Bertani (LB) broth

To 950ml of distilled water add :

Bactotryptone	10 g
Bacto Yeast Extract	5 g
Sodium Chloride	10 g

Shake to dissolve. Adjust pH to 7.0 with 5N NaOH (~0.2ml). Adjust to 1 litre with distilled water. Autoclave 20 minutes at 121 °C

Luria Bertani (LB) agar

As for Luria Bertani broth with technical agar added to 1.2%.

Luria Bertani High Salt (LBHS) medium (for marine organisms)

To 950ml of distilled water add :

Bactotryptone	10 g
Bacto Yeast Extract	5 g
Sodium Chloride	15 g

Shake to dissolve. Adjust pH to 7.0 with 5N NaOH (~0.2ml). Adjust to 1 litre with distilled water. Autoclave 20 minutes at 121 °C

terrific broth

In 900ml of water dissolve :

Tryptone	12 g
Yeast Extract	24 g
Glycerol	4ml

Autoclave as 90ml aliquots. Before use, add 10ml of separately autoclaved phosphate buffer (2.31 g of KH_2PO_4 and 12.54 g of K_2HPO_4 dissolved and made up to 100ml in distilled water).

APPENDIX 2 : SDS PAGE SOLUTIONS

Sample preparation

Solubilising buffer

Glycerol	5.0ml
20% SDS	2.5ml
β -- mercaptoethanol	0.5ml
0.5M Tris HCl, pH 6.8	2.0ml
Bromophenol blue (1mg ml^{-1})	0.5ml

Gel preparation

Separating gel	8%	13%	14%
30% acrylamide, 0.8% bis	4.0 ml	6.5 ml	7.0ml
distilled water	7.0 ml	4.5 ml	4.0ml
1.5M Tris HCl, pH 8.8	3.75ml	3.75ml	3.75ml
20% (w/v) SDS	75.0 µl	75.0 µl	75.0 µl
TEMED	7.5 µl	7.5 µl	7.5 µl
10% (w/v) APS	150.0 µl	150.0 µl	150.0 µl

Stacking gel	6%
30% acrylamide, 0.8% bis	2.0 ml
distilled water	5.4 ml
0.5M Tris HCl, pH 6.8	2.5 ml
20% (w/v) SDS	50.0 µl
TEMED	5.0 µl
10% (w/v) APS	100.0 µl

Gel running

Electrode buffer	(1 litre)
Tris	3.02g
Glycine	14.0 g
SDS	1.0 g

APPENDIX 3 : GEL STAINING

Coomassie blue

Concentrated Stain

Coomassie blue R250		1.25 g
Methanol		200 ml
Glacial acetic acid		50 ml
Distilled water	to	500 ml

Gel was soaked in concentrated stain for 20 minutes at room temperature, then transferred into destain. Destaining continued at room temperature, with frequent changes, until background was cleared.

Destain

Distilled water containing 40% methanol and 10% acetic acid. Gels were retained until photographed in distilled water.

APPENDIX 4 : MISCELLANEOUS RECIPES

Western blotting transfer buffer

Tris	5.82 g
Glycine	2.83 g
SDS	0.0375 g
Methanol	200 ml

Distilled water to 1 litre

PEG /NaCl

To 475ml water add :

PEG 8000	100 g
NaCl	116.9 g

Phosphate buffered saline (PBS)

Dulbecco A tablets from Oxoid, Basingstoke were used, 1 tablet per 100ml of distilled water.

PBS-BSA

PBS containing 0.1% bovine serum albumin (Sigma; Poole).

PBS-TWEEN

PBS containing 0.5% Tween 20 (Polyoxyethylenesorbitan monolaurate; Sigma; Poole)

Running buffer for sequencing

TBE (0.09M tris borate, 0.002M EDTA, pH 8.0)

Tris	135.0 g
Boric acid	68.75 g
EDTA	4.65 g

Makes 2.5 litres of 5 X. Dilute to single strength for use.

TE buffer

1mM EDTA in 10mM Tris/HCl, pH 8.0

Tris buffered saline (TBS)

Dissolve one physiological saline tablet (Oxoid; Basingstoke) in 475ml distilled water. Add 25ml of 1M Tris/HCl, pH 7.5.

TBS-TWEEN

TBS containing 0.5% Tween 20 (Polyoxyethylenesorbitan monolaurate)

TBS-G

TBS containing 0.2% gelatin.

APPENDIX 5 : AMINO ACIDS - SINGLE LETTER CODING

Showing frequency of codon appearance in the full and NNK genetic codes

		No. OF CODONS	
		FULL CODE	NNK
A	Alanine	4	2
C	Cysteine	2	1
D	Aspartic acid	2	1
E	Glutamic acid	2	1
F	Phenylalanine	2	1
G	Glycine	4	2
H	Histidine	2	1
I	Isoleucine	3	1
K	Lysine	2	1
L	Leucine	6	3
M	Methionine	1	1
N	Asparagine	2	1
P	Proline	4	2
Q	Glutamine	2	1
R	Arginine	6	3
S	Serine	6	3
T	Threonine	4	2
V	Valine	4	2
W	Tryptophan	1	1
Y	Tyrosine	2	1
	Stop	3	1

APPENDIX 6 : GENETIC CODE

Amino acids designated by single letters (see Appendix 5). **BOLD** codons constitute the NNK genetic code. (St. = stop codon)

T T T	**F**	**T C T**	**S**	**T A T**	**Y**	**T G T**	**C**
T T C	F	T C C	S	T A C	Y	T G C	C
T T A	L	T C A	S	T A A	st.	T G A	st.
T T G	**L**	**T C G**	**S**	**T A G**	**st.**	**T G G**	**W**
C T T	**L**	**C C T**	**P**	**C A T**	**H**	**C G T**	**R**
C T C	L	C C C	P	C A C	H	C G C	R
C T A	L	C C A	P	C A A	Q	C G A	R
C T G	**L**	**C C G**	**P**	**C A G**	**Q**	**C G G**	**R**
A T T	**I**	**A C T**	**T**	**A A T**	**N**	**A G T**	**S**
A T C	I	A C C	T	A A C	N	A G C	S
A T A	I	A C A	T	A A A	K	A G A	R
A T G	**M**	**A C G**	**T**	**A A G**	**K**	**A G G**	**R**
G T T	**V**	**G C T**	**A**	**G A T**	**D**	**G G T**	**G**
G T C	V	G C C	A	G A C	D	G G C	G
G T A	V	G C A	A	G A A	E	G G A	G
G T G	**V**	**G C G**	**A**	**G A G**	**E**	**G G G**	**G**

APPENDIX 7 : GENE III OF fd-tet

Base and amino acid sequences of protein III of phage fd-tet. The capitalised, underlined bases represent the annealing site of the primer OLIGO 2303.

SIGNAL PEPTIDE

```
atg aaa aaa tta tta ttc gca att cct tta gtt gtt cct ttc tat tct cac
 M   K   K   L   L   F   A   I   P   L   V   V   P   F   Y   S   H

tcc
 S
```

MATURE PROTEIN

```
gct gaa act gtt gaa agt tgt tta gca aaa cct cat aca gaa aat tca ttt
 A   E   T   V   E   S   C   L   A   K   P   H   T   E   N   S   F

act aac gtc tgg aaa gac gac aaa act tta gat CGT TAC GCT AAC TAT GAG
 T   N   V   W   K   D   D   K   T   L   D   R   Y   A   N   Y   E

GGc tgt cgt tgg aat gct aca ggc gtt gtg gtt tgt act ggt gac gaa act
 G   C   L   W   N   A   T   G   V   V   V   C   T   G   D   E   T

cag tgt tac ggt aca tgg gtt cct att ggg ctt gct atc cct gaa aat gag
 Q   C   Y   G   T   W   V   P   I   G   L   A   I   P   E   N   E

ggt ggt ggc tct gag ggt ggc ggt tct gag ggt ggc ggt tct gag ggt ggc
 G   G   G   S   E   G   G   G   S   E   G   G   G   S   E   G   G

ggt act aaa cct cct gag tac ggt gat aca cct att ccg ggc tat act tat
 G   T   K   P   P   E   Y   G   D   T   P   I   P   G   Y   T   Y

atc aac cct ctc gac ggc act tat ccg cct ggt act gag caa aac ccc gct
 I   N   P   L   D   G   T   Y   P   P   G   T   E   Q   N   P   A

aat cct aat cct tct ctt gag gag tct cag cct ctt aat act ttc atg ttt
 N   P   N   P   S   L   E   E   S   Q   P   L   N   T   F   M   F

cag aat aat agg ttc cga aat agg cag ggt gca tta act gtt tat acg ggc
 Q   N   N   R   F   R   N   R   Q   G   A   L   T   V   Y   T   G

act gtt act caa ggc act gac ccc gtt aaa act tat tac cag tac act cct
 T   V   T   Q   G   T   D   P   V   K   T   Y   Y   Q   Y   T   P

gta tca tca aaa gcc atg tat gac gct tac tgg aac ggt aaa ttc aga gac
 V   S   S   K   A   X   Y   D   A   Y   W   N   G   K   F   R   D
```

tgc gct ttc cat tct ggc ttt aat gag gat cca ttc gtt tgt gaa tat caa

C A F H S G F N E D P F V C E Y Q

ggc caa tcg tct gac ctg cct caa cct cct gtc aat gct ggc ggc ggc tct

G Q S S D L P Q P P V N A G G G S

ggt ggt ggt tct ggt ggc ggc tct gag ggt ggc ggc tct gag ggt ggc ggt

G G G S G G G S E G G G S E G G G

tct gag ggt ggc ggc tct gag ggt ggc ggt tcc ggt ggc ggc tcc ggt tcc

S E G G G S E G G G S G G G S G S

ggt gat ttt gat tat gaa aaa atg gca aac gct aat aag ggg gct atg acc

G D F D Y E K M A N A N K G A M T

gaa aat gcc gat gaa aac gcg cta cag tct gac gct aaa ggc aaa ctt gat

E N A D E N A L Q S D A K G K L D

tct gtc gct act gat tac ggt gct gct atc gat ggt ttc att ggt gac gtt

S V A T D Y G A A I D G F I G D V

tcc ggc ctt gct aat ggt aat ggt gct act ggt gat ttt gct ggc tct aat

S G L A N G N G A T G D F A G S N

tcc caa atg gct caa gtc ggt gac ggt gat aat tca cct tta atg aat aat

S Q M A Q V G D G D N S P L M N N

ttc cgt caa tat tta cct tct ttg cct cag tcg gtt gaa tgt cgc cct tat

F R Q Y L P S L P Q S V E C R P Y

gtc ttt ggc gct ggt aaa cca tat gaa ttt tct att gat tgt gac aaa ata

V F G A G K P Y E F S I D C D K I

aac tta ttc cgt ggt gtc ttt gcg ttt ctt tta tat gtt gcc acc ttt atg

N L F R G V F A F L L Y V A T F M

tat gta ttt tcg acg ttt gct aac ata ctg cgt aat aag gag tct taa

Y V F S T F A N I L R N K E S Stop

APPENDIX 8 : GENE VIII OF fd-tet

Base and amino acid sequences of protein VIII of phage fd-tet.

SIGNAL PEPTIDE

```
atg aaa aag tct tta gtc ctc aaa gcc tcc gta gcc gtt gct acc ctc gtt
 M   K   K   S   L   V   L   K   A   S   V   A   V   A   T   L   V

ccg atg ctg tct ttc gct
 P   M   L   S   F   A
```

MATURE PROTEIN

```
gct gag ggt gac gat ccc gca aaa gcg gcc ttt gac tcc ctg caa gcc tca
 A   E   G   D   D   P   A   K   A   A   F   D   S   L   Q   A   S

gcg acc gaa tat atc ggt tat gcg tgg gcg atg gtt gtt gtc att gtc ggc
 A   T   E   Y   I   G   Y   A   W   A   M   V   V   V   I   V   G

gca act atc ggt atc aag ctg ttt aag aaa ttc acc tcg aaa gca agc tga
 A   T   I   G   I   K   L   F   K   K   F   T   S   K   A   S  Stop
```

APPENDIX 9 : AMINO ACID SEQUENCE OF LT-B

Lysine residues (highlighted in bold type) are at positions 23, 34, 43, 62, 63, 69, 81, 84 and 91.

```
1     APQSITELCS EYHNTQIYTI NDKILSYTES MAGKREMVII TFKSGATFQV

51    EVPGSQHIDS QKKAIERMKD TLRITYLTET KIDKLCVWNN KTPNSIAAIS MEN
```

REFERENCES

Acheson, D.W.K. 'Enterotoxins in acute infective diarrhoea.' *Journal of Infection* **24 :** 225 - 245 (1992)

Adey, N.B., Mataragnon, A.H., Rider, J.E., Carter, J.M. and Kay, B.K. 'Characterisation of phage that bind plastic from phage-displayed random peptide libraries.' *Gene* **156 :** 27 - 31 (1995)

Alouf, J.E. 'A decade of progress in toxin research.' In *Bacterial Protein Toxins*, eds. J. Freer, R. Aitken, J.E. Alouf, G. Boulnois, P. Falmagne, F. Fehrenbach, C. Montecucco, Y. Piemont, R. Rappuoli, T. Wadstrom, and B. Witholt. 3 - 29. Gustav Fischer, Stuttgart (1994)

Amin, T. and Hirst, T.R. 'Purification of the B-subunit oligomer of *Escherichia coli* heat-labile enterotoxin by heterologous expression and secretion in a marine vibrio.' *Protein Expression and Purification* **5 :** 198 - 204 (1994)

Anonymous 'Dr. Koch`s sixth cholera report.' *British Medical Journal* March 22 : 568 - 569 (1884)

Anonymous 'Consensus development conference statement.' *Reviews of Infectious Diseases* **8 :** S227 - S233 (1986)

Anonymous *Weekly Epidemiological Record* **69 :** 322 - 323 (1994a)

Anonymous 'Cholera update, end of 1993.' *Weekly Epidemiological Record* **69 :** 13 - 17 (1994b)

Anonymous 'Diarrhoeal diseases in Bali.' *Weekly Epidemiological Record* **70 :** 86 - 87 (1995a)

Anonymous 'WHO global programme for vaccines and immunisation.' *Weekly Epidemiological Record* **70 :** 1 (1995b)

Bardwell, J.C.A., McGovern, K. and Beckwith, J. 'Identification of a protein required for disulfide bond formation *in vivo*..' *Cell* **67 :** 581 - 589 (1991)

Barrett, R.W., Cwirla, S.E., Ackerman, M.S., Olsen, A.M., Peters, E.A. and Dower, W.J. 'Selective enrichment and characterisation of high affinity ligands from collections of random peptides on filamentous phage.' *Analytical Biochemistry* **204 :** 357 - 364 (1992)

Bass, S., Greene, R. and Wells, J.A. 'Hormone phage : an enrichment method for variant proteins with altered binding properties.' *Proteins* **8 :** 309 - 314 (1990)

Becker, J.W., Cunningham, B.A., Reeke, G.N., Wang, J.L. and Edelman, G.M. 'The molecular structure of Concanavalin A.' In *Concanavalin A as a tool*, eds. H. Bittiger and H.P. Schnebli. 33 - 54. Wiley, New York (1976)

Beekwilder, J., Rakonjac, J., Jongsma, J. and Bosch, D. 'A phagemid vector using the *E. coli* phage shock promoter facilitates phage display of toxic proteins' *Gene* **228** : 23 - 31 (1999)

Björck, L. and Kronvall, G. 'Purification and some properties of Streptococcal Protein G, a novel IgG-binding reagent' *Journal of Immunology* **133** : 969 - 974 (1984)

Black, R.E. 'The epidemiology of cholera and enterotoxigenic *E. coli* diarrheal disease.' In *Development of Vaccines and Drugs against Diarrhea. 11th Nobel*

conference, Stockholm 1985, eds. J. Holmgren, A. Lindberg, and R. Mollby. 23 - 32. Studentlitteratur, Lund, Sweden (1986)

Black, R.E. 'Epidemiology of traveler's diarrhea and relative importance of various pathogens.' *Reviews of Infectious Diseases* **12 :** S73 - S79 (1990)

Black, R.E. 'Epidemiology of diarrhoeal disease: implications for control by vaccines.' *Vaccine* **11 :** 100 - 106 (1993)

Boeke, J.D. and Model, P. 'A prokaryotic membrane anchor sequence: Carboxyl terminus of bacteriophage f1 gene III protein retains it in the membrane.' *Proceedings of the National Academy of Sciences of the United States of America* **79 :** 5200 - 5204 (1982)

Bonnycastle, L.L.C., Mehroke, J.S., Rashed, M., Gong, X. and Scott, J.K. 'Probing the basis of antibody reactivity with a panel of constrained peptide libraries displayed by filamentous phage.' *Journal of Molecular Biology* **258 :** 747 - 762 (1996)

Booth, B.A., Boesman-Finkelstein, M. and Finkelstein, R.A. '*Vibrio cholerae* hemagglutinin / protease nicks cholera enterotoxin.' *Infection and Immunity* **45 :** 558 - 560 (1984)

Bradford, M.M. 'A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.' *Analytical Biochemistry* **72 :** 248 - 254 (1976)

Brock, T.D. '*Robert Koch. A life in medicine and bacteriology.*' Scientific Revolutionaries. A biographical series, Science Tech Publishers, Madison (1988)

Burritt, J.B., Bond, C.W., Doss, K.W. and Jesaitis, A.J. 'Filamentous phage display of oligopeptide libraries.' *Analytical Biochemistry* **238 :** 1 - 13 (1996)

Cesareni, G. 'Peptide display on filamentous phage capsids. A new powerful tool to study protein-ligand interaction.' *FEBS Letters* **307 :** 66 - 70 (1992)

Clackson, T. and Wells, J.A. '*In vitro* selection from protein and peptide libraries.' *Trends in Biotechnology* **12 :** 173 - 184 (1994)

Clements, J.D. and Finkelstein, R.A. 'Demonstration of shared and unique immunological determinants in enterotoxins from *Vibrio cholerae* and *Escherichia coli*.' *Infection and Immunity* **22 :** 709 - 713 (1978a)

Clements, J.D. and Finkelstein, R.A. 'Immunological cross-reactivity between a heat-labile enterotoxin(s) of *Escherichia coli* and subunits of *Vibrio cholerae* enterotoxin.' *Infection and Immunity* **21 :** 1036 - 1039 (1978b)

Clements, J.D. and Finkelstein, R.A. 'Isolation and characterization of homogeneous heat labile enterotoxins with high specific activity from *Escherichia coli* cultures.' *Infection and Immunity* **24 :** 760 - 769 (1979)

Corey, D.R., Shiau, A.K., Yang, Q., Janowski, B.A. and Craik, C.S. 'Trypsin display on the surface of bacteriophage.' *Gene* **128 :** 129 - 134 (1993)

Craig, L., Sanschagrin, P.C., Rozek, A., Lackie, S., Kuhn, L.A. and Scott, J.K. 'The role of structure in antibody cross-reactivity between peptides and folded proteins.' *Journal of Molecular Biology* **281** : 183 - 201 (1998)

Cwirla, S.E., Peters, E.A., Barret, R.W. and Dower, W. J. 'Peptides of phage : A vast library of peptides for identifying ligands.' *Proceedings of the National Academy of Sciences of the United States of America* **87 :** 6378 - 6382 (1990)

Dallas, W.S. 'Conformity between heat-labile toxin genes from human and porcine enterotoxigenic *Escherichia coli.' Infection and Immunity* **40 :** 647 - 652 (1983)

Dallas, W.S. and Falkow, S. 'Amino acid sequence homology between cholera toxin and *Escherichia coli* heat-labile toxin.' *Nature* **288 :** 499 - 501 (1980)

de la Cruz, V., Lal, L. and McCutchan, T. 'Immunogenicity and epitope mapping of foreign sequences *via* genetically engineered filamentous phage.' *Journal of Biological Chemistry* **263 :** 4318 - 4322 (1988)

De, S.N. 'Enterotoxicity of bacteria-free culture-filtrate of *Vibrio cholerae.' Nature* **183 :** 1533 - 1534 (1959)

De, S.N., Bhattacharya, K. and Sarkar, J.K. 'A study of the pathogenicity of strains of *Bacterium coli* from acute and chronic enteritis.' *Journal of Pathology and Bacteriology* **71 :** 201 - 209 (1956)

de Bruin, R., Spelt, K., Mol, J., Koes, R. and Quattrocchio, F. 'Selection of high-affinity phage antibodies from phage display libraries' *Nature Biotechnology* **17** : 397 - 399 (1999)

Dente, L., Sollazzo, M., Baldari, C., Cesareni, G. and Cortese, R. 'The pEMBL family of single-stranded vectors.' In *DNA Cloning : A Practical Approach*, ed. D.M. Glover. 101 - 107. 1. IRL Press, Oxford (1985)

Devlin, J.J., Panganiban, L.C. and Devlin, P.E. 'Random peptide libraries : A source of specific protein binding molecules.' *Science* **249 :** 404 - 406 (1990)

Dorner, F. '*Escherichia coli* enterotoxin.' *Journal of Biological Chemistry* **250 :** 8712 - 8719 (1975)

Dorner, F., Jaksche, H. and Stockl, W. '*Escherichia coli* enterotoxin: Purification, partial characterisation and immunological observations.' *Journal of Infectious Diseases* **133 :** S142 - S156 (1976)

Dotto, G.P., Enea, V. and Zinder, H.D. 'Functional analysis of bacteriophage f1 intergenic region.' *Virology* **114 :** 463 - 473 (1981)

Dotto, G.P. and Horiuchi, K. 'Replication of a plasmid containing two origins of bacteriophage f1.' *Journal of Molecular Biology* **153 :** 169 - 176 (1981)

Dougan, G. 'The molecular basis for the virulence of bacterial pathogens : Implications for oral vaccine development.' *Microbiology* **140 :** 215 - 224 (1994)

Dutta, N.K., Panse, M.V. and Kulkarni, D.R. 'Role of cholera toxin in experimental cholera.' *Journal of Bacteriology* **78 :** 594 - 595 (1959)

Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. and Cesareni, G. 'Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector.' *Journal of Molecular Biology* **222 :** 301 - 310 (1991)

Felici, F., Luzzago, A., Folgori, A. and Cortese, R. 'Mimicking of discontinuous epitopes by phage-displayed peptides, II. Selection of clones recognized by a

protective monoclonal antibody against the *Bordetella pertussis* toxin from phage peptide libraries.' *Gene* **128 :** 21 - 27 (1993)

Field, M., Fromm, D., Wallace, C.K. and Greenough, W.B. 'Stimulation of active chloride secretion in small intestine by cholera toxin.' *Journal of Clinical Investigation* **48 :** 24a (1969)

Finkelstein, R.A. and LoSpalluto, J.J. 'Production of highly purified choleragen and choleragenoid.' *Journal of Infectious Diseases* **121 :** S63 - S72 (1970)

Finkelstein, R.A., LaRue, M.K., Johnston, D.W., Vasil, M.L., Cho, G.J. and Jones, J.R. 'Isolation and properties of heat-labile enterotoxin(s) from enterotoxigenic *Escherichia coli*.' *Journal of Infectious Diseases* **133 :** S120 - S137 (1976)

Fishman, P.H. 'Mechanism of action of cholera toxin : studies on the lag period.' *Journal of Membrane Biology* **54 :** 61 - 72 (1980)

Fishman, P.H. 'Mechanism of action of cholera toxin.' In *ADP-ribosylating toxins and G-proteins. Insights into signal transduction*, eds. J. Moss and M. Vaughan. 127 - 140. ASM, Washington D.C. (1990)

Folgori, A., Tafi, R., Meola, A., Nuzzo, M., Motti, C., Galfre, G., Felici, F., Cortese, R., Monaci, P. and Nicosia, A. 'Selection of epitopes from phage displayed peptide libraries by using human sera: A new tool for the identification of antigenic and immunogenic mimotopes.' In *Bacterial Protein Toxins*, eds. J. Freer, R. Aitken, J.E. Alouf, G. Boulnois, P. Falmagne, F. Fehrenbach, C. Montecucco, Y. Piemont, R. Rappuoli, T. Wadstrom, and B. Witholt. 415 - 425. Gustav Fischer, Stuttgart (1994)

Friedrich, W. '*Vitamins*.' de Gruyter, Berlin (1988)

Fukuta, S., Magnani, J.L., Twiddy, E.M., Holmes, R.K. and Ginsburg, V. 'Comparison of the carbohydrate binding specificities of cholera toxin and *Escherichia coli* heat-labile enterotoxins LTh-I, LT-IIa and LT-IIb.' *Infection and Immunity* **56 :** 1748 - 1753 (1988)

Furste, J.P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M. and Lanka, E. 'Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector.' *Gene* **48 :** 119 - 131 (1986)

Geysen, H.M., Rodda, S.J. and Mason, T.J. '*A priori* delineation of a peptide which mimics a discontinuous antigenic determinant.' *Molecular Immunology* **23 :** 7 : 709 - 715 (1986)

Gill, D.M. 'The arrangement of subunits in cholera toxin.' *Biochemistry* **15 :** 1242 - 1248 (1976)

Gill, D.M. 'Bacterial toxins : A table of lethal amounts.' *Microbiological Reviews* **46 :** 86 - 94 (1982)

Gill, D.M. and King, C. 'The mechanism of action of cholera toxin in pigeon erythrocyte lysates.' *Journal of Biological Chemistry* **250 :** 6424 - 6432 (1975)

Gill, D.M., Evans, D.J. and Evans, D.G. 'Mechanism of activation of adenylate cyclase *in vitro* by polymyxin-released heat-labile enterotoxin of *Escherichia coli*.' *Journal of Infectious Diseases* **133 :** S103 - S107 (1976)

Gill, D.M. and Rappaport, R.S. 'Origin of the enzymically active A_1 fragment of cholera toxin.' *Journal of Infectious Diseases* **139 :** 674 - 680 (1979)

Gill, D.M., Clements, J.D., Robertson, D.C. and Finkelstein, R.A. 'Subunit number and arrangement in *Escherichia coli* heat-labile enterotoxin.' *Infection and Immunity* **33 :** 677 - 682 (1981)

Goldstein, I.J. 'Carbohydrate specificity of Concanavalin A.' In *Concanavalin A as a tool*, eds. H. Bittiger and H.P. Schnebli. 55 - 65. Wiley, New York (1976)

Gorbach, S.L., Carpenter, C.C.J., Grayson, R., Gryboski, J.D., Guerant, R.L., Hendrix, T.R., Hornick, R.B., Marr, J.S., Morris, J., Polk, B.F., Waters, S.R., Williams, G.W. and Wolfe, M.S. 'Consensus development conference statement.' *Reviews of Infectious Diseases* **8 :** S227 - S233 (1986)

Green, B.A., Neill, R.J., Ruyechan, W.T. and Holmes, R.K. 'Evidence that a new enterotoxin of *Escherichia coli* which activates adenylate cyclase in eucaryotic target cells is not plasmid mediated.' *Infection and Immunity* **41 :** 383 - 390 (1983)

Greenough, W.B., Pierce, N.F., AlAwqati, Q. and Carpenter, C.C.J. 'Stimulation of gut electrolyte secretion by prostaglandins, theophylline and cholera exotoxin.' *Journal of Clinical Investigation* **48 :** 32a (1969)

Greenwood, J., Willis, A.E. and Perham, R.N. 'Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from *Plasmodium falciparum* on a filamentous bacteriophage.' *Journal of Molecular Biology* **220 :** 821 - 827 (1991)

Griffiths, S.L. and Critchley, D.R. 'Characterisation of the binding sites for *Escherichia coli* heat-labile toxin type I in intestinal brush borders.' *Biochimica et Biophysica Acta* **1075 :** 154 - 161 (1991)

Gyles, C.L. '*Escherichia coli* cytotoxins and enterotoxins.' *Canadian Journal of Microbiology* **38 :** 734 - 746 (1992)

Gyles, C.L. and Barnum, D.A. 'A heat-labile enterotoxin from strains of *Escherichia coli* enteropathogenic for pigs.' *Journal of Infectious Diseases* **120 :** 419 - 426 (1969)

Hardy, S.J.S., Holmgren, J., Johansson, S., Sanchez, J. and Hirst, T.R. 'Co-ordinated assembly of multisubunit proteins : Oligomerization of bacterial enterotoxins *in vivo* and *in vitro*.' *Proceedings of the National Academy of Sciences of the United States of America* **85 :** 7109 - 7113 (1988)

Harris, S.L., Craig, L., Mehroke, J.S., Rashed, M., Zwick, M.B., Kenar, K., Toone, E.J., Greenspan, N., Auzanneau, F-I., Marino-Albernas, J-R., Pinto, B.M. and Scott, J.K. 'Exploring the basis of peptide-carbohydrate crossreactivity: Evidence for discrimination by peptides between closely related anti-carbohydrate antibodies' *Proceedings of the National Academy of Sciences of the United States of America* **94 :** 2454 - 2459 (1997)

Hill, R.H. and Stockley, P.G. 'Phage presentation.' *Molecular Microbiology* **20 :** 685 - 692 (1996)

Hirschhorn, N. and Greenough, W.B. 'Progress in oral rehydration therapy.' *Scientific American* **264 :** 16 - 22 (1991)

Hirst, T.R. 'Biogenesis of cholera toxin and related oligomeric enterotoxins.' In *Bacterial toxins and virulence factors in disease*, eds. J. Moss, B. Iglewski, M. Vaughan, and A.T. Tu. 123 - 184. Marcel Dekker, New York (1995)

Hirst, T.R., Hardy, S.J.S and Randall, L.L. 'Assembly *in vivo* of enterotoxin from *Escherichia coli*: : formation of the B subunit oligomer.' *Journal of Bacteriology* **153** : 21 - 26 (1983)

Hirst, T.R., Randall, L.L. and Hardy, S.J.S. 'Cellular location of heat-labile enterotoxin in *Escherichia coli*.' *Journal of Bacteriology* **157 :** 637 - 642 (1984a)

Hirst, T.R., Sanchez, J., Kaper, J.B., Hardy, S.J.S. and Holmgren, J. 'Mechanism of toxin secretion by *Vibrio cholerae* investigated in strains harbouring plasmids that encode heat-labile enterotoxins of *Escherichia coli*.' *Proceedings of the National Academy of Sciences of the United States of America* **81 :** 7752 - 7756 (1984b)

Hirst, T.R. and Holmgren, J. 'Conformation of protein secreted across bacterial outer membranes : a study of enterotoxin translocation from *Vibrio cholerae*.' *Proceedings of the National Academy of Sciences of the United States of America* **84 :** 7418 - 7422 (1987a)

Hirst, T.R. and Holmgren, J. 'Transient entry of enterotoxin subunits into the periplasm occurs during their secretion from *Vibrio cholerae*.' *Journal of Bacteriology* **169 :** 1037 - 1045 (1987b)

Hirst, T.R. and Welch, R.A. 'Mechanisms for secretion of extracellular proteins by gram-negative bacteria.' *Trends in Biological Sciences* **13 :** 265 - 269 (1988)

Hirst, T.R., Nashar, T.O., Eaglestone, S., Lencer, W.I., Webb, H.M. and Yu, J. 'Bacterial and host interactions during the biogenesis, toxicity and immunogenicity of *Escherichia coli* heat-labile enterotoxin.' *Biochemical Society Transactions* **22 :** 306 - 309 (1994)

Hoess, R.H. 'Phage display of peptides and protein domains.' *Current Opinion in Structural Biology* **3 :** 572 - 579 (1993)

Hoess, R., Brinkmann, U., Handel, T. and Pastan, I. 'Identification of a peptide which binds to the carbohydrate-specific monoclonal antibody B3.' *Gene* **128 :** 43 - 49 (1993)

Hofstra, H. and Witholt, B. 'Kinetics of synthesis, processing and membrane transport of heat-labile enterotoxin, a periplasmic protein in *Escherichia coli*.' *Journal of Biological Chemistry* **259 :** 15182 - 15187 (1984)

Holliger, P. and Riechmann, L. 'A conserved infection pathway for filamentous bacteriophages is suggested by the structure of the membrane penetration domain of the minor coat protein g3p from phage fd.' *Structure* **5** : 265 - 275 (1997)

Holmgren, J. 'Comparison of the tissue receptors for *Vibrio cholerae* and *Escherichia coli* enterotoxins by means of gangliosides and natural cholera toxoid.' *Infection and Immunity* **8 :** 851 - 859 (1973)

Holmgren, J., Lonnroth, I. and Svennerholm, L. 'Fixation and inactivation of cholera toxin by G_{M1} ganglioside.' *Scandinavian Journal of Infectious Diseases* **5 :** 77 - 78 (1973)

Holmgren, J., Lindblad, M., Freidman, P., Svennerholm, L. and Myrvold, H. 'Comparison of receptors for cholera and *Escherichia coli* enterotoxins in human intestine.' *Gastroenterology* **89 :** 27 - 35 (1985)

Honda, T., Tsuji, T., Takeda, Y. and Miwatani, T. 'Immunological non-identity of heat-labile enterotoxins from human and porcine enteropathogenic *Escherichia coli*.' *Infection and Immunity* **34 :** 337 - 340 (1981)

Hunt, P.D. and Hardy, S.J. 'Heat-labile enterotoxin can be released from *Escherichia coli* cells by host intestinal factors.' *Infection and Immunity* **59 :** 168 - 171 (1991)

Iannolo, G., Minenkova, O., Petruzzelli, R. and Cesareni, G. 'Modifying filamentous phage capsids: limits in the size of the major capsid protein.' *Journal of Molecular Biology* **248 :** 835 - 844 (1995)

Ichige, A., Oishi, K. and Mizushima, S. 'Isolation and characterisation of mutants of a marine *Vibrio* strain that are defective in production of extracellular proteins.' *Journal of Bacteriology* **170 :** 3537 - 3542 (1988)

Inoue, T., Tsuji, T., Koto, M., Imamura, S. and Miyama, A. 'Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic *Escherichia coli* is identical to that of human strain H 10407.' *FEMS Microbiology Letters* **108 :** 157 - 162 (1993)

Janicot, M., Clot, J. and Desbuquois, B. 'Interactions of cholera toxin with isolated hepatocytes. Effects of low pH, chloroquine and monensin on toxin internalisation, processing and action.' *Biochemical Journal* **253 :** 735 - 743 (1988)

Jespers, L.S., Messens, J.H., de Keyser, A., Eeckhout, D., van den Brande, I., Gansemans, Y.G., Lauwereys, M.J., Vlasuk, G.P. and Stanssens, P.E. 'Surface expression and ligand based selection of cDNA's fused to filamentous phage geneVI.' *Biotechnology* **13 :** 378 - 382 (1995)

Katz, B.A. and Cass, R.T. 'In crystals of complexes of streptavidin with peptide ligands containing the HPQ sequence the pK_a of the peptide histidine is less than 3.0.' *Journal of Biological Chemistry* **272 :** 13220 - 13228 (1997)

Kay, B.K., Adey, N.B., He, Y., Manfredi, J.P., Mataragnon, A.H. and Fowlkes, D.M. 'An M13 phage library displaying random 38-amino-acid peptides as a source of novel sequences with affinity to selected targets.' *Gene* **128 :** 59 - 65 (1993)

Kay, B.K. and Hoess, R.H. 'Principles and applications of phage display.' In *Phage display of peptides and proteins*, eds. B.K. Kay, J. Winter, and J. McCafferty. 21 - 34. Academic Press, London (1996)

Kay, B.K., Winter, J. and McCafferty, J., eds. *Phage display of peptides and proteins : a laboratory manual.* New York: Academic Press, 1996.

Kean, B.H. 'Travellers diarrhoea ; an overview.' *Reviews of Infectious Diseases* **8 :** S111 - S116 (1986)

Khine, A.A. and Lingwood, C.A. 'Capping and receptor-mediated endocytosis of cell-bound verotoxin (shiga-like toxin) 1: Chemical identification of an amino acid in the B subunit necessary for efficient receptor glycolipid binding and cellular internalization.' *Journal of Cellular Physiology* **161 :** 319 - 332 (1994)

Kishchenko, G., Batliwala, H. and Makowski, L. 'Structure of a foreign peptide displayed on the surface of bacteriophage M13.' *Journal of Molecular Biology* **241 :** 208 - 213 (1994)

Koivunen, E., Wang, B. and Ruoslahti, E. 'Isolation of a highly specific ligand for the $\alpha_5\beta_1$ integrin from a phage display library.' *Journal of Cell Biology* **124 :** 373 - 380 (1994)

Kranendonk, O. 'Letter to journal.' *American Journal of Tropical Medicine* **9 :** 536 - 538 (1960)

Kunkel, S.L. and Robertson, D.C. 'Purification and chemical characterisation of the heat-labile enterotoxin produced by enterotoxigenic *Escherichia coli*.' *Infection and Immunity* **25 :** 586 - 596 (1979)

Kurumizaka, H. and Shibata, T. 'Homologous Recognition by Rec A protein using non-equivalent three DNA-strand binding sites.' *Journal of Biochemistry* **119 :** 216 - 213 (1996)

Ladner, R.C. 'Constrained peptides as binding entities.' *Trends in Biotechnology* **13 :** 426 - 430 (1995)

Laemmli, U.K. 'Cleavage of structural proteins during the assembly of the head of bacteriophage T4.' *Nature* **227 :** 680 - 685 (1970)

Ledley, F.D., Mullin, B.R., Lee, G., Aloj, S.M., Fishman, P.H., Hunt, L.T., Dayhoff, M.O. and Kohn, L.D. 'Sequence similarity between cholera toxin and glycoprotein hormones : implications for structure-activity relationship and mechanism of action.' *Biochemical and Biophysical Research Communications* **69 :** 852 - 858 (1976)

Lee, C.M., Chang, P.P., Tsai, S.C., Adamik, R., Price, S.R., Kunz, B.C., Moss, J., Twiddy, E.M. and Holmes, R.K. 'Activation of *Escherichia coli* heat-labile enterotoxins by native and recombinant adenosine diphosphate-ribosylation factors,

20kDa guanine nucleotide-binding proteins.' *Journal of Clinical Investigation* **87 :** 1780 - 1786 (1991)

Leece, R. and Hirst, T.R. 'Expression of the B subunit of *Escherichia coli* heat-labile enterotoxin in a marine *Vibrio* and in a mutant that is pleiotropically defective in the secretion of extracellular proteins.' *Journal of General Microbiology* **138 :** 719 - 724 (1992)

Lencer, W.I., Constable, C., Moe, S., Jobling, M.G., Webb, H.M., Ruston, S., Madara, J.L., Hirst, T.R. and Holmes, R.K. 'Targetting of cholera toxin by KDEL.' *Journal of Cell Biology* **131 :** 951 - 962 (1995a)

Lencer, W.I., Moe, S., Rufo, P.A. and Madara, J.L. 'Transcytosis of cholera toxin subunits across model human membranes.' *Proceedings of the National Academy of Sciences of the United States of America* **92 :** 10094 - 10098 (1995b)

Lencer, W.I., Constable, C., Moe, S., Rufo, P.A., Wolf, A., Jobling, M.G., Ruston, S.P., Madara, J.L., Holmes, R.K. and Hirst, T.R. 'Proteolytic activation of cholera toxin and *Escherichia coli* labile toxin by entry into host epithelial cells - Signal transduction by a protease-resistant toxin variant.' *Journal of Biological Chemistry* **272 :** 15562 - 15568 (1997)

Leong, J., Vinal, A.C. and Dallas, W.S. 'Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons from *Escherichia coli* of human and porcine origin.' *Infection and Immunity* **48 :** 73 - 77 (1985)

Levine, M.M., Kaper, J.B., Black, R.E. and Clements, M.L. 'New knowledge on pathogenesis of bacterial enteric infections as applied to vaccine development.' *Microbiological Reviews* **47 :** 510 - 550 (1983)

Lonnroth, I. and Holmgren, J. 'Subunit structure of cholera toxin.' *Journal of General Microbiology* **76 :** 417 - 427 (1973)

Loregian, A., Hirst, T.R., Marsden, H.S. and Palú, G. 'Use of *Vibrio* spp. for expression of *Escherichia coli* enterotoxin B subunit fusion proteins: Purification and characterisation of a chimera containing a carboxyl--terminal fragment of DNA polymerase from herpes simplex virus type 1.' *Protein Expression and Purification* **8 :** 381 - 389 (1996)

Lowman, H.B., Bass, S.H., Simpson, N. and Wells, J.A. 'Selecting high-affinity binding proteins by monovalent phage display.' *Biochemistry* **30 :** 10832 - 10838 (1991)

Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. 'Protein measurement with the Folin phenol reagent.' *Journal of Biological Chemistry* **193 :** 265 - 275 (1951)

Luzzago, A., Felici, F., Tramontano, A., Pessi, A. and Cortese, R. 'Mimicking of discontinuous epitopes by phage-displayed peptides, I. Epitope mapping of human ferritin using a phage library of constrained peptides.' *Gene* **128 :** 51 - 57 (1993)

Macfarlane, M.G. and Knight, B.C.J.G. *Biochemical Journal* **35 :** 844 - 902 (1941)

Makowski, L. 'Structural constraints on the display of foreign peptides on filamentous bacteriophages.' *Gene* **128 :** 5 - 11 (1993)

Makowski, L. 'Phage display : structure, assembly and engineering of filamentous bacteriophage M13.' *Current Opinion in Structural Biology* **4 :** 225 - 230 (1994)

Malik, P., Terry, T.D., Gowda, L.R., Langara, A., Petukhov, S.A., Symmons, M.F., Welsh, L.C., Marvin, D.A. and Perham, R.N. 'Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage' *Journal of Molecular Biology* **260** : 9 - 21

Marcello, A., Loregian, A., Palú, G. and Hirst, T.R. 'Efficient extracellular production of hybrid *E. coli* heat-labile enterotoxin B subunits in a marine vibrio.' *FEMS Microbiology Letters* **117 :** 47 - 52 (1994)

Markwell, M.A.K., Haas, S.M., Bieber, L.L. and Tolbert, N.E. 'A modification of the Lowry procedure to simplify protein determination in membrane and lipoprotein samples.' *Analytical Biochemistry* **87 :** 206 - 210 (1978)

Martens, C.L., Cwirla, S.E., Lee, R.Y.-W., Whitehorn, E., Chen, E.Y.-F., Bakker, A., Martin, E.L., Wagstrom, C., Gopalan, P., Smith, C.W., Tate, E., Koller, K.J., Schatz, P.J., Dower, W.J. and Barrett, R.W. 'Peptides which bind to E-selectin and block neutrophil adhesion.' *Journal of Molecular Biology* **270 :** 21129 - 21136 (1995)

McConnell, S.J., Kendall, M.L., Reilly, T.M. and Hoess, R.H. 'Constrained peptide libraries as a tool for finding mimotopes.' *Gene* **151 :** 115 - 118 (1994)

McLafferty, M.A., Kent, R.B., Ladner, R.C. and Markland, W. 'M13 bacteriophage displaying disulphide-constrained microproteins.' *Gene* **128 :** 29 - 36 (1993)

McNeill, W.H. *'Plagues and People.'* Blackwell, Oxford (1976)

Merritt, E.A., Sarfaty, S., van den Akker, F., L'Hoir, C.L., Martial, J.A. and Hol, W.G.J. 'Crystal structure of cholera toxin B-pentamer bound to receptor G_{M1} pentasaccharide.' *Protein Science* **3 :** 166 - 175 (1994a)

Merritt, E.A., Sixma, T.K., Kalk, K.H., van Zanten, B.A.M. and Hol, W.G.J. 'Galactose-binding site in *Escherichia coli* heat-labile enterotoxin (LT) and cholera toxin (CT).' *Molecular Microbiology* **13 :** 745 - 753 (1994b)

Merritt, E.A., Sarfaty, S., Chang, T., Palmer, L.M., Jobling, M.G., Holmes, R.K. and Hol, W.G.J. 'Surprising leads for a cholera toxin receptor-binding antagonist: Crystallographic studies of CTB mutants.' *Structure* **3 :** 561 - 570 (1995)

Model, P. and Russel, M. 'Filamentous bacteriophage.' In *The Bacteriophages*, ed. R. Calendar. 375 - 456. 2. Plenum Press, New York (1988)

Montecucco, C., Papini, E. and Schiavo, G. 'Molecular models of toxin membrane translocation.' In *Sourcebook of Bacterial Protein Toxins*, eds. J.E. Alouf and J.H. Freer. 45 - 56. Academic Press, London (1991)

Montecucco, C., Papini, E. and Schiavo, G. 'Bacterial protein toxins penetrate cells *via* a four-step mechanism.' *FEBS Letters* **346 :** 92 - 98 (1994)

Moon, H.W., Whipp, S.C., Engstrom, G.W. and Baetz, A.L. 'Response of the rabbit ileal loop to cell-free products from *Escherichia coli* enteropathogenic for swine.' *Journal of Infectious Diseases* **121 :** 182 - 187 (1970)

Moss, J. and Richardson, S.H. 'Activation of adenylate cyclase by heat-labile enterotoxin. Evidence of ADP-ribosyl transferase activity similar to that of choleragen.' *Journal of Clinical Investigation* **62 :** 281 - 285 (1978)

Moss, J. and Vaughan, M. 'Activation of cholera toxin and *Escherichia coli* heat-labile enterotoxins by ADP-ribosylation factors, a family of 20 kDa guanine nucleotide binding proteins.' *Molecular Microbiology* **5 :** 2621 - 2627 (1991)

Neer, E.J. and Clapham, D.E. 'Role of G protein subunits in transmembrane signalling.' *Nature* **333 :** 129 - 134 (1988)

Neill, R.J., Ivins, B.E. and Holmes, R.K. 'Synthesis and secretion of the plasmid-coded heat-labile enterotoxin of *Escherichia coli* in *Vibrio cholerae*.' *Science* **221 :** 289 - 291 (1983)

O'Neil, K.T., Hoess, R.H., Jackson, S.A., Ramachandran, N.S., Mousa, S.A. and DeGrado, W.F. 'Identification of novel peptide antagonists for GP IIb / IIIa from a conformationally constrained phage peptide library.' *Proteins* **14 :** 509 - 515 (1992)

Oishi, K., Yokoshima, S., Tomiyama, T. and Aisa, K. 'Exohemagglutinins: new products of vibrios.' *Applied and Environmental Microbiology* **38 :** 169 - 172 (1979)

Oldenburg, K.R., Loganathan, D., Goldstein, I.J., Schultz, P.G. and Gallop, M.A. 'Peptide ligands for a sugar-binding protein isolated from a random peptide library.' *Proceedings of the National Academy of Sciences of the United States of America* **89 :** 5393 - 5397 (1992)

Orlandi, P.A. and Fishman, P.H. 'Orientation of cholera toxin bound to target cells.' *Journal of Biological Chemistry* **268 :** 17038 - 17044 (1993)

Osek, J., Lebens, M. and Holmgren, J. 'Improved medium for large-scale production of recombinant cholera toxin B subunit for vaccine purposes.' *Journal of Microbiological Methods* **24 :** 117 - 123 (1995)

Overbye Michel, L., Sandkvist, M. and Bagdasarian, M. 'Specificity of the protein secretory apparatus: Secretion of the heat-labile enterotoxin B subunit pentamers by different species of Gram - bacteria.' *Gene* **152 :** 41 - 45 (1995)

Parmley, S.F. and Smith, G.P. 'Antibody-selectable filamentous fd phage vectors : Affinity purification of target genes.' *Gene* **73 :** 305 - 308 (1988)

Pearson, G.D.N. and Mekalanos, J.J. 'Molecular cloning of *Vibrio cholerae* enterotoxin genes in *Escherichia coli* K-12.' *Proceedings of the National Academy of Sciences of the United States of America* **79 :** 2976 - 2980 (1982)

Peterson, J.W. and Ochoa, L.G. 'Role of prostaglandins and cAMP in the secretory effects of cholera toxin.' *Science* **245 :** 857 - 859 (1989)

Petrenko, V.A., Smith, G.P., Gong, X. and Quinn, T. 'A library of organic landscapes on filamentous phage.' *Protein Engineering* **9 :** 797 - 801 (1996)

Phalipon, A., Folgori, A., Arondel, J., Sgaramella, G., Fortugno, P., Cortese, R., Sansonetti, P.J. and Felici, F. 'Induction of anti-carbohydrate antibodies by phage library-selected peptide mimics' *European Journal of Immunology* **27** : 2620 - 2625 (1997)

Pickett, C.L., Twiddy, E.M., Belisle, B.W. and Holmes, R.K. 'Cloning of genes that encode a new heat-labile enterotoxin of *Escherichia coli*.' *Journal of Bacteriology* **165 :** 348 - 352 (1986)

Pronk, S.E., Hofstra, H., Groendijk, H., Kingma, J., Swarte, M.B.A., Dorner, F., Drenth, J., Hol, W.G.J. and Witholt, B. 'Heat-labile enterotoxin of *Escherichia coli*.' *Journal of Biological Chemistry* **260 :** 13580 - 13584 (1985)

Rappuoli, R. and Pizza, M. 'Structure and evolutionary aspects of ADP-ribosylating toxins.' In *Sourcebook of Bacterial Protein Toxins*, eds. J.E. Alouf and J.H. Freer. 1 - 21. Academic Press, London (1991)

Read, R.J. and Stein, P.E. 'Toxins.' *Current Opinion in Structural Biology* **3 :** 853 - 860 (1993)

Ribi, H.O., Ludwig, D.S., Mercer, K.L., Schoolnik, G.K. and Kornberg, R.D. 'Three-dimensional structure of cholera toxin penetrating a lipid membrane.' *Science* **239 :** 1272 - 1276 (1988)

Richardson, S.H. 'Factors influencing *in vitro* permeability factor production by *Vibrio cholerae*.' *Journal of Bacteriology* **100 :** 27 - 34 (1969)

Riechmann, L. and Holliger, P 'The C-terminal domain of TolA is the coreceptor for filamentous phage infection of *E. coli*. *Cell* **90** : 351 - 360 (1997)

Rodi, D.J. and Makowski, L. 'Phage-display technology - finding a needle in a vast molecular haystack.' *Current Opinion in Biotechnology* **10** : 87 - 93 (1999)

Sack, R.B. 'Human diarrheal disease caused by enterotoxigenic *Escherichia coli*.' *Annual Review of Microbiology* **29 :** 333 - 353 (1975)

Sack, R.B., Gorbach, S.L., Banwell, J.G., Jacobs, B., Chatterjee, B.D. and Mitra, R.C. 'Enterotoxic *Escherichia coli* isolated from patients with severe cholera-like disease.' *Journal of Infectious Diseases* **123 :** 378 - 385 (1971)

Sandkvist, M., Hirst, T.R. and Bagdasarian, M. 'Alterations at the carboxyl terminal change assembly and secretion properties of the B subunit of *Escherichia coli* heat-labile enterotoxin.' *Journal of Bacteriology* **169 :** 4570 - 4576 (1987)

Sandkvist, M., Hirst, T.R. and Bagdasarian, M. 'Minimal deletion of amino acids from the carboxyl terminus of the B subunit of heat-labile enterotoxin causes defects in its assembly and release from the cytoplasmic membrane of *Escherichia coli*.' *Journal of Biological Chemistry* **265 :** 15239 - 15244 (1990)

Sandkvist, M., Morales, V. and Bagdasarian, M. 'A protein required for secretion of cholera toxin through the outer membrane of *Vibrio cholerae*.' *Gene* **123 :** 81 - 86 (1993)

Sandkvist, M., Bagdasarian, M., Howard, S.P. and DiRita, V.J. 'Interaction between the autokinase EpsE and EpsL in the cytoplasmic membrane is required for extracellular secretion in *Vibrio cholerae*.' *EMBO journal* **14 :** 1664 - 1673 (1995)

Scott, J.K. 'Discovering peptide ligands using epitope libraries.' *Trends in Biological Sciences* **17 :** 241 - 245 (1992)

Scott, J.K. and Smith, G.P. 'Searching for peptide ligands with an epitope library.' *Science* **249 :** 386 - 390 (1990)

Scott, J.K., Loganathan, D., Easley, R.B., Gong, X. and Goldstein, I.J. 'A family of concanavalin A-binding peptides from a hexapeptide epitope library.' *Proceedings of*

the National Academy of Sciences of the United States of America **89 :** 5398 - 5402 (1992)

Sears, C.L. and Kaper, J.B. 'Enteric bacterial toxins : Mechanisms of action and linkage to intestinal secretion.' *Microbiological Reviews* **60 :** 167 - 215 (1996)

Sequenase Protocols. '*Step by step protocols for DNA sequencing with Sequenase version 2.0 T7 DNA polymerase.*' 9th Edition ed., United States Biochemical, (1994)

Sixma, T.K., Pronk, S.E., Kalk, K.H., Wartna, E.S., van Zanten, B.A.M., Witholt, B. and Hol, W.G.J. 'Crystal structure of cholera toxin-related heat-labile enterotoxin from *E. coli*.' *Nature* **351 :** 371 - 377 (1991)

Sixma, T.K., Pronk, S.E., Kalk, K.H., van Zanten, B.A.M., Berghuis, A.M. and Hol, W.G.J. 'Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography.' *Nature* **355 :** 561 - 564 (1992)

Sixma, T.K., Kalk, K.H., van Zanten, B.A.M., Dauter, Z., Kingma, J., Witholt, B. and Hol, W.G.J. 'Refined structure of *Escherichia coli* heat-labile enterotoxin, a close relative of cholera toxin.' *Journal of Molecular Biology* **230 :** 890 - 918 (1993)

Smith, G.P. 'Filamentous fusion phage : Novel expression vectors that display cloned antigens on the virion surface.' *Science* **228 :** 1315 - 1317 (1985)

Smith, G.P. 'Filamentous phage assembly : Morphogenetically defective mutants that do not kill the host.' *Virology* **167 :** 156 - 165 (1988a)

Smith, G.P. 'Filamentous phages as cloning vectors.' In *Vectors : A survey of molecular cloning vectors and their uses*, eds. R.L. Rodriguez and D.T. Denhardt. 61 - 83. Butterworth, Stoneham (1988b)

Smith, G.P. 'Surface display and peptide libraries.' *Gene* **128 :** 1 - 2 (1993)

Smith, G.P. 'Foreword : 1988 - a year of discovery.' In *Phage display of peptides and proteins : a laboratory manual*, eds. B.K. Kay, J. Winter, and J. McCafferty. xvii - xix. Academic Press, New York (1996)

Smith, G.P., Schultz, D.A. and Ladbury, J.E. 'A ribonuclease S-peptide antagonist discovered with a bacteriophage display library.' *Gene* **128 :** 37 - 42 (1993)

Smith, G.P. and Petrenko, V.A. 'Phage display.' *Chemical Reviews* **97 :** 391 - 410 (1997)

Smith, H.W. and Halls, S. 'Studies on *Escherichia coli* enterotoxin.' *Journal of Pathology and Bacteriology* **93 :** 531 - 543 (1967)

Smith, H.W. and Halls, S. 'The transmissible nature of the genetic factor in *Escherichia coli* that controls enterotoxin production.' *Journal of General Microbiology* **52 :** 319 - 334 (1968)

Smith, H.W. and Gyles, C.L. 'The relationship between two apparently different enterotoxins produced by enteropathic strains of *Escherichia coli* of porcine origin.' *Journal of Medical Microbiology* **3 :** 387 - 401 (1970)

Smith, H.W. and Sack, R.B. 'Immunologic cross reactions of enterotoxins from *Escherichia coli* and *Vibrio cholerae*.' *Journal of Infectious Diseases* **127 :** 164- 170 (1973)

Spangler, B.D. 'Structure and function of cholera toxin and the related *Escherichia coli* heat-labile enterotoxin.' *Microbiological Reviews* **56 :** 622 - 647 (1992)

Stephen, C.W., Helminen, P. and Lane, D.P. 'Characterisation of epitopes on human p53 using phage-displayed peptide libraries : Insights into antibody-peptide interactions.' *Journal of Molecular Biology* **248 :** 58 - 78 (1995)

Stephen, J. and Pietrowski, R.A. '*Bacterial Toxins*.' : Aspects of Microbiology, eds. J.A. Cole and C.J. Knowles. Thomas Nelson, Walton-on-Thames (1981)

Svennerholm, A. and Holmgren, J. 'Identification of *Escherichia coli* heat-labile enterotoxin by means of a ganglioside immunosorbent assay (G_{M1}) procedure.' *Current Microbiology* **1 :** 19 - 23 (1978)

Tayot, J.-L., Holmgren, J., Svennerholm, L., Lindblad, M. and Tardy, M. 'Receptor-specific large-scale purification of cholera toxin on silica beads derivatized with lysoG_{M1} ganglioside.' *European Journal of Biochemistry* **113 :** 249 - 258 (1981)

Tomasi, M. and Montecucco, C. 'Lipid insertion of cholera toxin after binding to G_{M1}-containing liposomes.' *Journal of Biological Chemistry* **256 :** 11177 - 11181 (1981)

Uesaka, Y., Otsuka, Y., Lin, Z., Yamasaki, S., Yamaoka, J., Kurazono, H. and Takeda, Y. 'Simple method of purification of *Escherichia coli* heat-labile enterotoxin

and cholera toxin using immobilised galactose.' *Microbial Pathogenesis* **16 :** 71 - 76 (1994)

van Heyningen, S. 'The ring on a finger.' *Nature* **351 :** 351 (1991)

van Heyningen, W.E. 'General characteristics.' In *Microbial Toxins,* eds. S.J. Ajl, S. Kadis, and T.C. Montie. 1 - 28. 1. Academic Press, New York (1970)

van Heyningen, W.E., Carpenter, C.C.J., Pierce, N.F. and Greenough, W.B. 'Deactivation of cholera toxin by ganglioside.' *Journal of Infectious Diseases* **124 :** 415 - 418 (1971)

Vasil, M.L., Holmes, R.K. and Finkelstein, R.A. 'Conjugal transfer of a chromosomal gene determining production of enterotoxin in *Vibrio cholerae*.' *Science* **187 :** 849 - 850 (1975)

Veh, R.W., Corfield, A.P., Sander, M. and Schauer, R. 'Neuraminic acid-specific modification and tritium labelling of gangliosides.' *Biochimica et Biophysica Acta* **486 :** 145 - 160 (1977)

Vieira, J. and Messing, J. 'Production of single-stranded plasmid DNA.' In *Methods in Enzymology*, eds. R. Wu and L. Grossman. 3 - 11. 153. Academic Press, London (1987)

Weber, P.C., Pantoliano, M.W. and Thompson, L.D. 'Crystal structure and ligand-binding studies of a screened peptide complexed with streptavidin.' *Biochemistry* **31 :** 9350 - 9354 (1992)

Webster, R.E. 'Biology of the filamentous bacteriophage.' In *Phage display of peptides and proteins*, eds. B.K. Kay, J. Winter, and J. McCafferty. 1 - 20. Academic Press, London (1996)

Wilchek, M. and Beyer, E.A. 'The avidin-biotin complex in bioanalytical applications.' *Analytical Biochemistry* **171 :** 1 - 32 (1988)

Wilson, D.R. and Finlay, B.B. 'Phage display: applications, innovations and issues in phage and host biology' *Canadian Journal of Microbiology* **44** : 313 - 329 (1998)

Winter, G., Griffiths, A.D., Hawkins, R.E. and Hoogenboom, H.R. 'Making antibodies by phage display technology.' *Annual Review of Immunology* **12 :** 433 - 455 (1994)

Wolf, A.A., Nibbering, K., Madara, J.L., Jobling, M.G., Holmes, R.K. and Lencer, W.I. 'Association of cholera toxin (CT) with caveolae in human intestinal T84 cells; Possible role in signal transduction.' *Gastroenterology* **112 :** SS, A420 (1997)

Yamamoto, T. and Yokota, T. 'Sequence of heat-labile enterotoxin of *Escherichia coli* pathogenic for humans.' *Journal of Bacteriology* **155 :** 728 - 733 (1983)

Young, D.B. and Broadbent, D.A. 'Biochemical characterisation of extracellular proteases from *Vibrio cholerae*.' *Infection and Immunity* **37 :** 875 - 883 (1982)

Yu, J., Webb, H. and Hirst, T.R. 'A homologue of the *Escherichia coli* DsbA protein involved in disulphide bond formation is required for enterotoxin biogenesis in *Vibrio cholerae*.' *Molecular Microbiology* **6 :** 1949 - 1958 (1992)

Yu, J., McLaughlin, S., Freedman, R.B. and Hirst, T.R. 'Cloning and active site mutagenesis of *Vibrio cholerae* DsbA, a periplasmic enzyme that catalyses disulfide bond formation.' *Journal of Biological Chemistry* **268** : 4326 - 4330 (1993)

Zacher, A.N., Stock, C.A., Golden, J.W. and Smith, G.P. 'A new filamentous phage cloning vector: fd-tet.' *Gene* **9** : 127 - 140 (1980)

GLASGOW UNIVERSITY LIBRARY