A Chemical Genomics Approach to Human Drug Target Discovery

With Test Of Principle Using Simvastatin

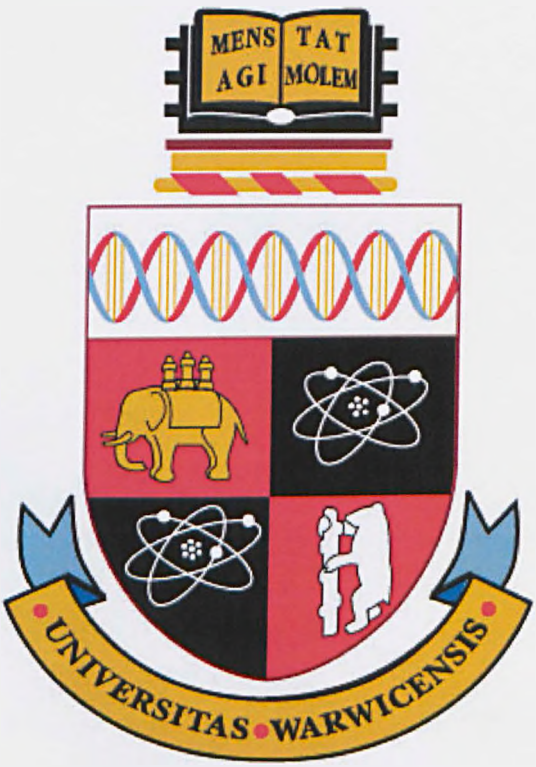
Katherine Casey-Green, MChem

Student Number: 0312091

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Chemistry

University Of Warwick, Department of Chemistry



January 2011

Contents

Table of Figures

Table of Structures and Schemes

Table of Tables

Table of Appendices

Acknowledgements

With thanks to Dr. Andy Marsh, Dr. Paul Taylor, Dr. Suzy Dilly and Professor Donald Singer for help, guidance, counsel, instruction and supervision over the three (and then some) years. Especially for giving up your time over the Christmas period to look over this thesis. Thank you all.

To Dr. Sweta Ladwa and Dr. Junlong Zhang for their instruction in microbiology techniques without which I would have been totally out of my depth.

Also to Mr. Wade Dimitri, for making tissue collection possible at the source, and his theatre team at the University Hospital, especially Faye and Alistair who looked after me, made me feel welcome and instructed me in the ways of the theatre. I appreciate every moment of time you gave me, as I know it is very precious to you.

To Ashan, David, James and John, who gave up their time to collected essential informed consent from individuals before their surgeries.

To Richard Napier for his level headed advice on phage behaviour, and Jesper Gronlund for spending an afternoon with me trying to photograph elusive phage plaques.

Lastly to Dave, Sarah, Nick, Jason, Tim and Ross for reminding me how to be human, and making sure I practised every once in a while.

Declaration

This thesis is the work of the author and has not been submitted for a degree at another university.

Abstract

Understanding drug-protein interactions and downstream effects of these interactions underpins much of clinical pharmacology. By studying the protein targets of a small molecule we can learn about the action of this compound in the body, and this information can lead to greater understanding of mechanisms and clinical effects. We probed the polypeptide interactions of five small molecules using a human vascular phage display library, with the intention to elucidate previously unknown protein targets for the small molecule in human vasculature. A method for studying the chemical nature of this interaction was also developed.

The photochemical immobilisation system used - Magic Tag® - was developed as a means of immobilising small molecules without concealing any facet of the molecule from the interaction study. Five different photochemistries are displayed in a multi-well format, to maximise diversity in the display of the small molecule. A human vascular tissue T7Select® bacteriophage display library was prepared from internal mammary artery tissues donated from patients undergoing coronary artery bypass surgery. Biopanning of the immobilised small molecule against this library allowed hypothesis generating analysis of small molecule-polypeptide interactions.

Because of the non-selective nature of the photochemical immobilisation of the small molecule, several regioisomers might be expected to form on the Magic Tag® surface. To be able to connect a protein interaction with a specific face of the small molecule, analysis of this regio-non-specific interaction must be undertaken. For this purpose a cleavable resin analogue of the Magic Tag® surface was prepared.

Abbreviations

°C	degrees celcius
δ	shift in ppm (NMR)
µL	microlitre
Acc.	accession
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
br	broad (NMR/IR)
Chem	chemistry
CHN	elemental analysis (carbon, hydrogen, nitrogen)
cm	centimetre(s)
cm^{-1}	wavenumber(s)
conc.	concentration
Contig	contiguous region
COX	cyclooxygenase
d	doublet (NMR)
DCM	dichloromethane
dd	doublet of doublets (NMR)
DHP	dihydropyran
DI RO	deionised reverse osmosis purification (water)
DIPEA	diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	Dimethylsulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EBI	European Bioinformatics Institute
EcoRI	endonuclease isolated from E.coli
EDC	N-Ethyl-N′-(3-dimethylaminopropyl)carbodiimide hydrochloride
ELISA	enzyme-linked immunosorbant assay
eq.	equivalents
ESI	electrospray ionisation
Et	Ethyl group
g	grams
h/hr(s)	hour(s)
HinDIII	restriction enzyme isolated from Haemophilus influenzae
HVTL	human vascular tissue library
Hz (MHz)	hertz (megahertz)
ID	identity
IL-1	interleukin-1
IPTG	isopropyl b-D-1-thiogalactopyranoside
IR	infrared
J	coupling constant (NMR)
LB	lysogeny broth
LCMS	liquid chromatography/mass spectroscopy
m	multiplet (NMR)

M	molar
m.p.	melting point
mg	milligrams
milliQ	purification/deionisation system by millipore corp.
min	minutes
mL	millilitre
MLC2	myosin regulatory light chain 2
mM	millimolar
mmol	millimole(s)
MN	number average molecular weight
mol	mole(s)
mRNA	messenger ribonucleic acid
MS	mass spectroscopy
MW	molecular weight
NCBI	National Center for Biotechnology Information
Neg	negative (control)
NMR	nuclear magnetic resonance
No.	number
NSAID	non-steroidal antiinflammatory drug
PBS	phosphate buffered saline
PCA	Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-4-amidobenzophenone
PCAN	PCA control well
PCB	Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-4-hydroxybenzophenone
PCBN	PCB control well
PCC	Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-1-(4-azidophenyl) 4-amidobutyric acid
PCCN	PCC control well
PCD	Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-4-isothiocyanophenyl azide
PCDN	PCD control well
PCE	Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-3-(3-acetoxyphenyl)-3-(trifluoromethyl)-3H-diazirine
PCEN	PCE control well
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming units
Phth	phthalimide
Pos	positive
RefSeq	NCBI Reference Sequence
Rf	retention factor
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
Seq	sequence
Simva	simvastatin
SPE	solid phase extraction
t	triplet (NMR)
TAE	Tris/acetate/EDTA buffer
TBS	Tris buffered saline

TEA	Triethylamine
TEG	triethylene glycol
TFA	trifluoroacetic acid
THF	tetrahydrofuran
THP	tetrahydropyran
TQEF	Teaching Quality and Education Fund
tRNA	transfer ribonucleic acid
Ts	p-toluenesulphonyl group
TXA_2	thromboxane A2
TXB_2	thromboxane B2
UTY	ubiquitously transcribed tetratricopeptide
UV	ultraviolet

Chapter 1 Introduction

1.1 Context

Understanding the interactions between drugs and proteins, along with downstream effects of these interactions, underpins much of clinical pharmacology, the science and clinical use of drugs.[1, 2] By studying the protein targets of a small molecule we can learn about the action of this compound in the body, and this information can lead to greater understanding of mechanisms and clinical effects. However, by the nature of conventional drug discovery methods, many drugs are discovered without knowledge of their interactions *in vivo*, rather on the basis of their ability to bring about a biological response.[3, 4] Understanding biological targets *in vivo* and fuller mode of action is often a later stage in the pharmacological study of a small molecule therapeutic. Unexpected or unintended activities of these small molecules can lead to off-target effects and hidden phenotypes,[5] whether positive - leading to potential clinical benefit - or detrimental. An understanding of the proteins and pathways contributing to drug activity can uncover hidden phenotypes, indicating unforeseen effects and potentially useful new applications of drugs.

This introduction will describe background to the need for development of new and safer therapeutics, especially for cardiovascular targets, current network pharmacology approaches to drug-target discovery and key methodological aspects of the research, including candidate drug selection, chemical genomic technologies and photochemical immobilisation system based biopanning approaches.[6]

1.2 Target-based Drug Discovery and Network Pharmacology

From the sequencing of the human genome at the turn of the century until 2005, in a move from traditional physiology-based approaches to drug discovery, target-based drug discovery established itself as the new paradigm in the genomic era.[4, 7, 8] Target-based drug discovery approaches made the assumption that a one-drug one-target therapeutic would give fewest off-target and detrimental side effects and allow greater screening throughput than the cell or animal based assays used in physiology based approaches.[7]

Unfortunately, most clinically useful drugs in fact have multiple targets, as is becoming more widely appreciated. The phenotypic robustness and network structure of most biological pathways has meant that targeted drug discovery has not led to the expected increase in new, better drugs. The robustness of phenotype can be understood in terms of contingent functions and alternative compensatory signalling routes. The major sources of attrition in target-based drug development are insufficient efficacy and unacceptable toxicity.[9]

Moving forward with the lessons learned from failing target-led approaches, a new approach beginning to take hold is network pharmacology.

The three aspects of the multidisciplinary field are biological networks, as described by Hopkins,[9] where the biological and kinetic profile of the drug is considered above the individual validation of targets of combinations of targets; polypharmacology, as considered by Tzakos *et al*,[10] where the many targets of any one drug are studied and considered as a gross effect against the patient; and chemogenomics as studied by Mestres *et al*,[11, 12] where large amounts of interaction data for small molecule drugs and protein targets is collected to allow virtual screening and further study using computational chemistry techniques.

This collected approach develops the understanding of the polypharmacology of the small molecule drug - that is, the action of the therapeutic against multiple protein targets in the body - and its effect on biological networks. The study of the network pharmacology is becoming recognised as a plausible way to improve efficacy and understand toxicity.[13]

1.3 Selection of Candidate Therapeutics

Globally, cardiovascular diseases are the number one single cause of death. Cardiovascular disease represented 30% of a global distribution of deaths in 2005. Of these deaths, 7.6 million were due to heart attack and 5.7 million were due to stroke. 80% of the deaths occurred in low and middle income countries.[14]

In atherosclerosis – one of the largest components of cardiovascular disease, present in the initial stages even in childhood[15] – dysfunction of the vascular endothelium through injury or previous medical condition leads to the formation of 'fatty streaks' in the intima of the vessel. The vascular wall (Figure 1.3.1) and the endothelium in particular are undergoing a constant process of injury and repair in response to mechanical and chemical injury.[16] The progression of an atherosclerotic plaque is described in Figure 1.3.2.[17] The accumulation of low density lipoprotein (LDL) in the intima (1) which is modified by oxidation or glycation (2), is followed by the incorporation of monocytes by adhesion molecules on the endothelial surface (3). These monocytes form macrophages upon combination with the LDL present (4), and lead to the

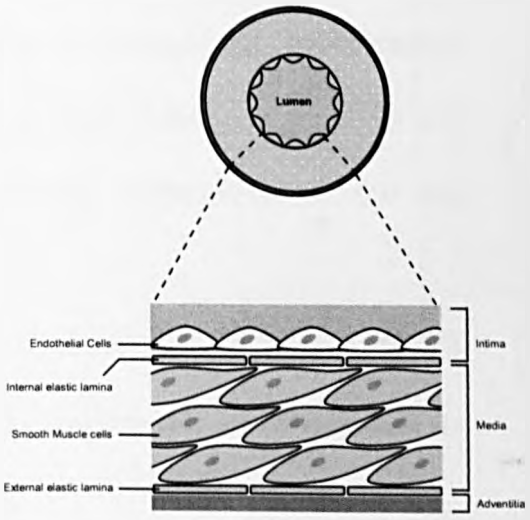


Figure 1.3.1: The vessel wall

formation of foam cells. Smooth muscle cells migrate into the intima (5) and proliferate (6), forming an extracellular matrix which promotes the growth of the atherosclerotic plaque through accumulation of foam cells (7). Within the plaque, a

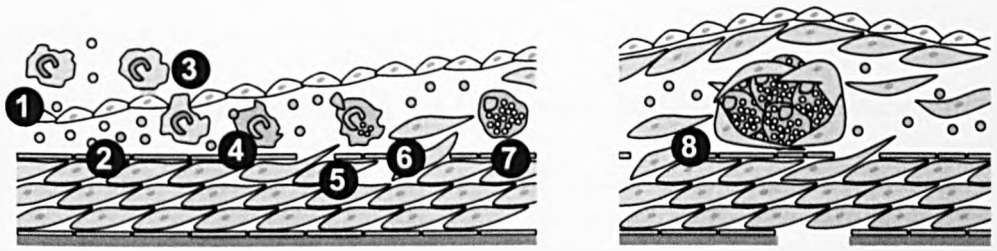
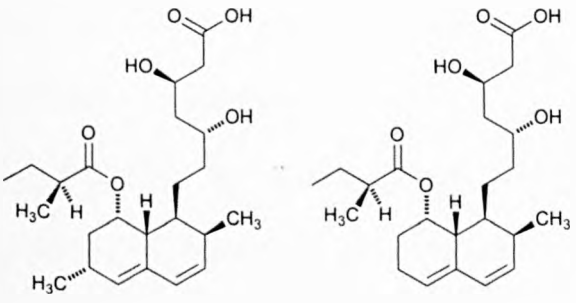


Figure 1.3.2: Progression of the atherosclerotic plaque in the vessel wall

lipid core is contained by a fibrous cap of migrated smooth muscle cells (8). This plaque can eventually reduce blood flow leading to ischemia or rupture leading to infarction or stroke.[17]

Very low density lipoprotein (VLDL) and low density lipoprotein (LDL) carry cholesterol into the arterial wall, an initiator for atheroma, and high density lipoprotein (HDL) clears cholesterol.[18] Dysregulation of these lipoproteins when coupled with other cardiovascular risk factors such as advanced age, hypertension, diabetes or smoking, lead to early manifestations of CVD. Treatment for CVD will often consist of a combination of therapies, including treatment of the CV risk factors and lipid management.[18]

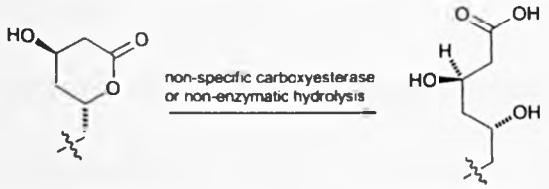
The statins as a drug family are primarily HMG-CoA reductase inhibitors. HMG-CoA reductase is an enzyme that mediates the conversion of HMG-CoA to mevalonate, a limiting step in cholesterol synthesis.



lovastatin (hydroxyacid form) mevastatin (hydroxyacid form)

Structures 1.3.1 - 2: The first statins extracted in the 1970s, lovastatin and mevastatin, displayed here in the active hydroxyl acid form

Mevastatin and lovastatin (Structures 1.3.1 - 2) were the first statins extracted in the 1970s from *Penicillium citrilium* and *Aspergillus terreus* fungus respectively.[19] Since then, several more statins have been extracted, and several synthetic statin family members have been prepared. All of these family members feature an open chain lactone (hydroxy acid) moiety which fills the HMG-CoA binding pocket, inhibiting reductase activity. In some structures a lactone pro-drug is delivered, which is metabolised by the cytochrome P450 enzymes CYP3A4 and, more rarely, CYP2C8 in the intestine wall, the liver and to some extent in the plasma.[20]



lactone form hydroxy acid form HMG-CoA

Structures 1.3.3 - 5: Statin pro drug and active form, and HMG-CoA, which the statins competitively inhibit

Statins as a class have very complex pleiotropic effects[21, 22] as inhibition of HMG-CoA reductase also blocks the production of isoprenoid intermediates farnesyl pyrophosphate and geranyl-geranyl pyrophosphate. These pyrophosphates are required for isoprenylation of specific proteins, particularly in signal transduction pathways and serve as important lipid attachments including the GTP-binding protein Ras, and other Ras-like proteins. The isoprenoids also have a role in the composition of transfer RNA, in glycoprotein synthesis, cell membrane function, electron transport, cellular signalling and differentiation.

There is also a wide range of apparently non-lipid related pleiotropic effects, including improvement of endothelial dysfunction, increased nitric oxide bioavailability, antioxidant effects, anti-inflammatory and immunomodulatory properties, stabilization of atherosclerotic plaques and inhibition of cardiac hypertrophy, many of which have no determined physiological basis as of yet.[23] It is these polypharmacological effects that make statins of interest to screen against a polypeptide library.

Cardiovascular disease has a major inflammatory component and inflammatory mechanisms contribute to atheroma formation.[24] Inflammation is primarily triggered by mechanical or microbial damage, and leads to the release of chemical signalling molecules called cytokines, which mediate further inflammation. These molecules lead to vasodilation, increased capillary permeability, monocyte accumulation, stimulation of sensory receptors, fever and contraction of smooth muscle.[25] These signalling molecules include prostaglandins, leukotrienes, prostacyclins and thromboxanes. Prostaglandins and thromboxanes are synthesised from phospholipids in the membrane of the endoplasmic reticulum. They stem from arachidonic acid, which interacts with the enzyme cyclooxygenase (COX).[26] There are three forms of the COX enzyme - COX-1 which is a housekeeping enzyme involved in gastric protection, COX-2 which is only found in inflammatory cells,[26] and COX-3 which is a transcript variant of COX-1 retaining one more intron.[27]

These inflammatory interactions are required for the stress response of cells, but in diseases such as arthritis and cardiovascular disease,[24] severe inflammation can be damaging to the bodily systems and can lead to chronic pain.

By inhibiting cyclooxygenase, hence blocking production of prostaglandins and thromboxane, an anti-inflammatory action can be obtained. Many non-steroidal anti-

inflammatory drugs (NSAIDs) are non-specific COX inhibitors. Due to the necessary actions of COX-1, these NSAIDs have adverse effects including gastric intestinal upset, peptic ulcers, allergy and renal impairment.[26] Specific COX-2 inhibitors have been developed, but unfortunately these appear to have other serious side effects, including detrimental cardiovascular effects.[25]

COX inhibition is not the only route to anti-inflammatory action in non-steroidal drugs. Inhibition of other pro-inflammatory bioactives, such as the interleukin family of polypeptides, and oxidative species such as superoxides can also give anti-inflammatory therapeutic action.

Despite the inflammatory effects of cardiovascular disease, non-steroidal anti-inflammatory drugs, especially COX-2 inhibitors, are known to lead to an increased incidence of cardiovascular disease.[28] The range of interest in these anti-inflammatory molecules led to interest in their binding partners and the decision to biopan a sample of these molecules against a protein library.

1.4 Affinity Chromatography and Immobilisation

Frequently, drug discovery starts with a disease target, identified either through clinical association or by outcomes of genomic assays,[29] and is followed by screening of this drug target against a combinatorial library of small molecules constructed on a solid support, searching for the molecular recognition of a non-covalently binding ligand. This method is the basis of affinity chromatography.[29]

Pre-validation of targets, however, remains a highly complex and time and labour-intensive problem and small molecules that are discovered in the lab to be active against validated targets do not always translate into the clinic.[9] By screening a natural product library for a desired phenotypic response or binding event, the

cellular proteins binding can be determined directly, removing this stage of the analysis.[30] Techniques for biological screening of libraries of compounds are the basis for many affinity chromatography experiments, used in drug discovery as well as binding assays and protein purification.

There has been extensive previous research into the chemical immobilisation of small molecules on to solid supports for biological investigation. Appropriate immobilisation allows screening to determine structure-activity relationships of ligand and receptor from high affinity specific recognition of proteomic targets.[31-34] Regardless of what type of biomacromolecule library is being used as a probe, the small molecule is almost always immobilised before the assay, or during the assay by a pull down method, to allow the most efficient removal of the non-binding material from the assay by washing. To immobilise the small molecule, or group of small molecules under study, a polymer surface or bead[31, 35] or occasionally an activated glass bead[36] or slide[37] is used. The surface itself is often derivatised to give specifically directed stereochemistry[37], or in preparation for forming a compound library through diversity oriented synthesis,[30, 32, 38] with the aim of preserving the biological activity of the small molecule.

One method for maximising the biologically accessible surface of the immobilised small molecule is the use of photochemical functionalities. These are most commonly used in photocrosslinking chemistries - where small molecule-protein,[39] protein-protein[40] and oligonucleotide-protein[41] interactions are determined in solution by the formation of covalent bonds between a photoactive conjugate of the smaller moiety and the larger protein while the binding site is filled. The requirements of these photoactive chemistries are that they are activated by a wavelength which will not damage the cell or protein partner,[42] demonstrate an

active lifetime short enough that the complex will not dissociate before crosslinking occurs which would lead to false positive signals, must be minimally selective for the same reason, and must be orthogonal to all functional groups within the experiment. The review by Tanaka[40] provides considerable details regarding the current chemistry for these types of photoimmobilisation systems. This chemical technology on the solid phase can be used to immobilise small molecules for affinity chromatography as shown by Dilly *et al.*[6, 39]

A variety of other immobilisation methods are available,[36, 43] however conventional matrices for affinity chromatography are often chemically unstable and incompatible with various fixation procedures, and purification efficiency is often low.[37]

After immobilisation, investigation of the interactions between the immobilised ligand and the biomacromolecule can be carried out by affinity chromatography,[30, 35, 36] where large collections of proteins are screened against collections or arrays of small molecules.[38] Combinatorial libraries of compounds may be screened for individual compounds capable of manipulating the functions of a single target protein,[32] or a library of proteins may be screened for those proteins which have high affinity binding to an immobilised compound.[43]

The detected binding events in this type of assay usually occur on functionally important surfaces of the protein, which allows a prediction of the *in vivo* effects of these binding events.[32] Even after successful target identification, it must be determined if the action and binding states translate to solution state, and whether the binding will result in any visible effect on the cell state.

Drawbacks include non-specific binding, the specific binding of low-potency proteins,[30] a limited choice of support for the small molecules and often an instability of the matrix which can cause problems in purification.[37] Therefore the

success of any affinity binding technique will depend on the strength of binding between the small molecule and the target to allow for washing procedures which will remove any non-specific binding products, and also an abundant protein expression for the amplification of the protein target to appear above background noise.[30]

Affinity chromatography studies the intermolecular interactions (hydrophobic, electrostatic, hydrogen bonding, etc.) between the stationary phases and mobile phases, which are similar to the interactions between ligands and receptors in biological environments leading to potential drug discovery uses and analysis of biomarkers for clinical diagnosis.[31, 33, 35, 44]

Protein target identification and characterisation techniques after protein affinity assay are many and varied, from labelling with heavy and light isotopes[30] through mass spectrometry, NMR spectroscopy and surface plasmon resonance.[44] No one technique gives all of the information about the target protein and its interaction with the ligand. Studies can be based on one compound binding one target or receptor, up to library screening.[44-46]

There are other methods for target identification that avoid some of the complexity of whole proteome interaction studies and protein-target identification. For example; chemi-proteomics,[44, 45] studying a manageable fraction of the proteome; oligonucleotide libraries to assay functional DNA and RNA molecules;[44] mRNA display libraries where the expressed protein is covalently bound to its mRNA sequence[33]; and phage display library methods, where a cDNA insert is spliced into the gene for one of the bacteriophage coat proteins, so that the foreign amino acid sequence is genetically fused to the endogenous amino acids of the coat protein to

make a hybrid 'fusion' protein.[47] The phage display methods used in this thesis will be described in more detail in Chapter 2.1.

1.5 Magic Tag Photochemistry Method

In 2007 Dilly *et. al.* published a paper detailing the photoactivated immobilisation of small molecules onto derivatised 96-well plates and the screening of those photoadducts against a phage display library.[6] The method utilises a number of different photochemical groups, allowing for the immobilisation of a number of different regioisomers of the drug (see Figure 1.5.1), and a reduction in the non-specific binding of proteins.[48]

The photoactive functional groups include two different benzophenones, two different phenyl azides and one diazirine chemistry. The benzophenone chemistries form a diradical under UV irradiation, which can abstract a hydrogen from nearby molecules and either react with the radical species formed or go on to react further with other molecules.[49] Phenyl azides form nitrenes under UV irradiation, and the diazirine chemistry forms a carbene.[50] These highly active species are formed in the singlet configuration, the nitrene can form the triplet configuration through inter system crossing, or a less active ketenimine through ring expansion.[51] The singlet nitrene or carbene insert into C-H, or heteroatom-H bonds and the triplet species behave as a diradical, abstracting a hydrogen from neighbouring species followed by further reaction or radical termination.[52, 53] The ketenimine is mildly reactive to basic chemistries.[51] These covalent bonds are formed with limited specificity, giving multiple regioisomers.[54-56]

By displaying different regions of the drug molecule, the method minimises the areas of the drug molecule kept hidden from the assay - as they would be in the case of

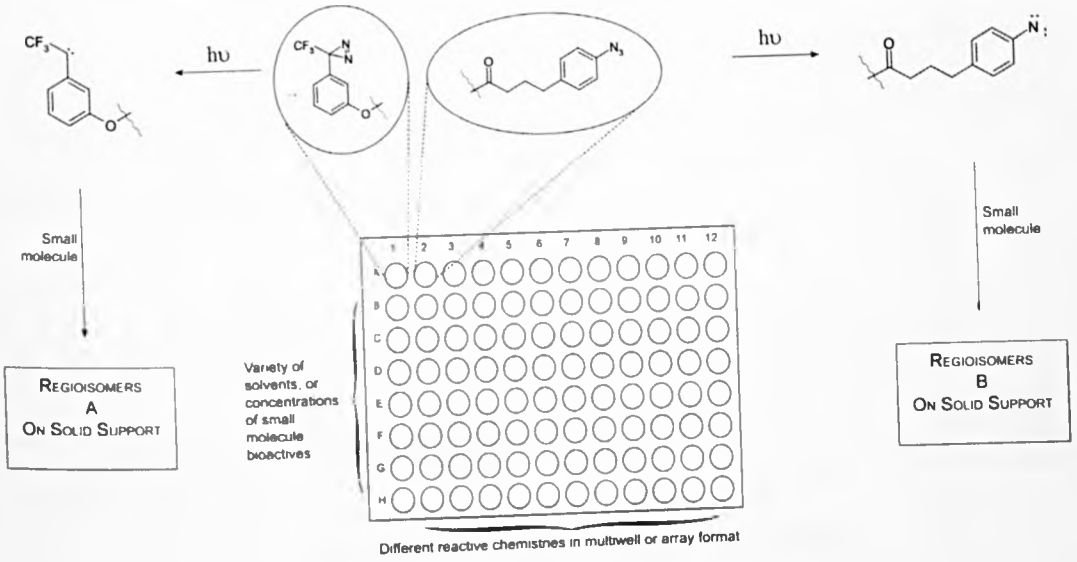


Figure 1.5.1: The Magic Tag® Plate Layout and UV activation[6]

specific immobilisation or solid phase synthesis, and can maximise the potential non-covalent interactions of the molecule with a protein target binding site. Using phage display technology as described below (Figure 1.5.2), high affinity interactions of drug and protein can be determined. This technology has since been trademarked under the name Magic Tag®. The photochemical immobilisation system used and later built upon - Magic Tag® - was developed as a means of immobilising small molecules without concealing any facet of the molecule from the interaction study. Five different photochemistries are displayed in a multi-well format, to maximise diversity in the display of the small molecule.

Because of the non-specific nature of the photochemical immobilisation of the small molecule, several regioisomers might be expected to form on the Magic Tag® surface. To be able to connect a protein interaction with a specific face of the small molecule, analysis of this interaction must be undertaken. For this purpose a cleavable resin analogue of the Magic Tag® surface was prepared.

To assay the immobilised small molecule on the Magic Tag® surface, a phage display library is prepared to represent the proteome of interest. Biopanning[46] (Figure 1.5.2) of the immobilised small molecule against this library allowed analysis of small molecule-polypeptide interactions.

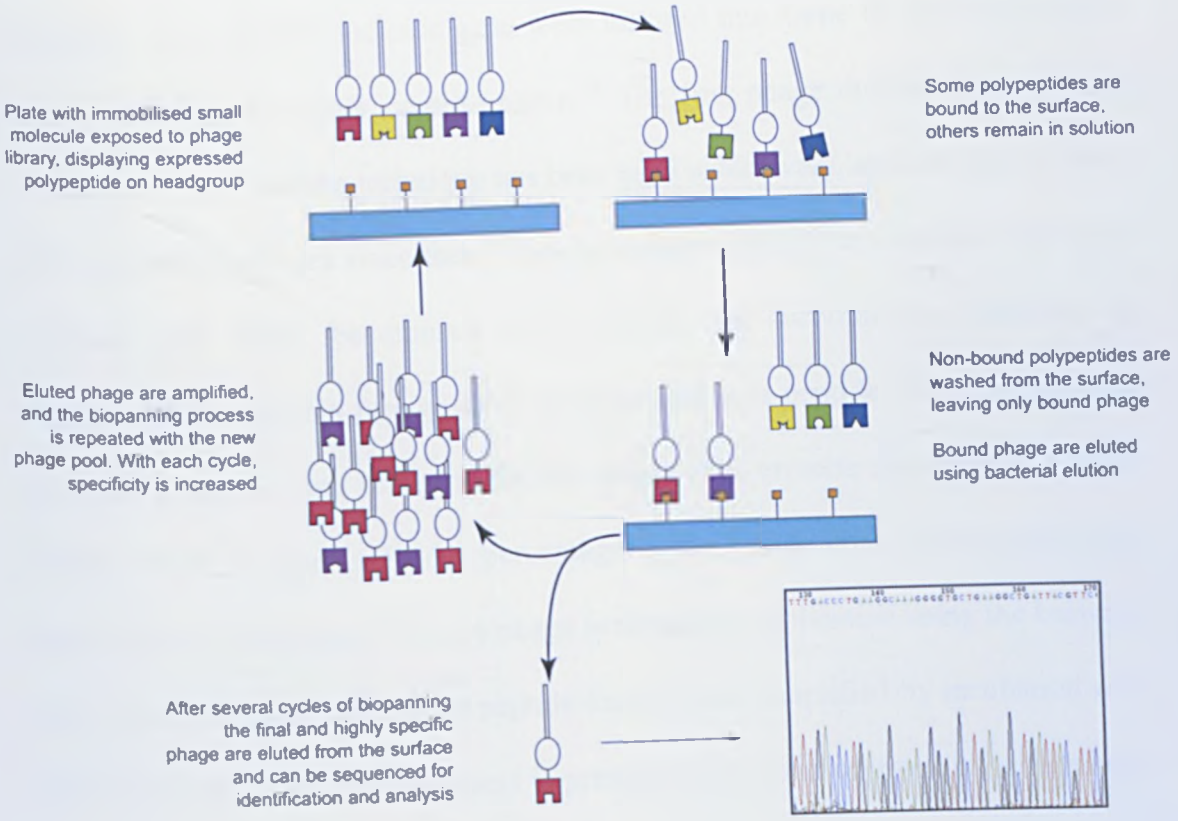


Figure 1.5.2: The biopanning system on an immobilised small molecule

In 2008, Ladwa *et. al.* published the results of a study of interactions between a human vascular lung phage library and Magic Tag® supported salbutamol[57] - a β2-adrenoreceptor agonist used in the treatment of asthma - interactions against a human vascular lung phage library. They determined that the phage clone showing the best binding to the immobilised salbutamol displayed a portion of the protein ATF4. Isoprenaline (an analogue of salbutamol) affects the expression of ATF4 in airway smooth muscle cells. Stressed cells, for example in asthma, were shown to arrest production of ATF4. Addition of isoprenaline restores expression (as monitored by

mRNA quantification). Isoprenaline had no effect on ATF4 expression under basal conditions.

1.6 Phage display

The first phage displayed proteins were prepared by Smith in 1985, when two fragments from an endonuclease gene were inserted into Gene III of a filamentous phage, encoding the minor capsid protein.[58] The first phage display libraries were prepared in 1991, and the technique has been used extensively, and has shown many different and novel uses since then.[59] Bacteriophages (phage) are viruses that infect bacterial cells. Their recombinant DNA vectors can accommodate segments of 'foreign' DNA, and this foreign DNA is expressed as a peptide. This foreign gene sequence is spliced into the gene for the phage coat protein, resulting in a fused peptide which is expressed on a percentage of the phage coat proteins when the phage particle is amplified.[47] Every phage is capable of replication using the bacterial replication machinery, and so the peptide insert can be amplified by incubation with an *E.coli* stock. As the peptide insert is presented into solution from the phage head group, these peptides retain most behaviour they would show in vivo, and so can be used to give some indication of biological peptide behaviour.[47] By PCR of the DNA insert, the coding cDNA sequence for these inserts can be identified by Sanger sequencing.

The ultimate aim of phage display is the selection of phage that bind to the target antigen of interest with high affinity from a large excess of phage clones which do not bind or do so with lower affinity. This is achieved by multiple rounds of binding the phage to the target, washing to remove non-bound phage followed by elution and amplification of specifically bound phage. Apart from the amount of ligand bait

present, during selection other parameters have to be taken into account depending on the nature of the ligand. Procedures for selecting against purified ligand rely upon the binding affinity of the phage for the ligand of interest relative to phage particles which do not bind and require the ligand to be attached to a solid support. Success of phage display selection depends significantly upon the quality of the original phage display library used, as polypeptide binding sites which are not represented cannot be isolated.[60]

There are two parameters for selection that can be manipulated to some extent. These are stringency - the degree to which peptides with higher fitness are favoured - and yield - the proportion of phage with a given fitness that survive selection. High yield is most important for the first round of selection whose input consists of a very broad library. In later rounds of selection, after the first amplification, yield can be decreased in favour of stringency.[47]

There are several different types of bacteriophage that can be used in this manner, but two of the most common are M13 and T7 phage. The M13 phage are filamentous phage which leave the bacterial cell though the secretory pathway. This has been shown to limit the peptide diversity of the library and shows positional amino acid biases.[61] Previous work done by the group had been carried out using T7 phage, which are lytic phage, leaving the host through host cell lysis. These displayed peptides do not have the same biological constraints as the M13 system and should show maximum potential diversity. It has been shown that the T7 systems show greater peptide diversity than the M13 systems.[61, 62] It is also useful to note that the T7 phage grow at a much faster rate than M13 phage, allowing three rounds of biopanning and amplification to be carried out over two days rather than over weeks.

Despite its successes, a common criticism of phage-displayed cDNA libraries is that the method cannot reproduce any posttranslational modifications found in many eukaryotic proteins and that proper protein folding may be inhibited by the lack of ER and appropriate chaperones. In addition, proteins may be expressed poorly due to differential codon usage between host and library.[63]

1.7 Cleavable chemistries

The Magic Tag® system is designed to maximise regiochemical diversity by using five different chemistries, and photoactivated immobilisation which is generally non-specific in its method of reaction. Some specificity has been identified in certain functional groups.[40, 64] Different arrangements of the small molecule on the surface will allow it to interact with different binding sites in the protein/phage library.

The low specificity of immobilisation means that it is difficult to assess which part of the molecule has interacted with the identified polypeptide target. If this target might result in a detrimental side effect it would be useful to be able to identify the part of the small molecule responsible for binding, so that the molecule might be modified to reduce the likelihood of this binding event occurring. To do this it would be useful to be able to identify the mode of binding of the small molecule on the surface.

The addition of cleavable linkers to the construction of the surface-supported moiety allows removal of the prepared molecule from the surface. This gives the opportunity for solution phase analysis after each reaction step. A substantial review of available cleavage methods from solid support has been reported by Guillier.[65] Cleavage also allows the estimation of loading as a percentage by weight of the resin.

1.8 Hypothesis, Aims and Objectives

By immobilising a small molecule ligand, generally a drug molecule, pro-drug or a molecule under clinical study, on Magic Tag® plates and carrying out three rounds of biopanning against a T7Select Phage display library prepared using human vascular tissue from patients suffering from ischaemic heart disease, we hope to observe the enrichment of some polypeptides out of the library. These observed interactions will suggest some leads for further bioinformatic and laboratory study, increasing the knowledge base for the network pharmacology of the ligand molecule.

An additional aim was to prepare a cleavable resin based photoactive chemistry based on the Magic Tag® immobilisation system. This cleavable system would allow the association of these binding events with regiochemistry observed after cleavage in the immobilised small molecule, leading to structure-activity relationships.

1.9 References

1. A. Atkinson, *Principles of clinical pharmacology*, Academic Press, 2007.
2. I. H. Raisfeld, *Annu. Rev. Med.*, 1973, **24**, 385-418.
3. P. Y. Lum, C. D. Armour, S. B. Stepaniants, G. Cavet, M. K. Wolf, J. S. Butler, J. C. Hinshaw, P. Garnier, G. D. Prestwich, A. Leonardson, P. Garrett-Engele, C. M. Rush, M. Bard, G. Schimmack, J. W. Phillips, C. J. Roberts and D. D. Shoemaker, *Cell*, 2004, **116**, 121-137.
4. J. Drews, *Science*, 2000, **287**, 1960-1964.
5. M. L. MacDonald, J. Lamerdin, S. Owens, B. H. Keon, G. K. Bilter, Z. D. Shang, Z. P. Huang, H. Yu, J. Dias, T. Minami, S. W. Michnick and J. K. Westwick, *Nat. Chem. Biol.*, 2006, **2**, 329-337.
6. S. J. Dilly, M. J. Bell, A. J. Clark, A. Marsh, R. M. Napier, M. J. Sergeant, A. J. Thompson and P. C. Taylor, *Chem. Commun.*, 2007, 2808-2810.
7. F. Sams-Dodd, *Drug Discov. Today*, 2005, **10**, 139-147.
8. J. Drews, *Drug Discov. Today*, 2000, **5**, 2-4.
9. A. L. Hopkins, *Nat. Chem. Biol.*, 2008, **4**, 682-690.
10. S. C. Janga and A. Tzakos, *Mol. BioSyst.*, 2009, **5**, 1536-1548.
11. M. Cases and J. Mestres, *Drug Discov. Today*, 2009, **14**, 479-485.
12. I. Vogt and J. Mestres, *Mol. Inf.*, 2010, **29**, 10-14.
13. J. Watkins, A. Marsh, P. C. Taylor and D. R. Singer, *Ther. Deliv.*, 2010, **1**, 651-665.
14. K. Strong, C. Mathers, S. Leeder and R. Beaglehole, *Lancet*, 2005, **366**, 1578.
15. C. L. Williams, L. L. Hayman, S. R. Daniels, T. N. Robinson, J. Steinberger, S. Paridon and T. Bazzarre, *Circulation*, 2002, **106**, 143-160.

16. A. Lerman and A. M. Zeiher, *Circulation*, 2005, **111**, 363-368.

17. L. S. Lilly, ed., *Pathophysiology of Heart Disease: A Collaborative Project of Medical Students and Faculty*, Lippincott Williams & Wilkins, 2007.

18. R. H. Knopp, *N. Engl. J. Med.*, 1999, **341**, 498-511.

19. A. Gaw, C. Packard and J. Shepherd, *Statins: the HMG CoA reductase inhibitors in perspective*, Martin Dunitz, 2000.

20. M. Igel, T. Sudhop and K. vonBergmann, *Eur. J. Clin. Pharmacol.*, 2001, **57**, 357-364.

21. R. P. Mason, M. F. Walter, C. A. Day and R. F. Jacob, *Am. J. Cardiol.*, 2005, **96**, 11F-23F.

22. M. Lechleitner, *Journal of Clinical and Basic Cardiology*, 2002, **5**, 205-208.

23. M. Jasiska, J. Owczarek and D. Orszulak-Michalak, *Pharmacol. Rep.*, 2007, **59**, 483-499.

24. P. Libby, *Nature*, 2002, **420**, 868-874.

25. H. McGavock, *How drugs work: basic pharmacology for healthcare professionals*, Radcliffe Publishing, 2005.

26. P. Winstanley and T. Walley, *Medical pharmacology: a clinical core text for integrated curricula with self-assessment*, Churchill Livingstone, 2002.

27. H. E. Vonkeman and M. A. F. J. van de Laar, *Semin. Arthritis Rheum.*, 2010, **39**, 294-312.

28. J. M. Ritter, I. Harding and J. B. Warren, *Trends Pharmacol. Sci.*, 2009, **30**, 503-508.

29. L. Sleno and A. Emili, *Curr. Opin. Chem. Biol.*, 2008, **12**, 46-54.

30. L. Burdine and T. Kodadek, *Chem. Biol.*, 2004, **11**, 593-597.

31. N. Shimizu, K. Sugimoto, J. W. Tang, T. Nishi, I. Sato, M. Hiramoto, S. Aizawa, M. Hatakeyama, R. Ohba, H. Hatori, T. Yoshikawa, F. Suzuki, A. Oomori, H. Tanaka, H. Kawaguchi, H. Watanabe and H. Handa, *Nat. Biotechnol.*, 2000, **18**, 877-881.

32. T. Kodadek and K. Bachhawat-Sikder, *Mol. BioSyst.*, 2006, **2**, 25-35.

33. M. McPherson, Y. F. Yang, P. W. Hammond and B. L. Kreider, *Chem. Biol.*, 2002, **9**, 691-698.

34. C. R. Lowe, *Int. J. Biochem.*, 1977, **8**, 177-181.

35. T. Mori, T. Takahashi, T. Shiyama, A. Tanaka, N. Hira, N. Tanaka and K. Hosoya, *Bioorg. Med. Chem.*, 2006, **14**, 5549-5554.

36. L. H. Chen, Y. S. Choi, J. Kwon, R. S. Wang, T. Lee, S. H. Ryu and J. W. Park, *Tetrahedron*, 2004, **60**, 7293-7299.

37. H. Uga, C. Kuramori, A. Ohta, Y. Tsuboi, H. Tanaka, M. Hatakeyama, Y. Yamaguchi, T. Takahashi, M. Kizaki and H. Handa, *Mol. Pharmacol.*, 2006, **70**, 1832-1839.

38. P. J. Hergenrother, K. M. Depew and S. L. Schreiber, *J. Am. Chem. Soc.*, 2000, **122**, 7849-7850.

39. N. Kanoh, K. Honda, S. Simizu, M. Muroi and H. Osada, *Angew. Chem.-Int. Edit.*, 2005, **44**, 3559-3562.

40. Y. Tanaka, M. R. Bond and J. J. Kohler, *Mol. BioSyst.*, 2008, **4**, 473-480.

41. S. Kuboe, M. Yoda, A. Ogata, Y. Kitade, Y. Tomari and Y. Ueno, *Chem. Commun.*, 2010, **46**, 7367-7369.

42. F. Kotzybahibert, I. Kapfer and M. Goeldner, *Angew. Chem.-Int. Edit. Engl.*, 1995, **34**, 1296-1312.

43. Y. Ohtsu, R. Ohba, Y. Imamura, M. Kobayashi, H. Hatori, T. Zenkoh, M. Hatakeyama, T. Manabe, M. Hino, Y. Yamaguchi, K. Kataoka, H. Kawaguchi, H. Watanabe and H. Handa, *Anal. Biochem.*, 2005, **338**, 245-252.

44. R. J. Tian, S. Y. Xu, X. Y. Lei, W. H. Jin, M. L. Ye and H. F. Zou, *Trends Anal. Chem.*, 2005, **24**, 810-825.

45. U. Rix and G. Superti-Furga, *Nat. Chem. Biol.*, 2009, **5**, 616-624.

46. Y. Takakusagi, K. Takakusagi, F. Sugawara and K. Sakaguchi, *Expert. Opin. Drug Discov.*, 2010, **5**, 361-389.

47. G. P. Smith and V. A. Petrenko, *Chem. Rev.*, 1997, **97**, 391-410.

48. S. J. Dilly, M. P. Beecham, S. P. Brown, J. M. Griffin, A. J. Clark, C. D. Griffin, J. Marshall, R. M. Napier, P. C. Taylor and A. Marsh, *Langmuir*, 2006, **22**, 8144-8150.

49. G. D. Prestwich, G. Dorman, J. T. Elliott, D. M. Marecak and A. Chaudhary, *Photochem. Photobiol.*, 1997, **65**, 222-234.

50. M. S. Platz, *Acc. Chem. Res.*, 1995, **28**, 487-492.

51. C. Moody and G. Whitham, *Reactive intermediates*, Oxford University Press, 1992.

52. K. A. Schnapp, R. Poe, E. Leyva, N. Soundararajan and M. S. Platz, *Bioconjug. Chem.*, 1993, **4**, 172-177.

53. W. T. Borden, N. P. Gritsan, C. M. Hadad, W. L. Karney, C. R. Kemnitz and M. S. Platz, *Acc. Chem. Res.*, 2000, **33**, 765-771.

54. P. J. A. Weber and A. G. BeckSickinger, *J. Pept. Res.*, 1997, **49**, 375-383.

55. N. Kanoh, T. Nakamura, K. Honda, H. Yamakoshi, Y. Iwabuchi and H. Osada, *Tetrahedron*, 2008, **64**, 5692-5698.

56. S. Peddibhotla, Y. Dang, J. O. Liu and D. Romo, *J. Am. Chem. Soc.*, 2007, **129**, 12222-12231.

57. S. R. Ladwa, S. J. Dilly, A. J. Clark, A. Marsh and P. C. Taylor, *ChemMedChem*, 2008, **3**, 742-744.

58. G. P. Smith, *Science*, 1985, **228**, 1315-1317.

59. M. B. Zwick, J. Q. Shen and J. Scott, *Curr. Opin. Biotechnol.*, 1998, **9**, 427-436.

60. J. Smith, R. E. Kontermann, J. Embleton and S. Kumar, *Faseb J.*, 2005, **19**, 331-341.

61. L. R. H. Krumpe, A. J. Atkinson, G. W. Smythers, A. Kandel, K. M. Schumacher, J. B. McMahon, L. Makowski and T. Mori, *Proteomics*, 2006, **6**, 4210-4222.

62. S. C. Sharma, A. M. Memic, C. N. Rupasinghe, A. C. E. Duc and M. R. Spaller, *Biopolymers*, 2009, **92**, 183-193.

63. A. M. Piggott, A. M. Kriegel, R. D. Willows and P. Karuso, *Bioorg. Med. Chem.*, 2009, **17**, 6841-6850.

64. M. Wiegand and T. K. Lindhorst, *Eur. J. Org. Chem.*, 2006, 4841-4851.

65. F. Guillier, D. Orain and M. Bradley, *Chem. Rev.*, 2000, **100**, 3859-3859.

Chapter 2 Phage Library Formation

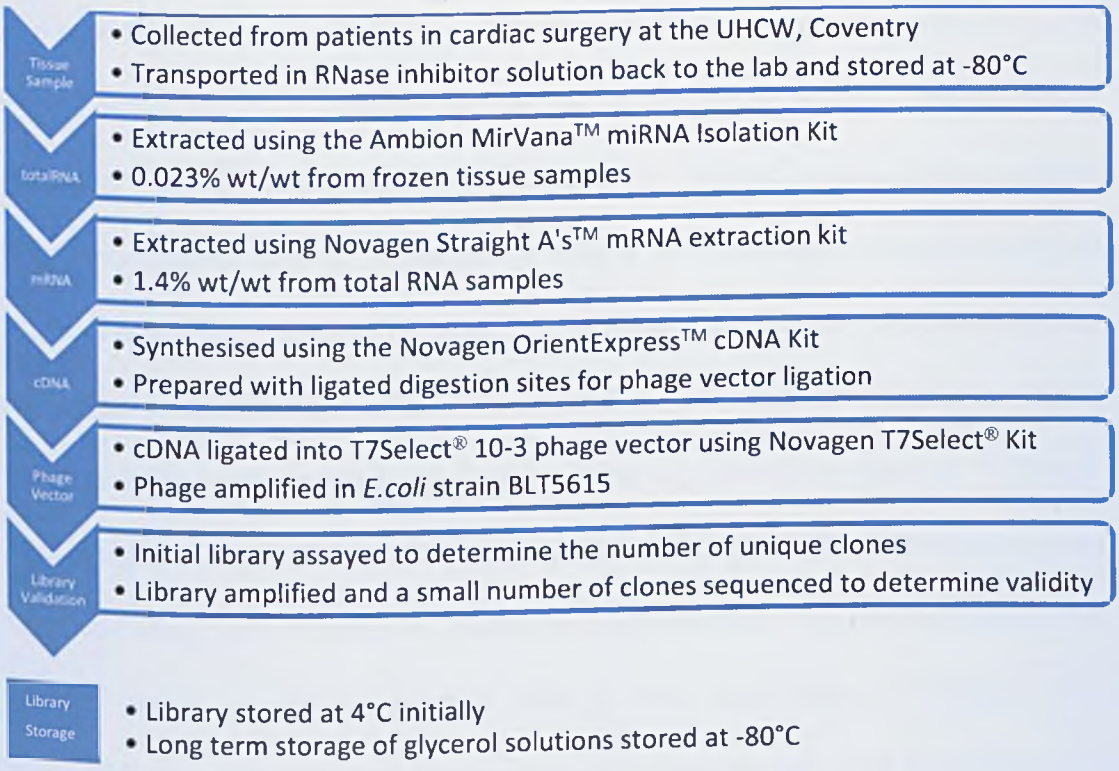
2.1 Introduction

2.1.1 Aims

Identifying a protein target directly from a cell lysate can be difficult,[1] with low quantities and low solubility of the binding protein, and the size of many receptor molecules demands an often unattainable sensitivity of the spectroscopic techniques if the ligand-receptor complex is to be observed directly.[2] Phage display techniques allow the binding amino acid sequence to be linked to its coding sequence.[3, 4] The ultimate aim of phage display is the selection of phage presenting polypeptides which bind to the small molecule of interest with high affinity from a large excess of phage clones which do not bind or do so with lower affinity.[5] A single phage particle from an output pool can be selected, the insert DNA representing the polypeptide amplified and nucleotide sequenced, meaning that even single binding events, provided they can be isolated from background noise, can be identified.[6]

By creating a phage library from human vascular tissues we hope to obtain a sample proteomic description of the vascular cell, albeit with no posttranslational modifications or structured protein folding.

2.1.2 Scheme for preparation of phage display library



2.1.3 T7Select® 10-3

There are currently several different types of bacteriophage used in phage display, the respective qualities of which are described in the introduction. The use of T7 bacteriophage lytic phage allows fast amplification (2-4 hours), limited biological constraints introduced by secretion pathways and less specific and positional amino acid bias than has been seen with the M13 bacteriophage.[7]

Preparation of a T7 phage display library from purified mRNA is well known, and can be carried out following commercial kit procedures (see details in Methods and Materials, Chapter 2.4). With additional use of commercial kits for extraction of RNA from tissues, formation of a cDNA library and packaging of library into phage vector arms can be carried out using straight forward and prevalidated techniques.

The T7 bacteriophage library kit used in this project was the T7Select® 10-3 Phage Display System from Novagen (T7Select® 10-3 OrientExpress™ cDNA Cloning

System, Oligo(dT) Cat# 70581-3, T7Select® 10-3 Cloning Kit Cat# 70550-3, Novagen User Protocols TB247 and TB178). This kit was used in combination with *E. coli* strain BLT5615 selected for the reasons described below (Glycerol Stock, Novagen, Cat# 69905). There are three types of T7Select® phage display vectors: high copy number display of peptides (T7Select®415), mid-copy number display of peptides and larger proteins (T7Select®10) and low copy number display of peptides and larger proteins (T7Select®1).[8] The inserted genetic material is cloned on the C-terminus of the phage capsid 10B protein which is produced by gene 10. T7Select® 10-3 phage have a modified s10 translation initiation signal (ribosome binding site) to reduce capsid gene expression, leading to a copy number of expressed inserts at 5-15 per phage particle on average. Due to this modification T7Select® 10-3 bacteriophage require an exogenous source of capsid protein, and BLT5615 is an ampicillin resistant *E. coli* which provides large amounts of capsid 10A protein.[9] In *E. coli* strain BLT5615, the expression of gene 10A, producing the major T7 capsid protein, is controlled by a LacUV5 promoter which can be induced by isopropyl β-D-1-thiogalactopyranoside, IPTG, a molecular mimic of allolactose, that triggers transcription of the *lac* operon.[8, 9] This allows amplification of the bacterial stock without the retardation of growth caused by capsid protein overproduction.

2.1.4 Source Human Tissue

In order to take a representative sample of the proteome of a patient suffering from cardiovascular disease, the focus of study in this thesis, the tissue samples collected for preparation of the library were from patients undergoing coronary artery bypass graft surgery for ischemic heart disease. The Human Tissue Act 2004 regulates the collection of consent from patients for use of human tissues for research, as well as

all storage, use and disposal of human tissue for the purposes of research. The CSRI, University of Warwick is a licensed centre with the Human Tissue Authority and all work with human tissues was carried out under their license.

Consent of the patient for collecting the surplus tissue is not collected by any researcher likely to be at a later date working with the material or any extracted genetic information, and the researcher is not in possession, and not likely to come into possession of information that identifies the person from whom it had come.[10, 11] Stored samples must be labelled with a code which allows connection of that patient back to the collected consent. Current recommendations are that consent is gathered for non-specific, non time-restricted research, allowing for use of the material in further studies without collecting new consent.[10, 11] No human tissue sample was stored for more than a week without consent from the patient, and no human tissue sample was used or processed to total RNA without consent from the patient.[12]

By the nature of the vascular tissue starting material, extracted from individuals suffering ischaemic heart disease, the prepared polypeptide phage library would be expected to be representative of the vascular cell proteome of the disease state individual. Modifications in protein expression as a response to the vascular and resulting heart disease, or those which have led to the vascular disease or are a response to the underlying cardiovascular risk factors, all might be expected to affect the quantities of individual proteins present in the individual's vascular cells.

By preparing the human vascular tissue phage display library in this way, the interactions observed by biopanning against an immobilised small molecule can be considered to be more biologically relevant than the use of cultured cells, for example human vascular endothelial cell cultures, HUVEC, which would be healthy cells.

The ability to track the clinical history of each sample, including drug use, cardiovascular risk factors and other disease influences will also be very important in connecting any potential drug-protein hits back to individual interactions. This continued link between individual samples, individual libraries and the protein interactions allows more reliable assertions to be made as to the nature of the interactions observed from the biopanning and their relevance *in vivo*.

2.1.5 Personalisation of Phage Display Library Preparation

If suitably sensitive, it should be possible to prepare a phage display library from tissues originating from an individual or small cohort. A personalised polypeptide phage display library would allow drug/probe/tissue target interaction studies for rare diseases and uncommon side effects, and could also allow for development of improved personalised treatment, with increased efficacy and reduced toxicity.

2.1.6 mRNA extraction from human tissues

All cells contain endogenous RNases which break down unprotected RNA in the cell. These enzymes break down unneeded RNA and protect against RNA viruses,[13] but they are also a major source of degradation of RNA in the extraction process. To avoid loss or damage to RNA between harvesting of tissues and the lysis of the cells for RNA extraction, the tissues are transported in an RNase inhibitor solution on ice, and then stored at -80°C, effectively halting any enzymatic action. As far as possible, time between removal of the cells from -80°C storage and processing is minimised. Disruption of cells is carried out in RNase inhibitors to immediately inactivate the RNase released from the cell.[14]

The extraction of mRNA can be carried out in two steps via a total RNA extraction.[4, 15-17] The Nanodrop spectrophotometer (ThermoScientific) is used to assess the concentration and purity of extracted nucleotide solutions by spectroscopic analysis. Nanodrop spectroscopy determines the absorbance of the nucleotide solution at three different wavelengths over a fixed path length. RNA and DNA have characteristic absorbance curves, as do proteins and other contaminants such as phenol and alcohols.[18] The absorbance maximum for nucleotides is at 260 nm, and the concentration of the solution can be measured at this wavelength using the Beer-Lambert Law. These curves can be affected by contamination, which is determined by the ratio of absorbance at two further wavelengths.

Protein has an absorbance maximum at 280 nm, and protein, EDTA, guanidine HCl, salt and phenol all show strong absorbance at 230 nm.[18] For purified RNA, $A_{260/280}$ should be 2.0. Ratios lower than this ideal suggest contamination of the RNA with protein.[18] Pure nucleic acids produce a $A_{260/280}$ ratio of 2.3 - 2.5,[18] though the Nanodrop technical specifications suggest a ratio of 2.0 - 2.2 are to be expected.

After purification of total RNA, further extraction of mRNA from this material is required. Almost all mRNA is polyadenylated at the 3' end, poly(A)$^+$ mRNA,[19] the polyA tail being essential for nuclear transport, stability and translation of mRNA.[20] This polyA tail gives an easy method for the capture of this material from mixed nucleotide solutions using oligo(dT) resins or capture matrices. After capture the mRNA can be washed and then eluted from the resin. Extracted mRNA can be assayed for concentration and purity by Nanodrop as described above.

2.1.7 cDNA library formation

There are two methods for priming the cDNA first strand synthesis. Either oligodeoxythymidylic acid, commonly known as oligo(dT) - primers or random (hexamer) primers with two thymine residues can be used to prime the reverse transcription reaction, carried out by the enzyme reverse transcriptase, in the presence of three standard deoxynucleotide triphosphates (dNTPs) and a methylated deoxycytosine (5-Me dCTP).[4, 21] These dNTPs are incorporated into the new DNA strand, which complements the original mRNA. The methylation on the cytosine residue protects the cDNA product from premature restriction or cleavage. The reverse transcriptase used is isolated from retroviruses such as Maloney Murine Leukaemia Virus (MMLV RT), and can use RNA as a template to synthesise a DNA strand in the 5' to 3' direction.[21]

The choice between oligo(dT) priming and random priming is based on several factors. Oligo(dT) primers are designed to prime the 3' polyA region of the mRNA present. This means that they are very specific for poly(A)$^+$ mRNA, but because the reverse transcriptase must start at the 3' end of the molecule, it is possible it will not reach the 5' end, which would be a particular problem for very long mRNAs.[22] It is possible for the primer to bind at the most 3' end of the polyA tail, requiring transcription of the polyA tail, which can be as long as 200-250 base pairs, and reducing the length of relevant transcribed material further.[22] There is also the loss of any mRNA without a polyA tail to take into account.[22] Random primers can bind throughout the mRNA molecule, which would lead to better coverage of the whole molecule. However, the risk with random primers is that they will prime any RNA present in the sample, and so starting material mRNA must be very pure. Random

primers also offer the ability to modify average insert size by adjusting the ratio of primers to mRNA.[23]

New dNTPs, including a methylated cytosine residue are incorporated into the new DNA strand, which complements the original mRNA. No purification is carried out before the second strand synthesis.

This first product is a RNA-DNA hybrid, which is treated with RNase H (an *E. coli* enzyme) during second strand synthesis to digest the RNA strand. These RNA pieces remain hybridised to the first cDNA strand in order to prime synthesis of the second DNA strand. Second strand synthesis is carried out using the enzyme DNA polymerase, which replaces all RNA with DNA except for a small section at the 5' end. This cDNA strand is not contiguous, as nicks at the point where the reaction was primed remain, and are ligated along with the restriction site linkers in the next step.

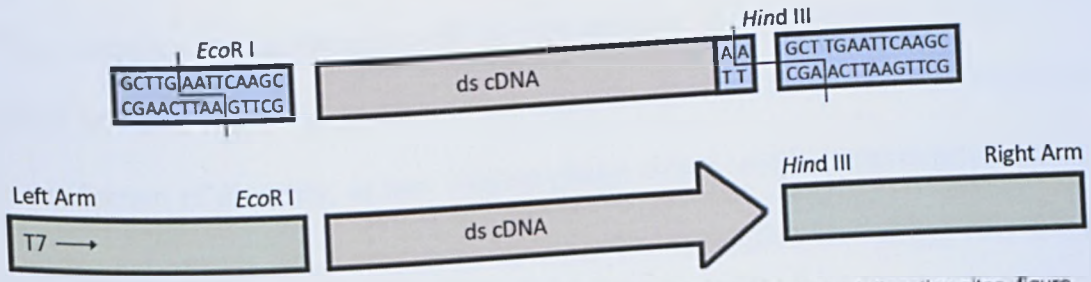


Figure 2.1.7: Insert of prepared cDNA into T7Select® vector arms with formed prepared digestion sites, figure reproduced from Novagen OrientExpress™ cDNA Manual

As there will have been many mRNA molecules extracted from the original tissue sample, and oligo(dT) priming would have occurred for each of these mRNA molecules, the resultant cDNA library will be large and highly diverse in identity and DNA length.

2.1.8 Phage Display Library Preparation and Validation

The phage library is prepared by ligation of the cDNA library into T7Select® 10-3 vector arms.[8] Validation of the library requires determination of the diversity of the library and the initial copy number of the library after amplification, followed by sequencing of a small number of clones.

Amplification of the phage is carried out in the host *E. coli* strain BLT5615, which upon induction with IPTG produces the major T7 capsid protein, of which the T7Select® 10-3 vector is deficient. The *E. coli* stock solution is allowed to grow to an optical density of 0.5 before introduction of 1% 100mM IPTG solution, which ensures the production of the capsid protein does not inhibit growth of the *E. coli*. The amplification is carried out in a flask large enough to ensure that oxygen deprivation does not limit the *E. coli* growth, and agitation maximises amount of dissolved oxygen in the media.[24]

Phage aliquots are introduced to the *E. coli* solution at a multiplicity of infection (MOI) of 0.001 - 0.01.[8] If the MOI is too high, competition for *E. coli* cells will lead to a reduction of diversity, as less virulent phage clones will be outcompeted during amplification. The number of plaques increases linearly with added phage only when the MOI is much less than one.[3]

PCR can be carried out on phage samples, either neat or as 'picks' from a phage plaque on an *E. coli* lawn. PCR on these sequences is carried out using two commercial primers specific to the T7Select® sequence, identified as 'UP' and 'DOWN' - see below.[8] This amplifies the sequence up to and covering the site of insertion (between the left and right arms, highlighted in the sequence below), amplifying the sequence which has been inserted, or amplifying the vector sequence in the case of failed insertion.

Figure 2.1.8: T7Select®10-3b Vector Sequence

```
20301  TGCTAACTTC CAAGCGGACC AGATTATCGC TAAGTACGCA ATGGGCCACG
                              T7SelectUP primer # 70005-3
20351  GTGGTCTTCG CCCAGAAGCT GCAGGAGCTG TCGTATTCCA GTCAGGTGTG
                    LEFT ARM (EcoR I site)
20401  ATGCTCGGGG ATCCGAATTC TCCTGCAGGG ATATCCCGGG AGCTCGTCGA
                RIGHT ARM (HinD III site)
20451  CAAGCTTGCG GCCGCACTCG AGTAACTAGT TAACCCCTTG GGGCCTCTAA
              T7SelectDOWN primer # 70006-3
20501  ACGGGTCTTG AGGGGTTTTT TGCTGAAAGG AGGAACTATA TGCGCTCATA
20551  CGATATGAAC GTTGAGACTG CCGCTGAGTT ATCAGCTGTG AACGACATTC
```


Table 2.1.8: Sequences for T7Select® UP and DOWN primers	
T7Select® UP primer # 70005-3	T7Select® DOWN primer # 70006-3
5' - GGAGCTGTCGTATTCCAGTC - 3'	5' - AACCCCTCAAGACCCGTTTA - 3'

Sequencing of T7Select® phage PCR products is specifically primed to amplify from one side of the insert site - using only the 'UP' T7Select® primer. Sequencing allows quantification of vectors in which there is no insert, and allows validation of those inserts which have been successfully cloned.

Chain terminator sequencing, or Sanger sequencing,[25, 26] uses dideoxynucleotide triphosphates (ddNTPs) as DNA chain terminators. In four separate reactions, the single stranded DNA template is amplified using a preexisiting DNA primer in a reaction with DNA polymerase and all four standard dNTPs and one of the four ddNTPs, which chain terminate the reaction and result in DNA fragments of different lengths. These fragments can then be separated by denaturing polyacrylamide-urea gel electrophoresis in four separate lanes, allowing

determination of the position of chain termination - indicating the presence of the base. By labelling the ddNTPs fluorescently,[27] radioactively,[25] or using dyes,[28] it is possible to visualise the result and either visually or computationally extract the sequence data.

Lasergene® SeqMan is DNASTAR software for sequence assembly, allowing the alignment of a large number of sequences into a defined number of contiguous regions. This allows any sequences that appear more than once to be analysed in one group. It also allows lower quality sequences to be compared to better quality sequence, improving the sequence information. The aligning regions give a consensus sequence that can then be analysed using tools such as BLAST. Lasergene® SeqMan was used to analyse the sequences output by the sequencing method.

2.1.9 Regulation on Storage and use of GMOs

The HSE regulate the storage and use of GMOs under 'contained use' within the laboratory environment.[29]

Genetically modified organisms (GMOs), specifically the genetically modified microorganisms (GMMs) used throughout this project are regulated for contained use by the HSE. Genetic modification includes any case where the DNA has been removed from the cell, manipulated and reinserted into the same or a different organism.

For this project, DNA has been removed from human tissues, inserted into a bacteriophage vector and packaged into the phage units for amplification in *E. coli*.

The *E. coli* purchased from Novagen is also a genetically modified microorganism.

Contained use covers any situation where control measures are used to limit contact between the GMOs and humans and the environment so as to provide a high level of safety.

By exercising control measures at source and supplementing these with appropriate personal protective clothing and equipment when necessary, providing appropriate training, displaying biohazard signs, providing written standard operating protocols (SOPs), providing safe storage for materials and contaminated laboratory equipment and having effective disinfectants and specified disinfectant procedures, the hazards associated with GMM use can be protected against. Disinfectant procedures are essential to inactivate the GMMs through complete or partial destruction, to ensure any contact between the GMM and humans or the environment is limited.[29]

With the use of phage in the GMO lab at Cherwell Innovation Centre, disinfectant procedures required first Virkon treatment of all equipment and bench surfaces, followed by 70% isopropanol treatment of equipment and surfaces. Infected solutions were treated first with Virkon solution overnight, followed by autoclaving of the remainder before disposal. The area was restricted to authorised personnel, and PPE worn inside the lab was not used outside the GMO lab. Lab equipment was also restricted to the GMO lab, and not used interchangeably between labs.[29]

The labs in the Horticultural Research Institute, Wellesbourne, and Cherwell Innovation Centre's GMO lab are registered as Class 2 GMO handling labs.

2.2 Results and Discussion

2.2.1 Tissue Sample Collection

Informed consent for enduring generic research use of surplus tissues was collected from patients for tissue harvesting in collaboration with Mr. Wade Dimitri and his

team at University Hospital, Walsgrave, Coventry. The Lead Investigator for this project was Professor Donald Singer. This project was approved by the Coventry NHS Research Ethics Committee (REC reference number: 07/H1210/126, see appendix 8.1 for approval details).

After consent was taken the patient was given a number based on a coding system, which could at a later date be linked back to clinical details relevant to the extracted tissue. Once consent was acquired, this was confirmed with the researcher and the tissue sample was collected. Tissue samples were linked to their consent by the date and time of the surgical procedure. In future it may be more efficient to report the code assigned to that patient back to the individual collecting the tissue, making this linking process more streamlined.

Surplus tissue was collected during coronary artery bypass surgery or aortic valve replacement surgery where additional bypass surgery was also required.

The material collected was primarily internal mammary artery with some surrounding tissue, material which would have otherwise have been discarded to clinical waste during the surgical procedure. During some procedures it was also possible to collect a sample of saphenous vein, or the removed non-functioning aortic valve.

The surplus tissue was received in theatre with surgical staples attached to all branching vasculature, and often with fatty deposits attached. These were difficult to remove after freezing and would have caused issues with tissue homogenisation later in the preparation of the sample, so were removed before leaving theatre, and disposed of in the theatre clinical waste.

Initial samples were placed immediately into liquid nitrogen for quick-freeze, but transporting liquid nitrogen to and from the theatre in University Hospital proved

unfeasible. The sample was subsequently transported in an RNase inhibitor solution (RNAlater® Solution, Ambion, Cat# AM7020) in ice. This solution allowed the sample to be transported on ice, and also allowed thawing of the sample after storage at -80°C for weighing before homogenisation for RNA extraction. The solution was entirely removed before storage of the sample at -80°C.

Table 2.2.1: Average sample sizes and tissue sizes after collection

Tissue Type	Average Sample Size (mg)	Average Tissue Sample (mg)
Mammary Artery	278.6	957.3
Saphenous Vein	119.8	164.7
Aortic Valve	N/A	N/A

A log of human tissue samples conforming to HTA guidelines was kept, describing sample's storage location, sample's unique identifier (this code is different to the consent code assigned to the patient), sample type, date and time of sample collection, amount of sample (mass), names of Lead Investigator and researcher, method of transportation from theatre and the dates on which the sample was processed to an acellular state. A second sample log was kept describing the dates on which the sample was processed to a non-cell state and which methods was used, including what quantities of this product nucleotide solution was used and when, and what remaining solution was stored. See appendices 8.2 for these sample logs.

On one occasion a large proportion of the internal mammary artery was considered unfit for use in the surgical procedure and would have been discarded to clinical waste, and so was harvested for the project. This one sample tends to skew the stated average tissue size for internal mammary artery samples, which is otherwise 880 mg.

2.2.2 mRNA extraction

mRNA extraction was carried out in two parts, firstly with extraction of total RNA (including small RNAs) using Ambion's mirVana® miRNA Isolation Kit, (Cat# AM1560), followed by extraction of mRNA using Novagen's Straight A's™ mRNA Isolation System, (Cat# 69962-3, User Protocol TB103).

As the sample had been transported in RNAlater® solution, it was possible to remove the tissue sample from -80°C storage and weigh at room temperature before homogenisation[14] with minimal degradation of the RNA by endogenous RNases. Homogenisation was carried out in the presence of lysis buffer, which contained RNase inhibitors and guanidinium thiocyanate to lyse the cell and denature RNase and DNase enzymes. This homogenisation process was carried out as quickly as possible to further minimise this degradation.

Table 2.2.2a: Analysis into collected human tissue samples

Tissue Type	Average Sample Size (mg)	Average Tissue Sample (mg)	No. of Samples per Individual	Total Number of Samples	Total Number of Individuals
Mammary Artery	279	957	3.5	177	56
Saphenous Vein	120	165	1.4	10	8
Aortic Valve	N/A	N/A	2.3	14	6

The total RNA was extracted using a phenol:chloroform separation method, with the additive bromochloropropane which aids in the separation of the layers, and the isolated nucleotide was then filtered and washed to remove all salts and other remaining contaminants. After extraction the total RNA was dissolved in RNase-free water, and the quantity and quality of the total RNA was assayed by Nanodrop. See Appendix 8.2 for full details of individual samples.

Table 2.2.2b: Analysis into the extraction of total RNA from human tissue samples					
Tissue Type	Av. % total RNA extracted (wt/wt)	Average 260/280	Average 260/230	Ratio Quality Assessment	Total Number of Samples extracted
Mammary Artery	0.023	1.9 no fail	1.9 incl. fail 1.9 excl. fail	93:16	110
Saphenous Vein	0.023	1.9 incl. fail 1.9 excl. fail	1.7 incl. fail 1.8 excl. fail	5:6	11

Based on expected A_{260}/A_{280} and A_{260}/A_{230} ratios of for RNA, 2.0 and 2.0-2.2 respectively, any samples with absorbance ratios lower than 1.7 in either ratio were considered of inadequate quality for further use and excluded from any further work. This failure rate was 16% for the internal mammary artery samples, and 45% for the saphenous vein samples. After this, the average ratios were A_{260}/A_{280} 1.92 and A_{260}/A_{230} 1.95 for internal mammary artery samples and A_{260}/A_{280} 1.91 and A_{260}/A_{230} 1.84 for the saphenous vein samples. This loss of quality in the saphenous vein samples is likely to be due to the relative size of the samples, as pipetting to remove the aqueous layer from the separation step must be more accurate to remove a smaller layer, and there is a tendency to use a larger excess of alcohol to precipitate the RNA in order to maximise the yield, leading to protein and alcohol contamination in the final product.

The purified total RNA was suspended in RNase-free water and stored at -80°C. An average of 58 µg ± 35 µg total RNA was extracted from each internal mammary artery sample of average 273 mg ± 136 mg and an average of 29 µg ± 14 µg extracted from each saphenous vein sample of average mass 120 mg ± 56 mg, using the Ambion mirVana™ miRNA Isolation Kit.

The aortic valve samples were not processed for extraction of total RNA.

The extraction of mRNA was carried out using Novagen's Straight A's™ method, using "Magnetight" oligo(dT) magnetic particles. These particles allowed capture of the particle-complexed poly(A)$^+$ RNA using a magnetic stand, and these beads could then be washed to remove any non-complexed material before the poly(A)$^+$ RNA was eluted from the surface using RNase-free water at 60°C. The quantity of magnetic particles used was calculated based on the input quantities of total RNA, with the calculation in the kit instructions. The final mRNA sample is suspended in 50 µl RNase-free water and stored at -80°C.

Table 2.2.2c: Analysis into the extraction of mRNA from total RNA samples

Tissue Type	Av. % mRNA extracted (wt/wt)	Average 260/280	Average 260/230	Ratio Quality Assessment	Total Number of Samples extracted
Mammary Artery	1.4	1.69	2.42	1:3	4

This extraction had a very high failure rate - on one occasion yielding no mRNA after purification. The sample which was considered of appropriate quality to take through to the next stage of the phage display library production - MT1801 - gave the following results: yield 1.34% (wt/wt), A_{260}/A_{280} 1.74 and A_{260}/A_{230} 2.03. Expected yield of mRNA from total RNA starting material is 1-2%,[30] so the extraction can be considered to have proceeded successfully. The total RNA sample pool for this mRNA sample is described below.

Twelve total RNA samples from internal mammary artery, extracted from seven individuals were taken to prepare MT1801. The details for these samples were as follows:

Table 2.2.2d: total RNA samples used to prepare mRNA sample M1801

Sample	Conc. of total RNA (ng/µl)	A_{260} /A_{280}	A_{260} /A_{230}	Volume pooled (µl)	Mass pooled (µg)	Proportion of final pool %	Proportion of final pool (patient %)
MT0101	640	1.91	1.63	10	6.4	4	4
MT0301	337	1.94	1.87	20	6.74	4	13
MT0302	427	1.91	1.9	20	8.54	5	
MT0303	348	1.93	1.82	20	6.96	4	
MT0401	850	1.91	2.16	20	17	10	22
MT0402	630	1.98	2.02	20	12.6	8	
MT0403	340	1.96	1.95	20	6.8	4	
MT0602	286	1.93	1.99	20	5.72	3	7
MT0603	259	1.95	1.82	20	5.18	3	
MT0701	878	1.93	1.95	50	43.9	26	26
MT0801	224	1.9	1.92	20	4.48	3	3
MT1002	864	1.87	1.85	50	43.2	26	26

To improve the yield of total RNA extracted from vascular tissues there are some methods to maximise the output from phenol:choloroform extraction. By fast-freezing the samples in liquid nitrogen after collection the RNA is protected from damaging conditions, as the RNase inhibitor cannot access all tissues while the cells are intact. This would involve better preparation and arrangement to carry liquid nitrogen into theatre, but is achievable. By vortexing the phenol and chloroform mixture more thoroughly the mixture and therefore the extraction can be maximised, and by making several extractions from the phenol:chloroform layer using aqueous solvents (back extraction) the yield could be maximised. Reasons for the low

yielding mRNA extractions are harder to define, though this may be due to inexperience with the Straight A's™ kit and its components.

2.2.3 cDNA formation

A sample of the pooled sample Poly(A)$^+$ mRNA noted above was used to prepare a cDNA library. The formation of a cDNA library was carried out using Novagen's T7Select® 10-3 OrientExpress™ cDNA Cloning System, Cat# 70581-3. For this study, oligo(dT) primed cDNA synthesis was chosen.

Reverse transcription and second strand synthesis was carried out to prepare the cDNA library, and *Hin*d III and *Eco*R I restriction enzyme sites were added to the ends of the cDNA. These were digested sequentially with *Hin*d III and *Eco*R I, and gel column fractionation was used to separate out any excess linkers and small cDNA synthesis products.

Of the five fractions which were collected from the column, the recommended fraction for collection was fraction three. Initial assaying of the fraction three output (F3a), containing the longest cDNA fragments, suggested that the concentration of this fraction was very low and the quality was poor (see details below). The required concentration for the next stage of phage library preparation was 20 - 40 ng/µl. This fraction was re-precipitated and dissolved in a smaller amount of solvent, giving a better quality result (F3b), but still poor concentration for the next step of the phage library preparation. The next fraction (F4) was much more concentrated.

It was this fraction that was taken forward to the next stage of library preparation, although it was noted that this fraction may have contained some smaller length DNA fragments, which would have lowered the average insert size of the library.

Table 2.2.3: Outputs from cDNA fractionation

Fraction	Concentration (ng/µl)	Volume (µl)	A_{260}/A_{280}	A_{260}/A_{230}
F3a	8.1	20	1.78	1.66
F3b	11.0	10	1.99	2.21
F4	33.3	10	1.88	2.46

2.2.4 Packaging into Phage Vector Arms

For this study, a mid-copy number display vector was used. This display vector (T7Select® 10-3) displays 5-15 copies of the insert on the head group of the phage. The control over this copy number is brought about by modification of the number of phage capsid proteins produced by the phage. The remaining capsid proteins are supplemented by the *E. coli* host, strain BLT5615.

The prepared cDNA library (1 µl, 33 ng, A_{260}/A_{280} 1.88 and A_{260}/A_{230} 2.46) was inserted into vector arms for the T7Select® 10-3 vector, at a molar ratio of 2:1. This was within the recommended molar ratio of insert:vector of between 1:1 and 3:1 for maximal cloning efficiency. Higher molar ratios of the insert are used to minimise the number of vectors which self-ligate, leading to 'empty' vectors with no insert. These are likely to amplify at a faster rate than vectors containing an insert, as the infected *E. coli* molecular machinery will not need to construct the insert.[31, 32]

2.2.5 Phage Library Validation

There are several steps to the validation of the prepared library. The diversity of the library (number of unique clones) must be determined before any amplification takes place. After amplification of the library there will be many copies of each unique

clone. At this point a sample of clones is taken and the sequences determined by PCR. Sequencing allows quantification of vectors in which there is no insert, and allows validation of those inserts which have been successfully cloned.

Two samples were taken from the phage library preparation and were titred by plaque counting assay, but when grown, the *E. coli* lawns on both of these plates were entirely lysed after 3 hours - suggesting a population of greater than 5,000 plaques for 100 µl at a dilution of 10^{-2}. This suggests a diversity of greater than 1.5 x 10^5 unique clones.

The remaining library was amplified in two halves in log phase *E. coli* culture, strain BLT5615 at OD_{600} ~0.5. (14 µl phage solution containing 70,000 pfu phage in 2.5 ml LB media containing $5x10^8$ *E. coli* cells assuming an OD_{600} of 0.5 exactly, multiplicity of infection of $1.4x10^{-4}$). This is an erroneously low MOI, as recommended MOI for amplification is 0.001 to 0.01. To avoid disproportionation destroying the entire library diversity, the initial amplification was carried out in two halves. After confirmation of the diversity of both halves of the library, these were later combined. The two amplified libraries were plated at a variety of concentrations, and were determined to have amplified to $5x10^7$ pfu/ml. This titre is low, as a healthy amplified library should have a titre of 10^{10} pfu/ml, and this low titre is likely due to the low multiplicity of infection.

Sixteen isolated plaques were picked from the surface of the agar plates and the genetic material from these picks were amplified by touchdown PCR. Agarose gel electrophoresis was run on the PCR products to confirm that PCR had been successful.

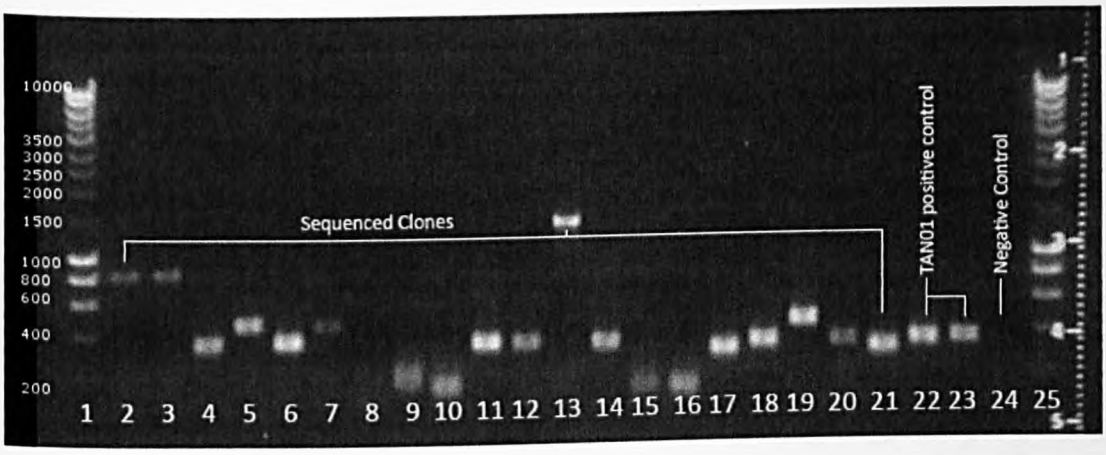


Figure 2.2.5: Ethidium bromide stained agarose gel showing a sample of clones from the phage display library after PCR to amplify the insert DNA. Hyperladder in lanes 1 and 25 allows estimation of insert length. PCR controls in lanes 22, 23 and 24 (positive, positive and negative).

Further plaques were picked and the libraries were amplified once again but were again shown to have amplified to $5x10^7$ pfu/ml. Further samples collected from the second amplification were picked and their genetic information amplified by PCR. 16 picks from this second amplification were cleaned up by filtration and sequenced. The returned sequences were aligned using the MegaBLAST algorithm against the coding region for the T7Select® 10-3 vector. (Results available in Appendix 8.3.2) Any sequences shown to be entirely vector with no insert were recorded to ensure that these did not interfere with later bioinformatics.

Table 2.2.5a: Sequencing results from the library

Empty Vector	Failed Sequencing	Further Bioinformatic Analysis
2	9	5

All returned sequences were then entered into the DNAStar Lasergene® software SeqMan Pro. This program aligned any highly similar sections of clone sequence within set parameters to identify any clones which appear more than once, trimmed all low quality ends and removed any T7Select® vector contamination from the sequences, so that the final consensus sequence can be used to accurately represent

the sequence present in the phage clone.

Table 2.2.5b: Alignment of sequences into contigs		
Clone ID	Section	Contig
Lib2	(54>194)	Lib Contig 3
Lib7	(41>202)	Lib Contig 4
Lib8	(84>180)	Lib Contig 4
Lib12	(80>508)	Lib Contig 1
Lib14	(90>247)	Lib Contig 2

The output from this alignment is a series of consensus sequences from these contiguous regions, or 'contigs'. For parameters and full information, see appendix 8.3.3. Remaining results were assayed by Basic Local Alignment Search Tool (BLAST), for more information see chapter 3.1.5. The top hit from this BLAST analysis is given below, full BLAST analysis is available in E-Appendix 8.3.4.

Table 2.2.5c: MegaBLAST algorithm analysis of contigs				
Contig	Accession No.	Name	Score (bits)	E.value
Lib Contig 1	NM_006028.3	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 3B (HTR3B), mRNA	73.4	$4x10^{-12}$
Lib Contig 2	NM_001184735.1	Homo sapiens astrotactin 2 (ASTN2), transcript variant 6, mRNA	187	$2x10^{-46}$
Lib Contig 3	NM_000440.2	Homo sapiens phosphodiesterase 6A, cGMP-specific, rod, alpha (PDE6A), mRNA	124	$4x10^{-24}$
Lib Contig 4	NM_000440.2	Homo sapiens phosphodiesterase 6A, cGMP-specific, rod, alpha (PDE6A), mRNA	142	$6x10^{-29}$

2.2.6 Individual Libraries

In a move towards personalised medicine, we would like to be able to prepare polypeptide libraries tailored to one individual, or small cohort of patient's tissue. A personalised polypeptide phage display library would allow interaction studies for rare diseases and uncommon side effects, and could also allow for development of improved personalised treatment, with increased efficacy and reduced toxicity.

Based on the average total sample size from any single patient (957 mg) and the average extraction of total RNA (200 µg) and mRNA (3.1 µg) a single patient would

give enough mRNA, provided that all tissue was processed successfully to mRNA, to produce a phage display library. Most of the failures in the total RNA extraction process were contamination from protein or alcohol, and could be corrected by reprecipitation and dissolving in fresh solvent, but identification of the reason behind the failure in the mRNA extraction would be essential in order to avoid this problem when only one tissue sample would be available for the patient.

From this initial data, it should be possible to modify the phage display library preparation to provide a phage library displaying peptides representing the proteome for one individual, as a real step towards personalised medicine.

2.3 Conclusions

2.3.1 Methodology

A bank of human vascular tissue was collected with appropriate consent, and total RNA was extracted from many samples. From a pool of these total RNA samples, mRNA was extracted and a cDNA library formed. A T7Select® Phage display library was prepared and validated.

2.3.2 Validation

The library diversity was shown to be greater than 5×10^5. Sixteen clones were picked from the library, their genetic material amplified by PCR and their sequences determined. The sequences returned included two empty vector sequences and five further sequences which returned hits with homo sapiens mRNA when analysed with the megaBLAST program.

2.3.3 Personalisation

Initial data suggests that it should be possible to use this method of phage display library preparation to provide a phage display library representing the proteome of one single individual. In this way, the assays carried out would be personalised to that individual, allowing investigation into rare drug interactions as well as rare diseases.

2.4 Methods and Materials

2.4.1 General Experimental

All consenting and tissue collection was carried out under HTA guidelines, with written informed consent, at University Hospital, Walsgrave, Coventry. Tissue storage under HTA guidelines and all tissue processing to mRNA took place at the Clinical Sciences Research Institute, Walsgrave, Coventry. Researcher were protected with appropriate vaccinations and HTA researcher registration with the CSRI HTA license Designated Individual, Prof. John Davey. Phage library preparation from mRNA and validation carried out under the HSE's GMO (contained use) guidelines at Tangent Reprofilling's labs at HRI, Wellesbourne.

All plasticware was classified RNase-free. RNA*later* RNase inhibitor solution from Ambion[33] (Catalogue #7020) was used to transport the tissues in ice from theatre to the biological lab, with a lag time of maximum 30 minutes between extraction of the tissue and processing, or freezing to -80°C depending on the process in question. *mir*Vana™ miRNA Isolation Kit from Ambion[16, 34] (Catalogue #1560, 1561) was used to extract total RNA from the tissues obtained. Straight A's™ mRNA extraction Kit from Novagen (Catalogue #69962-3) was used to extract mRNA from the total RNA samples. The OrientExpress™ Oligo(dT) Primer cDNA synthesis kit from

Novagen (Catalogue #69992-3) and the T7 Select® Phage Display System from Novagen (Catalogue #70550-3) were used to prepare the T7 Phage Display Library. Water used in total RNA and mRNA extraction was RNase free water, purchased from Ambion. Water used in cDNA and Phage Library preparation was milli-Q filtered sterile water prepared on site.

Lysogeny broth was made up as follows: to tryptone (8 g), yeast (4 g), NaCl (8 g), water was added to make up to 800 ml and stirred until all material dissolved. This mixture was autoclaved and after cooling to below ~40°C ampicillin (800 ul, 50 mg/ml) was added and after cooling to room temperature the media was stored at 4C.

LB agarose and LB agar were made up as follows: to 200 ml liquid LB made up as above prior to autoclaving was added 1.2 g agarose. This mixture was autoclaved and after cooling to room temperature was stored at 4°C. To 800 ml liquid LB made up as above prior to autoclaving was added 12 g agar. This mixture was autoclaved and after cooling to below ~40°C ampicillin (800 ul, 50 mg/ml) was added. This molten solution was then poured into disposable plastic petri dishes, to half full, under a laminar flow hood to prevent contamination and allowed to set under flow. Plates were placed in sealed plastic bags and stored at 4°C.

To prepare *E. coli* BLT5615 working plates and working stocks: *E. coli* strain BLT5615 (Glycerol Stock, Novagen, Cat# 69905) was taken from the -80°C freezer and kept on ice to minimise defrosting of the tube. Using a sterile loop the top of the frozen glycerol stock was scraped and the loop then streaked across the top of an agar plate containing ampicillin, along one edge only without damaging the agar surface. Using a new sterile loop, the plate was turned by 60° and the plate was streaked again, spreading the *E. coli* solution across a fresh edge of the plate. This turning and streaking was repeated a further two times and the plate was sealed with

parafilm, inverted and incubated overnight at 37°C. After growth, the plates were stored at 4°C. To prepare a working (overnight) stock of *E. coli* one cleanly defined colony was removed from the plate using a P1000 pipette tip and placed in 100 ml liquid LB in a 250 ml flask to allow adequate aeration. This flask was agitated at 37°C overnight to give an OD_{600} ~ 2.0. This stock *E. coli* solution was stored at 4°C for up to two weeks.

'Old method' PCR mastermix (per sample): 5 µl 10X Nova*Taq*™ buffer with $MgCl_2$ (Novagen, Cat#71005), 1 µl T7Select® Up primer (5 pmol/ µl), 1 µl T7Select® Down primer (5 pmol/ µl) (Sigma Genosys Primers), 1 µl dNTP mix (10mM each dNTP) (Invitrogen, Cat#18427-013), 0.25 µl Nova*Taq*™ DNA polymerase (1.25Ux12) (Novagen, Cat#71005), made up to 49 µl with water. **Touchdown PCR mastermix** (per sample): 5 µl 10X Nova*Taq*™ buffer with $MgCl_2$, 1 µl T7Select® Up primer (5 pmol/ µl), 1 µl T7 Select Down primer (5 pmol/ µl), 1 µl dNTP mix (10mM each dNTP), 0.25 µl Nova*Taq*™ DNA polymerase (1.25Ux12), made up to 40 µl with water.

Two positive control phage solutions (TAN01 and TAN02) were used to confirm the success of the PCR reaction in agarose gel electrophoresis.

'Old method' PCR: Phage pick solution (1 µl) and PCR mix (49 µl) was added to the wells of the PCR plate and PCR was set up to run: 3 mins, 94°C, 35 cycles: 0.5mins, 94°C; 1min, 50°C; 1min, 72°C, final extension: 6 mins, 72°C; hold at 25°C.

For agarose gel electrophoresis, gels were prepared as follows – TAE (100 ml, 1X) and agarose (1 g) were melted in microwave (scaled up for larger gels). Ethidium bromide (10 µl, 1 mg/ml, 10 µg) was added and poured into gel frame and allowed to set with combs in place to prepare the running lanes. Ethidium bromide (50 µl) was added to the running solution (1X TAE, 500ml).

After the gel was set the comb was removed and the PCR product (8 µl) was loaded with 20% volume Crystal 5X DNA loading buffer blue (Bioline, Cat# BIO-37045). 5 ul Hyperladder™ I (Bioline, Cat# Bio-33025) was added to the lanes on either side of the samples being assayed. Where two combs were used, two lanes of Hyperladder™ I were added on both rows. The gel was run in the electrophoresis module at 100-130 V for 30-45 mins. Gel was visualised under UV light and photographed.

PCR products were subjected to cleanup using Millipore filter plates (Millipore, Cat# LSKMPCR10) under vacuum before sequencing.

Sequencing mastermix (per sample): 2 µl Big Dye® Terminator v3.1 (Applied Biosystems, Cat# 4337454), 0.4 µl T7Select® Up Primer, water to make up to 9 µl. 1 µl cleaned DNA solution added before submission to the HRI sequencing service.

2.4.2 Tissue Sample Collection

Artery and vein material was collected in cryogenic tubes containing a RNase inhibitor solution RNAlater®, for transport in ice back to the research lab (30 minutes maximum transport time).

Sample size ranged from 3896 - 61 mg for the internal mammary artery samples, average 957 mg ± 670 mg and 548 - 43 mg for the saphenous vein samples, average 165 mg ± 170 mg. These were divided into smaller samples (998 - 56 mg for the internal mammary artery samples, average 279 mg ± 130 mg and 213 - 43 mg for the saphenous vein samples, average 120 mg ± 56 mg) for later processing. Samples were not weighed until work with the respective sample was about to start, to avoid prolonged exposure of the sample to room temperature, risking degradation of any nucleic material.

Samples were labelled MT (project code) xx (Patient Number) xx (sample number). Any RNase inhibitor was removed and the sample was frozen and stored at -80°C.

2.4.2 Total RNA extraction

The sample was weighed and 10:1 volume:mass (ml:g) lysis/binding buffer was added to the tissue. The sample was homogenised using a mechanical homogeniser (details) until no further lumps of tissue were visible. Homogenate additive (0.1 eq. volume of lysis/binding buffer) was added and mixed by brief vortexing. The solution was separated into two tubes, and any remaining lumps of adipose tissue were left behind. Phenol:chloroform:isoamyl alcohol (1 eq. volume of lysis/binding buffer) was added, followed by 1-bromo-3-chloro-propane (0.1 eq. volume lysis/binding buffer). The tubes were mixed thoroughly by inverting and vortexing, then centrifuged at 12Kxg for 5 mins. The aqueous layer was removed without disturbing the protein band at the interface and placed in a new tube, recording the approximate volume of material. Ethanol (neat, 1.25 volumes) was added to precipitate the nucleic acids. While a tube of RNase free water was being heated to 95°C, the ethanol solution was added to a filter cartridge in a collecting tube in 700 µl volumes and the solution was drawn through under centrifuge (12K xg, 10 secs). The run-through from the filter was discarded and after all the material had been added, the product was washed with Wash 1 solution (700 µl, ethanol based solution containing guanidinium thiocyanate to denature any remaining proteins) and drawn through under centrifuge (12K xg, 5 secs), then Wash solution 2/3 (500 µl x 2, ethanol based solution containing EDTA to protect RNA from damage caused by any metal present when eluted from the filter) and drawn through under centrifuge (12K xg, 5 secs). Any remaining solution was removed from the filter by centrifuge

(12K xg, 1 minute) and the collecting tube was discarded. The filter was placed in a fresh tube and the RNA was eluted with 95°C RNase free water (50-150 µl).

The concentration and purity of RNA was determined by NanoDrop[35] assay and the RNA solution was stored at -80°C. An average of 58 µg ± 35 µg total RNA was extracted from each internal mammary artery sample of average 273 mg ± 136 mg and an average of 29 µg ± 14 µg extracted from each saphenous vein sample of average mass 120 mg ± 56 mg, using the Ambion mirVana™ miRNA Isolation Kit.

2.4.3 mRNA extraction[4]

1mM DTT and the total RNA samples were thawed in the fridge and on ice respectively. The volume of particles required for the extraction was calculated from the known amount of total RNA, and the volume of lysis buffer required to make the concentration of samples up to 20% of total volume was also calculated. The samples of total RNA were pooled and if necessary concentrated to a concentration of over 0.5 mg/ml. (Concentration was carried out by adding 2 volume equivalents of ethanol to the pooled samples on dry ice and leaving to precipitate for ten minutes. The solution was then spun at 12K xg for 20 mins at 4°C. All solution was removed and the sample was redissolved in the appropriate amount of RNase free water.) NanoDrop assay was carried out before continuing to confirm sample concentration. Four particle volumes of lysis buffer were made up (1 ml buffer to 10 µl DTT). One volume of the particles was pipetted out, captured in the magnetic stand and the supernatant removed. The particles were then washed with two sequential volumes of the lysis buffer. One further volume of lysis buffer plus the extra lysis buffer to make the concentration of total RNA up to 20% were added to the particles and they were incubated at room temperature for 5 mins. The particles were captured and the

supernatant removed. The particles were then washed with 2 volumes of wash buffer and all solution was removed from the tube. 0.5 ml RNase free water was added and the tube was incubated for ten minutes at 60°C to elute the mRNA. The supernatant, now containing the mRNA, was transferred to a clean tube and 2 µl glycogen, 50 µl sodium acetate and 330 µl isopropyl alcohol were added. The tube was then centrifuged for ten minutes at 12K xg, following which the supernatant was removed (the pellet is not visible at this point). 0.5 ml 70% ethanol was added and the tube was spun for a further 5 mins. The supernatant was removed and the mRNA was dissolved in 0.02 ml RNase free water. The concentration of the mRNA was determined by NanoDrop[15] assay and stored at -80°C.

Twelve samples from seven patients were sampled, and a pool of 290 µl of 0.58 ng/µl total RNA was collected. From this pooled total RNA, 2.48 µg PolyA+ mRNA was extracted using the Novagen Straight A's™ mRNA Isolation Kit.

2.4.4 cDNA Formation

The formation of a cDNA library from an mRNA sample was carried out using Novagen's T7Select®10-3 OrientExpress™ cDNA Cloning System and mRNA starting material MT1801. Nanodrop: 123.5 ng/µl (A260/280 = 1.81 and A260/230 = 2.29). For this study, oligo(dT) primed cDNA synthesis was chosen.

mRNA (1.99 µg, 16.1 µl MT1801) and oligo(dT) primer (1 µg at 1 µg/µl) was mixed, and made up to 20 µl with water. This was incubated at 70°C for ten minutes and chilled on ice, followed by 12xg centrifuge for ten seconds.

First strand buffer (5X, 10 µl), DTT (100mM, 5 µl), methylation dNTP mix (10X, 2.5 µl) was added and made up to 46 µl with water. Solution was mixed and equilibrated for 1 min at 37°C. MMLV reverse transcriptase (4 µl, 800 units) was

added and incubated at 37°C for a further 60 mins. The reaction mixture was then heated at 70°C for ten minutes and cooled on ice. Centrifuged briefly to collect product.

Second strand buffer (5X, 50 µl), DTT (100mM, 6 µl), methylation dNTP mix (10X, 2 µl), DNA polymerase (5 µl, 50units) and RNase H (1.6 µl, 1.6 units) were added and made up to 250 µl total volume. This was then incubated at r.t. for 90 minutes.

TE buffered phenyl:choloroform:isoamyl alcohol (25:24:1, 250 µl) was added and vortexed for 30 seconds, followed by centrifugation for 1 minute at 12xg. The aqueous phase was removed to fresh tube. Glycogen (1 µl), ammonium acetate (4M, 250 µl) and isopropanol (300 µl) were added to precipitated the nucleotide material. The reaction mixture was then inverted several times and centrifuged at 12xg for 8 mins. No pellet was visible, so the reaction was centrifuged for a further 3 mins and the supernatant was removed. The pellet was washed with 70% ethanol and centrifuged for 3 mins. Repeated with 100% ethanol. The washed pellet was then suspended in TE buffer (20 µl) and left overnight at -20°C.

Flush buffer (3 µl), DTT (100mM, 1.5 µl), dNTPs (3 µl, 1mM made up from 10mM solution – 1 µl in 9 µl water), T4 DNA polymerase (0.6 µl, 1.5units) were added with water to make up to 30 µl. Solution was mixed gently and allowed to incubate at r.t. for 20 mins. TE buffer (20 µl) was added followed by TE buffered phenyl:choloroform:isoamyl alcohol (50 µl). This mixture was vortexed for 30 seconds and centrifuged at 12K xg for 1 min. The aqueous layer was removed to a fresh tube. This was repeated with further TE buffered phenyl:choloroform:isoamyl alcohol (50 µl). Glycogen (1 µl), ammonium acetate (4M, 50 µl) and ethanol (250 µl) were added and mixed well and left at -20°C for 5 h. Product was centrifuged at 12K xg for 10 mins. Supernatant was removed and pellet was washed with 70%

ethanol and centrifuged for 3 mins. Repeated with 100% ethanol. Suspended in TE buffer (10 µl).

Ligation buffer (10X, 2 µl), ATP (2 µl, 1mM made up from 10mM solution – 1 µl in 9 µl water), DTT (100mM, 2 µl), EcoRI (2 µl), T4 polynucleotide kinase (0.5 µl, 5units) were added and made up to 20 µl (including volume addition of T4 DNA ligase in later step). The mixture was incubated for 5 mins at 37°C, and placed on ice. T4 DNA ligase (1.75 µl, 7 units) were added. This was incubated at room temperature overnight.

The ligase was inactivated by incubating the reaction at 70°C for ten minutes. *Hin*dIII buffer (10X, 10 µl), *Hin*d III (5 µl, 100 units) were added and made up to a total volume of 100 µl. This mixture was incubated at 37°C for 2 h. 10 µl *Eco*RI adjustment buffer and *Eco*RI (5 µl, 100units) added. This was incubated at 37°C for 4 h.

A mini-column was prepared for the sample. Gel filtration resin (2 ml) was added to column and allowed to settle and pack. The column was equilibrated with column buffer (5x1 ml 1X made up as 700 µl buffer in 6.3 ml milliQ water). The cDNA sample was extracted with phenol:choloroform:isoamyl alcohol (115 µl), and the aqueous layer was removed to prepared column. The sample was allowed to settle into the column (fraction 1). Washed through with 200 µl column buffer (fraction 2). Washed through with 2x250 µl column buffer (fractions 3 and 4), washed through with 250 µl water (fraction 5).

To fractions 3 and 4 was added glycogen (1 µl), isopropyl alcohol (150 µl) - salt is already present in column. The precipitation was allowed to incubate for 5 mins at room temperature. This mixture was spun for ten minutes at 12K xg and washed with 70% and 100% ethanol. The final pellet was dissolved in TE Buffer (10 µl).

Fraction 3 was initially dissolved in 20 µl TE buffer but reprecipitated in glycogen, ammonium acetate and ethanol at -20°C, followed by washing with 70% and 100% ethanol.

The resulting cDNA was analysed by nanodrop and stored at -20°C.

MT31F3: *Nanodrop:* 11.0 ng/µl, $A_{260/280}$ = 1.99, $A_{260/230}$ = 2.21

MT31F4: *Nanodrop:* 33.3 ng/ml, $A_{260/280}$ = 1.88, $A_{260/230}$ = 2.46

2.4.5 Packaging into Phage Vector Arms

To MT31F4 (1 µl, 3.33 ng) was added ligation buffer (0.5 µl, 10X), ATP (0.5 µl, 10mM), DTT (0.5 µl, 100mM), T7 Select Vector Arms (0.8 µl, 0.5 µg, 0.02 pmol), T4 DNA ligase (1 µl, 0.4 U / µl, diluted from 4 U / µl: 1 µl in 9 µl water) and made up to a total volume of 5 µl. Mixed by pipette tip and incubated at 18°C overnight. Reaction mixture was added to T7 Phage Packaging Extract (25 µl). This was incubated at 20°C for 2 h.

2x1 µl aliquots (D1 and D2) were taken from the mixture and stored for diversification assay – see below. Stock *E. coli* strain BLT5615 (500 µl, OD_{600} ~2.0 in LB medium with ampicillin - see general experimental for recipe) was added to LB medium (5 ml) with ampicillin (1 µl/ml) for 2 h at 37°C. At 1.5 h IPTG (500 µl, 100 mM) was added. The remaining reaction mixture was split into two halves (14 µl each), and each was added to *E. coli* (2.5 ml, OD_{600} ~0.5) as prepared above, giving two amplification reactions. These were shaken at 37°C for 3.5 h and then centrifuged at 8000 rpm for 10 mins. The two supernatants were decanted off and stored separately at 4°C as Library 1 and Library 2. These two libraries were eventually combined and after amplification a series of samples were taken to

prepare glycerol stocks for storage at -80°C. Glycerol stocks were prepared as 1:9 80% glycerol solution to LB phage solution.

2.4.6 Phage Library Validation

As a diversification assay the two 1 µl aliquots taken from the unamplified library - D1 and D2 were added to water (99 µl) and then added to *E. coli* solution (100 µl, OD_{600} ~2.0) and IPTG (100 µl, 100 mM). Molten LB agarose (3 ml, ~50°C) was added and the mixture was spread across the top of a prewarmed agar plate containing ampicillin (1 µl/ml). See general experimental for preparation of LB agarose and agar plates. This agarose layer was allowed to set, and the plate was inverted and incubated at 37°C for 4 h. These *E. coli* lawn on these plates were completely lysed, giving an undefined diversity of $>10^6$.

To quantify the amplified library, the amplified phage solution was diluted to four different dilutions (10^{-2} - 10^{-8}) and 100 µl of the dilution was added to *E. coli* solution (100 µl, OD_{600} ~2.0) and IPTG (100 µl, 100 mM). Molten LB agarose (3 ml, ~50°C) was added and the mixture was spread across the top of a prewarmed agar plate containing ampicillin (1 µl/ml). This agarose layer was allowed to set, and the plate was inverted and incubated at 37°C for 4 h (or until plaques appeared evenly across the plate). In the dilutions of 10^{-2} and 10^{-4} the *E. coli* lawn was completely lysed, but in the case of the 10^{-6} and 10^{-8} dilutions, there were a countable number of plaques visible. Several plaques were picked from each of the two libraries were taken (1 plug of agarose in 100 µl water) and a PCR mastermix was made up - see General Experimental.

Phage pick solution (1 µl) and PCR mix (49 µl) was added to the wells of the PCR plate and PCR was run - cycle program described in general experimental.

An agarose gel was prepared for electrophoresis. Ethidium bromide (80 µl) was added to the running solution (1X TAE, 800 ml). PCR was repeated using touchdown PCR technique, see general experimental for mastermix and cycle program. See gel result from touchdown PCR in main text.

Amplification was repeated on the two libraries (100 µl of amplified library in log phase *E. coli* 5 ml, OD_{600} ~0.5) and these were poured onto plates in dilutions as above. 16 new plaques were picked. See main text for gel results.

The PCR product was introduced into a 96-well Millipore filter plate (Millipore, Cat# LSKMPCR10) and filtered under vacuum. Water (100 µl) was added and the plate was agitated for 5 mins before collection of the clean PCR product solution. Big Dye™ sequencing master mix was prepared as described in the general experimental and 9 µl was added to each well, followed by 1 µl DNA solution. This plate was submitted to the HRI sequencing service.

Picked clone solutions were archived as glycerol stocks (1:9 80% glycerol solution to phage solution) and stored at -80°C.

2.5 References

1. P. J. Hergenrother, K. M. Depew and S. L. Schreiber, *J. Am. Chem. Soc.*, 2000, **122**, 7849-7850.
2. J. R. Knowles, *Acc. Chem. Res.*, 1972, **5**, 155-&.
3. C. F. Barbas, D. R. Burton, J. K. Scott and G. J. Silverman, *Phage Display: A Laboratory Manual*, CSHL Press, 2004.

4. J. Sambrook, E. F. Fritsch and T. Maniatis, *Molecular Cloning, A Laboratory Manual*, CSHL Press, New York, 1989.

5. J. Smith, R. E. Kontermann, J. Embleton and S. Kumar, *Faseb J.*, 2005, **19**, 331-341.

6. G. P. Smith and V. A. Petrenko, *Chem. Rev.*, 1997, **97**, 391-410.

7. L. R. H. Krumpe, A. J. Atkinson, G. W. Smythers, A. Kandel, K. M. Schumacher, J. B. McMahon, L. Makowski and T. Mori, *Proteomics*, 2006, **6**, 4210-4222.

8. Novagen, 70014 T7Select® Packaging Kit http:// www.merckbiosciences.co.uk/product/70014, Accessed 15/11/2010.

9. Novagen, 69905 BLT5615 Glycerol Stock http://www.merck-chemicals.com/united-kingdom/life-science-research/blt5615-glycerol-stock/EMD_BIO-69905/p_wqGb.s1OB5MAAAEjMRx9.zLX? PortalCatalogID=merck4biosciences, Accessed 12/12/10, 2010.

10. HTA, HTA Code Of Practice 1 - Consent, 2009, http://www.hta.gov.uk/legislationpoliciesandcodesofpractice/codesofpractice/code1consent.cfm.

11. HTA, HTA Code Of Practice 9 - Research, 2009, http://www.hta.gov.uk/legislationpoliciesandcodesofpractice/codesofpractice/code9research.cfm.

12. HTA, HTA Code Of Practice 5 - Disposal of Human Tissue, 2009, http://www.hta.gov.uk/legislationpoliciesandcodesofpractice/codesofpractice/code5disposal.cfm.

13. H. Kessler, *Molecular Diagnostics of Infectious Diseases*, De Gruyter, 2010.

14. Ambion, Make Time Stand Still with RNAlater®, http://www.ambion.com/techlib/resources/RNAlater/index.html, Accessed 15/11/2010.

15. J. B. de Kok, R. W. Roelofs, B. A. Giesendorf, J. L. Pennings, E. T. Waas, T. Feuth, D. W. Swinkels and P. N. Span, *Lab. Investig.*, 2005, **85**, 154-159.

16. P. Chomczynski and N. Sacchi, *Anal. Biochem.*, 1987, **162**, 156-159.

17. P. Chomczynski and N. Sacchi, *Nat. Protoc.*, 2006, **1**, 581-585.

18. J. Kieleczawa, *DNA sequencing II: optimizing preparation and cleanup*, Jones and Bartlett Publishers, 2006.

19. M. D. Lopez and T. Samuelsson, *RNA-Publ. RNA Soc.*, 2008, **14**, 1-10.

20. J. Guhaniyogi and G. Brewer, *Gene*, 2001, **265**, 11-23.

21. Howe, *Gene Cloning and Manipultion*, Cambridge University Press, 1995.

22. R. Farrell, *RNA Methodologies: Laboratory Guide for Isolation and Characterization*, Elsevier Science & Technology, 2009.

23. Novagen, 69993 OrientExpress™ Oligo(dT) cDNA Synthesis Kit, http://www.merckbiosciences.co.uk/product/69993, Accessed 15/11/2010.

24. S. Y. Lee, *Trends Biotechnol.*, 1996, **14**, 98-105.

25. F. Sanger and A. R. Coulson, *J. Mol. Biol.*, 1975, **94**, 441-&.

26. F. Sanger, S. Nicklen and A. R. Coulson, *Proc. Natl. Acad. Sci. U. S. A.*, 1977, **74**, 5463-5467.

27. L. M. Smith, J. Z. Sanders, R. J. Kaiser, P. Hughes, C. Dodd, C. R. Connell, C. Heiner, S. B. H. Kent and L. E. Hood, *Nature*, 1986, **321**, 674-679.

28. X. H. C. Huang, M. A. Quesada and R. A. Mathies, *Anal. Chem.*, 1992, **64**, 2149-2154.

29. HSE, The Genetically Modified Organisms(Contained Use) Regulations, 2000, http://www.legislation.gov.uk/uksi/2000/2831/contents/made.

30. Ambion, Maximize Your RNA Yield — What Yield to Expect, http://www.ambion.com/techlib/tn/83/8311.html, Accessed 12/12/2010, 2010.

31. H. M. E. Azzazy and W. E. Highsmith, *Clin. Biochem.*, 2002, **35**, 425-445.

32. R. A. Kramer, F. Cox, M. van der Horst, S. van den Oudenrijn, P. C. M. Res, J. Bia, T. Logtenberg and J. de Kruif, *Nucleic Acids Res.*, 2003, **31**.

33. W.-H. Wang, L. G. McNatt, A. R. Shepard, N. Jacobson, D. Y. Nishimura, E. M. Stone, V. C. Sheffield and A. F. Clark, *Mol. Vis.*, 2001, **7**, 89-94.

34. R. Boom, C. J. A. Sol, M. M. M. Salimans, C. L. Jansen, P. M. E. Wertheim-Van Dillen and J. Van Der Noordaa, *J. Clin. Microbiol.*, 1990, **28**, 495-503.

35. H. Kidwell, *labtechnologist.com*, March 2008.

Chapter 3 Biopanning Human Vascular Tissue Phage Library against HMG-CoA Reductase inhibitor Simvastatin

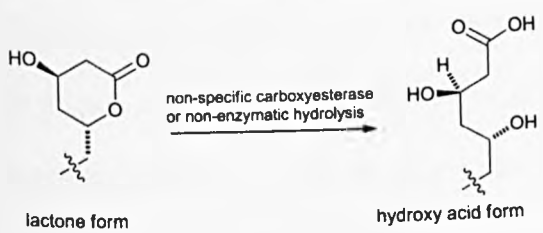
3.1 Introduction

3.1.1 Aims

HMG-CoA reductase inhibitors, or statins, are a billion-dollar per annum class of lipid lowering drugs.[1] In the formation of atherosclerotic plaques, one of the largest components of cardiovascular disease, low density and very low density lipoprotein escort cholesterol into the arterial wall.[2] These atherosclerotic plaques can build up and block the vessel, leading to ischemia, or rupture leading to infarction and stroke. Dysregulation of cholesterol and these lipoproteins increases the risk of serious cardiovascular events.

Statins reduce cholesterol production by inhibiting the rate limiting enzyme in cholesterol synthesis, HMG-CoA reductase. The resultant reduction in the overall amount of cholesterol in the body leads to diminished LDL uptake by the vascular smooth muscle cells[2] and increased LDL uptake by liver cells.[3] These actions are protective against atherosclerosis.

In addition, the statins have been show to have other effects, including interactions outside the known target proteins.[4-11] Some of this pleiotropy is due to effects downstream of the major inhibitory action.



Structures 3.1.1.1 - 3: Statin pro drug and active form, and HMG-CoA, which the statins competitively

Lovastatin and simvastatin, originally marketed by Merck as Mevacor® and Zocor® respectively, but now off-patent, are delivered in the lactone form as a prodrug, which is then hydrolysed by non-specific carboxyesterases or by non-enzymatic processes to the active hydroxy acid. Further metabolism of the drug is carried out by cytochrome P_{450} in the liver.[12, 13] Lactonisation of the hydroxyacid form of all statins can occur by the action of UDP-glucuronic acid.[14]

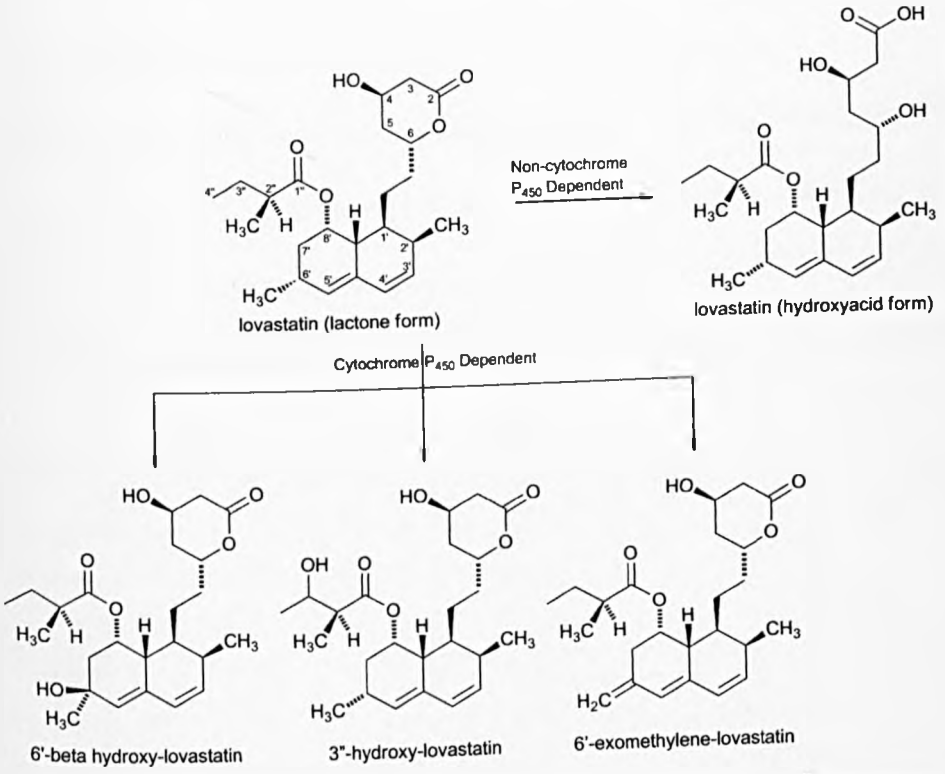
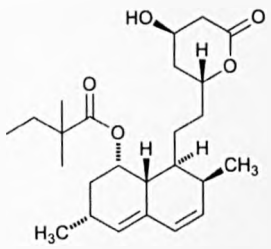


Figure 3.1.1: The metabolism of lovastatin, both cytochrome P_{450} dependant and independant

The accumulation of the statin lactone form has been tentatively associated with higher proportions of myotoxicity, which can lead to the fatal rhabdomyolysis which is a serious side affect of statins occurring in 1 in 10,000 patients on the drug.[15, 16] Evidence that simvastatin therapy is more likely than other statin types to show myotoxic effects is starting appear in the literature.[17]

The responses to statin therapy are complex, and with one in four over the age of forty being prescribed statins on the NHS, investigation into their interactions in the

body continues to be important. This chapter describes the molecular interaction study of the Magic Tag® immobilised simvastatin lactone prodrug against a phage display library prepared from human vascular tissue samples from individuals undergoing surgical treatment for ischemic heart disease.



Simvastatin

Structure 3.1.1.4: HMG-CoA reductase inhibitor simvastatin in its prodrug form

3.1.2 Immobilisation of small molecule onto Magic Tag®

The immobilisation of small molecules onto the Magic Tag® surface is carried out by activation of the surface using UV light at 254 nm. (Figure 3.2.1[18, 19]) This activation produces a highly reactive species which forms a covalent bond with the small molecule in solution. There are five different chemistries in the Magic Tag® system.

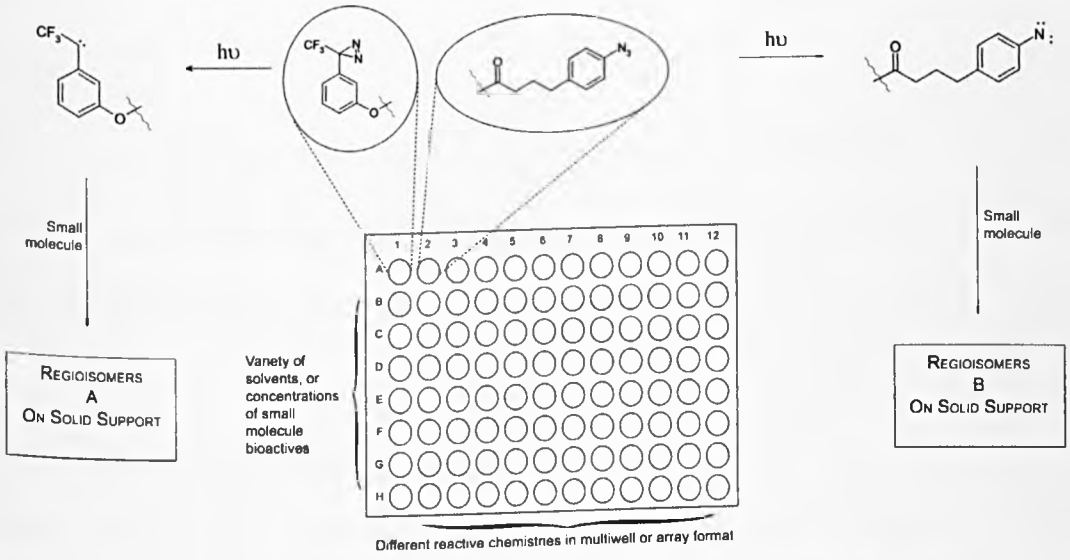
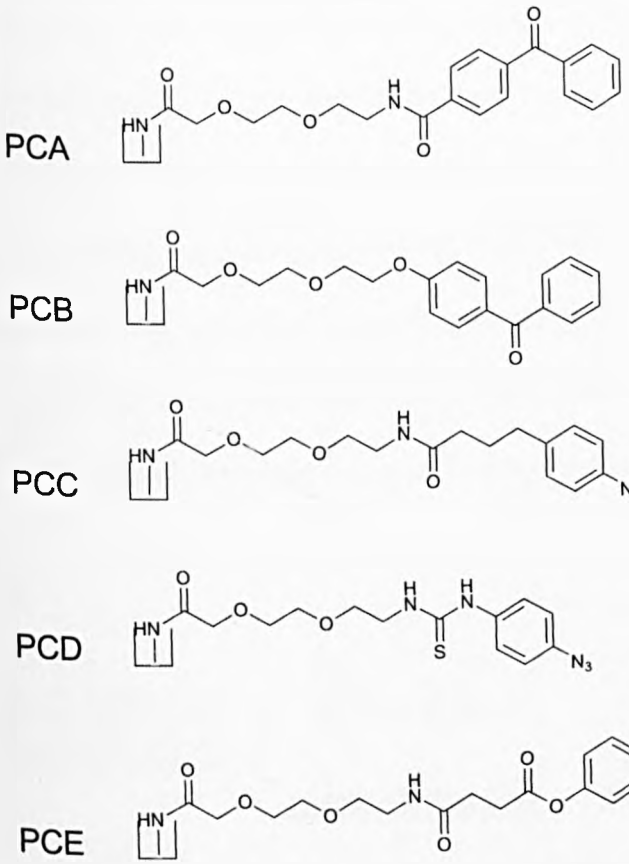


Figure 3.1.2: Layout of a standard Magic Tag® plate, showing activation of two of the five photochemistries and small molecule binding

These are shown as structures 3.1.2.1-5. Each of the different chemistries might be expected to form that covalent bond at a different point in the small molecule, displaying a different area into solution. As many small molecules have poor solubility in aqueous solution, the solution is prepared with a small percentage of

DMSO to encourage dissolution. Magic Tag® surfaces which are irradiated in the presence of the solvent are prepared as control wells.



Structures 3.1.2.1 - 5: Five Magic Tag® chemistries. Amine-functionalised polystyrene plates functionalised with triethylene glycol linkers and photoactive chemical (PC) functional group.

PCA - Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-4-amidobenzophenone

PCB - Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-4-hydroxybenzophenone

PCC - Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-1-(4-azidophenyl) 4-amidobutyric acid

PCD - Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-4-isothiocyanophenyl azide

PCE - Stripwell™ plate-supported 2-(2-[2-acetimidoethoxy]ethoxy)-3-(3-acetoxyphenyl)-3-(trifluoromethyl)-3H-diazirine

One of the five Magic Tag® chemistries, PCE, is so sensitive that preparation of the tag and introduction of the drug solution must be carried out under red light so that no premature activation takes place.

3.1.3 Biopanning against Phage Library

Using phage display technology to carry out the biopanning process as described below (Figure 3.1.4[18]), interactions between drug and phage can be determined, and using bioinformatic tools can suggest the sequence identity of the phage-displayed peptide which is mediating the interaction.[18]

A generalised method for biopanning of the phage library against a Magic Tag® immobilised small molecule is described here.

The stored library must first be freshly amplified in exponential (or log) phase *E. coli*, the point at which the culture growth is characterised by cell doubling, to ensure that the peptide insert is present, complete and unmodified. The small molecule is immobilised into sufficient Magic Tag® wells for three rounds of biopanning by UV irradiation. All of the excess solution is removed and the wells washed with buffered surfactant solution. Duplicate wells irradiated with only solvent are also prepared.[18] A sample of amplified phage display library is introduced to the prepared Magic Tag® wells and interaction is allowed to take place. Excess non-binding phage is washed from the wells using a buffered surfactant solution, leaving only binding

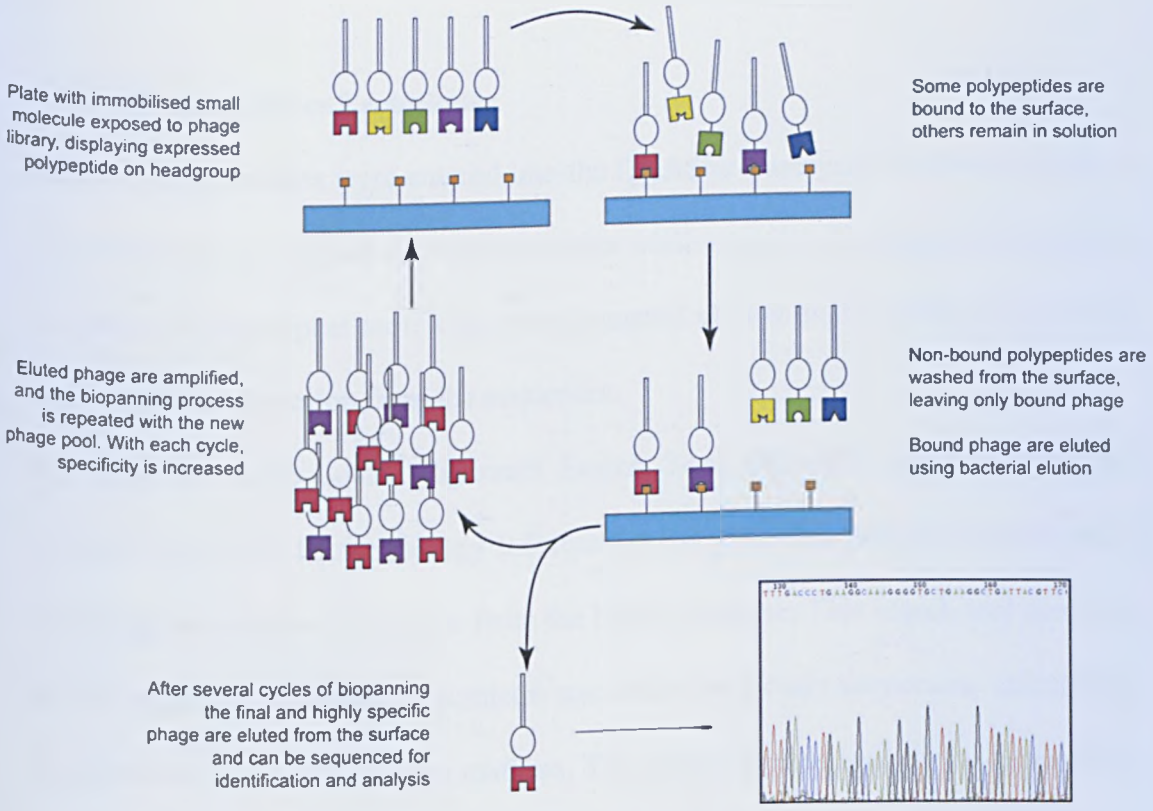


Figure 3.1.4: The biopanning system on an immobilised small molecule

phage. This remaining phage is eluted by infection of log phase *E. coli* strain BLT5615,[20] and the contents of each well are amplified individually at an appropriate MOI (0.001-0.01).[21] A sample of the products of this amplification, now

specific for the well chemistries, are introduced into the appropriate round two wells. The interaction, washing and elution are repeated, followed by amplification of the round two outputs. These were introduced to the round three wells and allowed to interact, the wells washed and the bound phage eluted as before. Round three outputs are not amplified directly as with rounds one and two, but instead are diluted and sampled on an *E. coli* lawn grown on agarose/agar plates. This allows a sample of plaques - each consisting of a single clone to be picked from the surface. By performing PCR on these clones using a primer designed to amplify the insert region and carrying out a simple clean up on the PCR product by filtration, the insert can be sequenced and it is possible to infer a possible polypeptide sequence from this data.

3.1.4 Bioinformatics Techniques

All returned sequences were entered into the DNAStar Lasergene® software SeqMan Pro. This program aligned any highly similar sections of clone sequence to identify any clones which appear more than once, trimmed all low quality ends and removed any vector contamination from the sequences.

By using the Basic Local Alignment Search Tool (BLAST) maintained by the National Center for Biotechnology Information (NCBI), these sequences could then be aligned with known sequences from the human genome. This search tool accesses several sequence databases to compare nucleotide or protein sequences, calculating the statistical significance of the matches. This allows comparison of the determined DNA sequence of the phage displayed insert with the human genome or translated proteome. This data can allow further investigation of the interactions between the small molecule under study with the implicated proteins or mRNA sequence.

3.2 Results and Discussion

3.2.1 Summary of Results

Three rounds of biopanning of human vascular tissue library against immobilised simvastatin in the lactone form were carried out. 120 phage clone picks from the output of round three against simvastatin and 60 phage clone picks from the output of round three against control wells were collected, the DNA insert amplified by PCR, cleaned by filtration and sequenced. Bioinformatics tools were used to identify possible peptide sequences which have shown binding to the immobilised simvastatin.

3.2.2 Immobilisation of simvastatin and control

Immobilisation of simvastatin (lactone form) onto five Magic Tag® chemistries, PCA, PCB, PCC, PCD, PCE, (see structures 3.1.2.1-5) in sufficient wells for three rounds of biopanning, was carried out under UV light for ten minutes. Duplicate Magic Tag® wells were prepared with water immobilised under the same conditions as the simvastatin solution.[18]

For the control wells, the surface was irradiated in the presence of the solvent or solvent mixture used to dissolve the small molecule. In this experiment the solvent should have been DMSO in water as in the control wells, but the solvent used to prepare the control wells was water only. This may have led to some false positive signals, as polypeptides which bind surface modifications originating from the DMSO molecules will not be recognised as background noise.

3.2.3 Biopanning

Three rounds of biopanning were carried out using a freshly amplified sample of the prepared human vascular tissue phage library against one well of each chemistry of the Magic Tag® immobilised simvastatin (lactone form) and the same in the water control wells.[18]

After the third round of biopannning, the phage output should not normally have been amplified, however in this case the amplification was carried out. The amplified well outputs were then diluted and infected *E. coli* lawns prepared at a dilution of 10^{-2} and 10^{-3}. From the most appropriate dilution plate from each well output 24 plaques were picked from the plates prepared from simvastatin immobilised wells (see Figure 3.2.3 for picking method) and 12 plaques were picked from control wells.

Thorough washing of the wells between rounds allows removal of non-specifically bound phage, but there is also the possibility it will remove a proportion of any low-affinity clones, whose expressed sequences may be expected to bind to the small

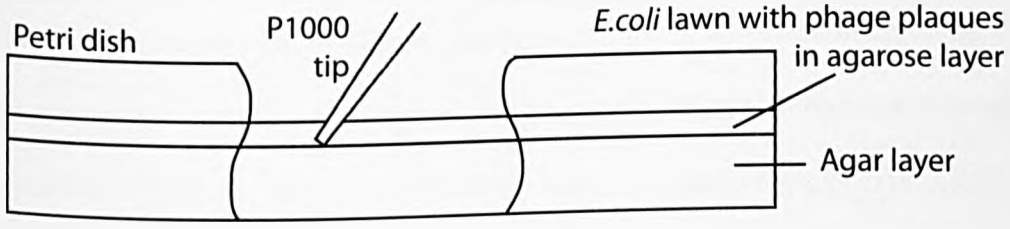


Figure 3.2.3: Picking method for phage plaques on agarose/agar plates

molecule in man and result in some biological effect. Information can also be lost in strong binders, although the use of *E. coli* elution means that the phage clones bound can infect the *E. coli* even without being released from the surface. Hence, by this elution method, even very strong binders might be represented in the output pool.

We do not expect to observe every relevant binding event between the polypeptide phage display library and the immobilised small molecule, rather we hope to observe

some interactions of interest and classify them, leading to hypothesis generation for further study.

3.2.4 Sequencing Results

All sequence output files were inspected visually using sequence viewer software to determine if the sequencing had been successful. Sequences are available in E-Appendix 8.4.1.

DNA quality is the single most important factor in obtaining high quality sequence data.[22] No output data is often caused by too low a concentration of the input DNA, which can also lead to poor data when the quantity of DNA is higher. Poor data can also be caused by conflicting sequences when the DNA under assay is not one single purified sequence, or more than one annealing site for the primer. Dye 'blobs', which are often seen at the very start of the sequence or within the first 150 base pairs are caused by an incorrect concentration of salts and alcohol in the precipitation step, or a too dilute concentration of *Taq* polymerase used in the amplification process.[22]

The returned sequences were aligned using the MegaBLAST alignment tool against the coding region for the T7Select® 10-3 vector. (results: E-Appendix 8.4.2) Any sequences shown to be entirely vector with no insert were recorded to ensure that these did not interfere with later bioinformatics.

Table 3.2.4: Sequencing results from simvastatin biopanning round 3

Empty Vector	Failed Sequencing	Further Bioinformatic analysis
18	90	52

All returned sequences were then entered into the DNAStar Lasergene® software SeqMan Pro. This program aligned any highly similar sections of clone sequence to identify any clones which appear more than once, trimmed all low quality ends and

removed any vector-related sequences from the result.

The output from this alignment is 31 consensus sequences from these contiguous regions, or 'contigs'. For parameters and full information, see E-Appendix 8.4.3.

The 60 control sequences were used to compare against the positive (drug immobilised) sequences. If there was a consensus between a control and a positive clone sequence the associated clone was considered to be non-selectively or non-specifically bound to the surface.

From the initial biopan, 120 clones were sequenced from the positive wells, 90 picks showed no information on sequencing, and 18 showed vector only. After contiguous regions were analysed, 31 contigs were observed, 23 of which were found to have no similarity to those sequences from the control wells within the parameters of the alignment.

The remaining contigs were then searched against the RefSeq protein sequence database, using the BLASTX algorithm, where a nucleotide sequence is used to query a protein database in all six frames. The Reference Sequence (RefSeq) collection is a comprehensive, integrated, non-redundant, well-annotated set of sequences, including genomic DNA, transcripts, and proteins.[23] (BLAST parameters available in E-Appendix 8.4.4)

The top annotated protein BLAST hit (not including 'predicted' or 'analogue' protein hits) is listed below. The full BLAST output data is available in E-Appendix 8.4.4.

```
Database: Homo sapiens RefSeq protein

Query= Simvastatin Contig 3
Length=302
>ref|NP_001002.1| 40S ribosomal protein S7 [Homo sapiens]
Length=194
 Score =  171 bits (432),  Expect = 2e-43
 Identities = 85/85 (100%), Positives = 85/85 (100%), Gaps = 0/85 (0%)
 Frame = +3

Query  3    TKNKQKRPRSRTLTAVHDAILEDLVFPSEIVGKRIRVKLDGSRLIKVHLDKAQQNNVEHK  182
            TKNKQKRPRSRTLTAVHDAILEDLVFPSEIVGKRIRVKLDGSRLIKVHLDKAQQNNVEHK
Sbjct  110  TKNKQKRPRSRTLTAVHDAILEDLVFPSEIVGKRIRVKLDGSRLIKVHLDKAQQNNVEHK  169
```

```
Query  183  VETFSGVYKKLTGKDVNFEFPEFQL  257
            VETFSGVYKKLTGKDVNFEFPEFQL
Sbjct  170  VETFSGVYKKLTGKDVNFEFPEFQL  194

Query= Simvastatin Contig 7
Length=218
>ref|NP_001158312.1| LYR motif-containing protein 4 isoform 2 [Homo sapiens]
Length=130
 Score = 89.4 bits (220),  Expect = 7e-19
 Identities = 43/61 (71%), Positives = 46/61 (76%), Gaps = 0/61 (0%)
 Frame = -3

Query  213  EMKSHSVTQPGVQWCNLS*LQPPPPRFKQFFCLSLPSSWDYRHAPSCMANFCIFSRDGIS  34
            +M SHSV Q GV W +LS LQP PP FKQF CLSLPSSWDYR  P  +ANFCI SRD IS
Sbjct  69   QMDSHSVAQAGVHWNDLSSLQPLPPWFKQFSCLSLPSSWDYRRTPPRLANFCILSRDVIS  128

Query  33   L  31
            L
Sbjct  129  L  129

Query= Simvastatin Contig 14
Length=387
>ref|NP_001158011.1| disrupted in schizophrenia 1 protein isoform c [Homo sapiens]
Length=755
 Score = 95.1 bits (235),  Expect = 1e-20
 Identities = 51/69 (74%), Positives = 53/69 (77%), Gaps = 1/69 (1%)
 Frame = -2

Query  326  SLQPPPPEFTPFSCLSLPSS*DYRRPPACLALFFVFFVEIGFRHVAQAGLGFLNSSDPPA  147
            SLQP PPEF  FSCLSL SS DYR PP CLA F VF VE+GF HV Q GL  L SSDPP+
Sbjct  678  SLQPLPPEFKQFSCLSLRSSWDYRCPPPCLANF-VFLVEMGFYHVDQTGLKLLTSSDPPS  736

Query  146  SASQSAGIT  120
            SASQSAGIT
Sbjct  737  SASQSAGIT  745

Query= Simvastatin Contig 15
Length=369
>ref|NP_872601.1| histone demethylase UTY isoform 1 [Homo sapiens]
Length=1079
 Score = 38.9 bits (89),  Expect(3) = 2e-11
 Identities = 17/19 (90%), Positives = 18/19 (95%), Gaps = 0/19 (0%)
 Frame = +1

Query  64    VETGFHHVGQAALKLLTSG  120
            VETGFHHVGQA L+LLTSG
Sbjct  1041  VETGFHHVGQACLELLTSG  1059

 Score = 37.0 bits (84),  Expect(3) = 2e-11
 Identities = 14/20 (70%), Positives = 17/20 (85%), Gaps = 0/20 (0%)
 Frame = +2

Query  2     LPSSWDYRRPPSRPSNFCIF  61
            LP+SW+YR  PS P+NFCIF
Sbjct  1021  LPNSWNYRHLPSCPTNFCIF  1040

 Score = 27.7 bits (60),  Expect(3) = 2e-11
 Identities = 13/16 (82%), Positives = 13/16 (82%), Gaps = 0/16 (0%)
 Frame = +3

Query  132   SASQSGGITGVSHCTR  179
            SASQS GITGVSH  R
Sbjct  1064  SASQSAGITGVSHHAR  1079

Query= Simvastatin Contig 23
Length=164
>ref|NP_001158312.1| LYR motif-containing protein 4 isoform 2 [Homo sapiens]
Length=130
 Score = 80.1 bits (196),  Expect = 4e-16
 Identities = 37/53 (70%), Positives = 40/53 (76%), Gaps = 0/53 (0%)
 Frame = -1

Query  164  EMKSHSVTQPGVQWCNLS*LQPPPPRFKQFFCLSLPSSWDYRHAPSCMANFCI  6
            +M SHSV Q GV W +LS LQP PP FKQF CLSLPSSWDYR  P  +ANFCI
Sbjct  69   QMDSHSVAQAGVHWNDLSSLQPLPPWFKQFSCLSLPSSWDYRRTPPRLANFCI  121
```

```
Query= Simvastatin Contig 26
Length=460
>ref|NP_000423.2| myosin regulatory light chain 2, ventricular/cardiac muscle isoform
[Homo sapiens]
Length=166
 Score =  151 bits (382),  Expect = 1e-37
 Identities = 73/73 (100%), Positives = 73/73 (100%), Gaps = 0/73 (0%)
 Frame = +2

Query  2    DPEETILNAFKVFDPEGKGVLKADYVREMLTTQAERFSKEEVDQMFAAFPPDVTGNLDYK  181
            DPEETILNAFKVFDPEGKGVLKADYVREMLTTQAERFSKEEVDQMFAAFPPDVTGNLDYK
Sbjct  94   DPEETILNAFKVFDPEGKGVLKADYVREMLTTQAERFSKEEVDQMFAAFPPDVTGNLDYK  153

Query  182  NLVHIITHGEEKD  220
            NLVHIITHGEEKD
Sbjct  154  NLVHIITHGEEKD  166

Query= Simvastatin Contig 28
Length=155
>ref|NP_963998.2| thromboxane A2 receptor isoform beta [Homo sapiens]
Length=407
 Score = 35.8 bits (81),  Expect = 0.009
 Identities = 16/22 (73%), Positives = 19/22 (87%), Gaps = 0/22 (0%)
 Frame = -1

Query  155  ASASQSAGITGVSHCA*PIFLY  90
            ASAS++AGITGVSHCA P  L+
Sbjct  359  ASASRAAGITGVSHCARPCMLF  380

Query= Simvastatin Contig 29
Length=175
>ref|NP_853516.1| gap junction gamma-3 protein [Homo sapiens]
Length=279
 Score = 52.4 bits (124),  Expect = 9e-08
 Identities = 25/25 (100%), Positives = 25/25 (100%), Gaps = 0/25 (0%)
 Frame = -1

Query  76  MCGRFLRRLLAEESRRSTPVGRLLL  2
           MCGRFLRRLLAEESRRSTPVGRLLL
Sbjct  1   MCGRFLRRLLAEESRRSTPVGRLLL  25

Query= Simvastatin Contig 31
Length=470
>ref|NP_060190.2| signal-transducing adaptor protein 2 isoform 1 [Homo sapiens]
Length=449
 Score = 60.1 bits (144),  Expect = 4e-10
 Identities = 31/40 (78%), Positives = 32/40 (80%), Gaps = 1/40 (2%)
 Frame = +2

Query  2    ETGFHPVCQAGLELLTSGSPPTSSSQSAGWITGVSHHIRP  121
            E GFH V QAGLELLTS  PPTS+SQSAG ITGVSHH  P
Sbjct  359  EKGFHHVAQAGLELLTSSDPPTSASQSAG-ITGVSHHTWP  397
```

Some of the contigs did not give alignments with annotated proteins. This can be due to: the sequence arising from a non-coding region of mRNA when the cDNA library was produced; the sequence coding for a protein which has not yet been annotated, or annotation which has not yet been curated and entered into the RefSeq database. Other issues such as frame slippage, mis-transcribed nucleotides and missing data due to poor sequences are, within reason, managed by the search tool.

Sequences outside the annotated protein pool may give more information when that pool has been further expanded, and so these contigs with their annotated chemistry and small molecule should be carefully recorded and stored for periodic repeat searching using the BLAST algorithm.

3.2.5 Bioinformatics

Where contigs appear in more than one sequence or multiple contigs hit the same BLAST result, they represent multiple copies of the same clone or heavily overlapping sections of polypeptide being expressed on the phage head group. This could be due to strong specific binding of that expressed polypeptide, strong non-specific binding of the same, or due to higher amplification rates of these clones over other clones - known as competitive amplification.[24]

By carrying out further bioinformatics, several hits of interest appeared. Phosphorylation of myosin regulatory light chain (contig 26) has been observed to be attenuated by simvastatin,[25] expression of thromboxane A_2 has been reported to be reduced by simvastatin (thromboxane A_2 receptor observed in contig 28),[26, 27] and gap junction proteins (gap junction gamma-3 observed in contig 29) have been seen to be down-regulated by statins leading to a potentially antiproliferative effect.[28]

Further discussion and conclusions drawn from these hits and those from further biopanning experiments are collected in Chapter 6.

3.2.6 Determining Relative Binding Strength

While every hit found in positive wells and not in the control wells must be considered to be of interest at this stage, ranking of the binding strength between the different clones is a good way of focusing interest. Ranking binding strength was attempted on a sample of clones from the simvastatin biopan by a competitive elution method. (See Appendix 8.4.5 for clone identities). This method compares the quantity of phage clones retained by small molecule-immobilised wells when eluted using a solution of the small molecule against elution using a solvent solution, reported as a ratio of drug:solvent elution from the surface.

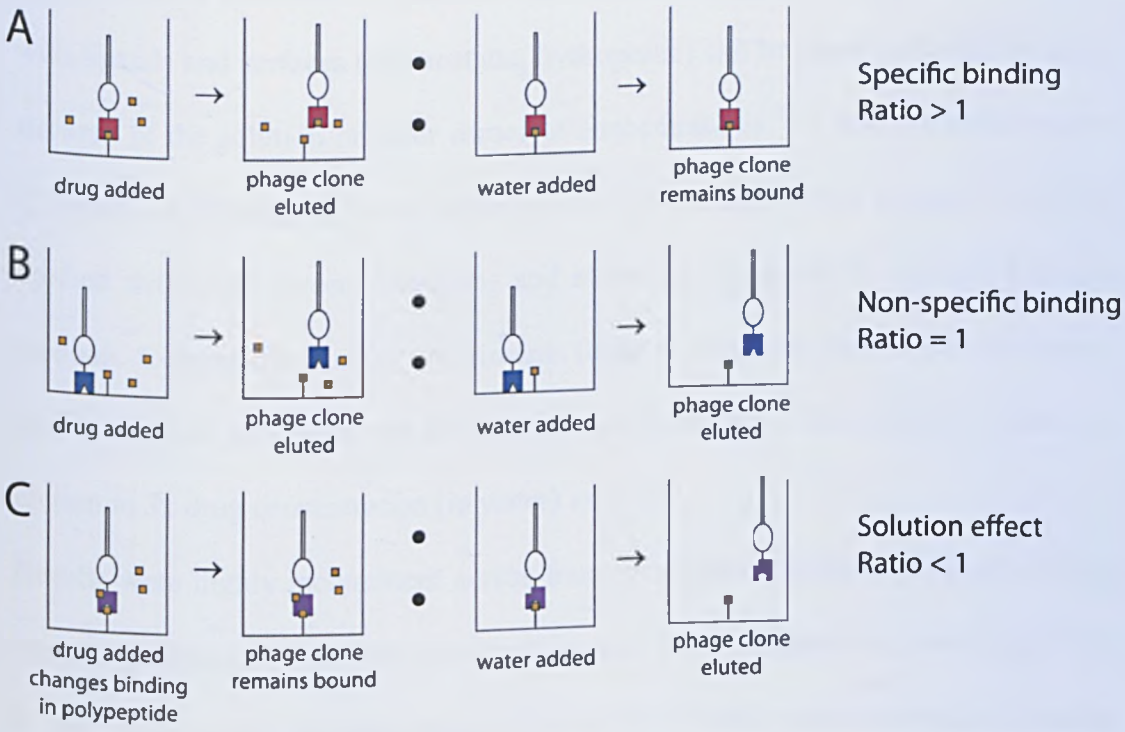


Figure 3.2.5: Competitive Elution results analysed

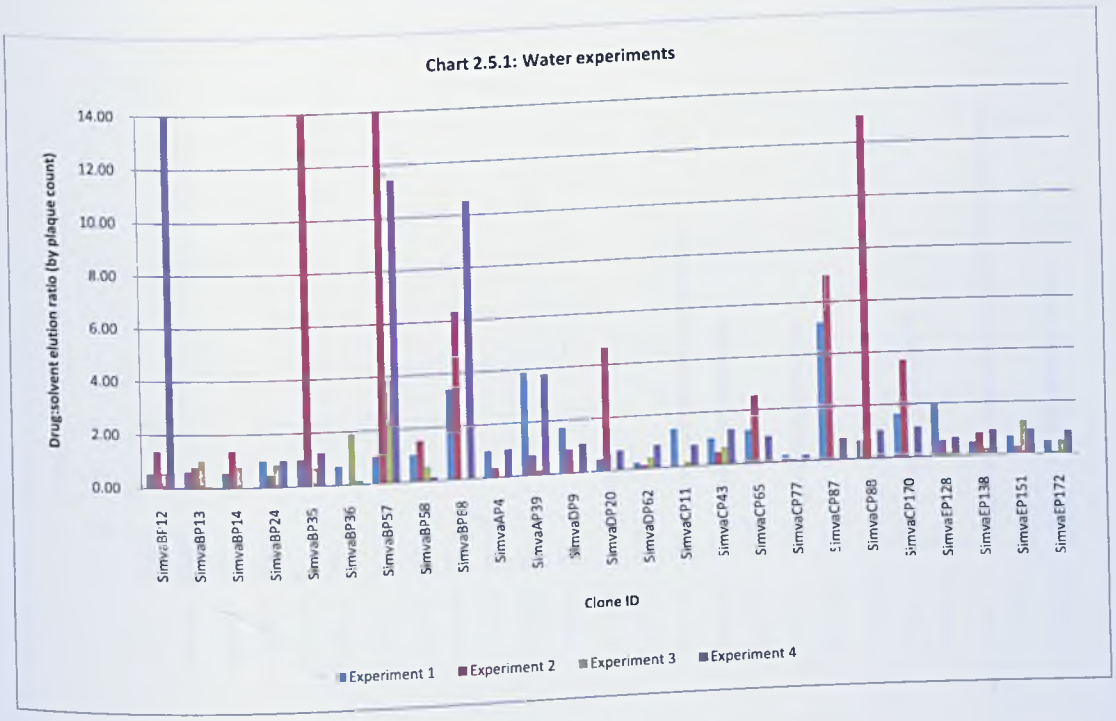
A ratio of drug:solvent=1 after elution implies that the polypeptide is non-selectively bound (B in Figure 3.2.5) – it is eluted from the plate surface as easily with water as it is with the drug solution. If the plaque count values (not the ratio) are very low, this could alternatively imply that the protein is bound so strongly that it is not eluted with either solution-phase drug or by water.

A high ratio, greater than one, implies some specificity of binding (A in figure 3.2.5). The bound polypeptide is eluted by solution phase drug molecules competing with the immobilised drug molecule, but not eluted by water. These are the phage clones of highest interest.

A lower ratio, less than one, implies some interference with the elution by the drug molecule (C in figure 3.2.5). This could be due to a change in the ionic strength of the solution increasing the strength of binding – whether that binding is to the drug molecule or to the plate surface – or some interaction between the drug molecule in solution and other parts of the polypeptide increasing binding.

The folding of proteins, including phage displayed peptides, and their interactions with ligands and surfaces (electrostatic, hydrophilic) can be greatly affected by ionic strength of the solution or other osmolyte concentrations.[29-31] Initial binding assays (3 rounds of biopanning) were carried out in LB medium (ionic strength ~0.15M), washed with TBS Tween® solutions and eluted by log phase *E. coli* stock in LB medium. Competitive elution experiments (drug *vs.* negative control) should mimic this binding to determine the strength of initial binding. The protocol called for elution in 3x drug concentration (in water) *vs.* water.

Results were highly inconsistent across four competitive elution experiments using water (see Chart 2.5.1, for full data see Appendix 8.4.5). It should be noted that if the *E. coli* lawn showed too many phage to count these plates were recorded as showing 10,000 or 20,000 plaques by visual estimate. 10,000 plaques were recorded when the edges of plaques (traces of *E. coli* lawn remaining) were visible, and 20,000 was recorded when no edges were visible.



To determine the effect of ionic strength on this competitive elution step, the experiment was repeated with 3x drug concentration (in TBS – ionic strength ~0.17M) *vs.* TBS. Comparisons between water elutions and TBS elutions run simultaneously can be seen in Chart 2.5.3. There was little solubility difference between the two solutions. Consistency continued to be a serious issue in each repeat of these competitive elutions (Chart 2.5.2).

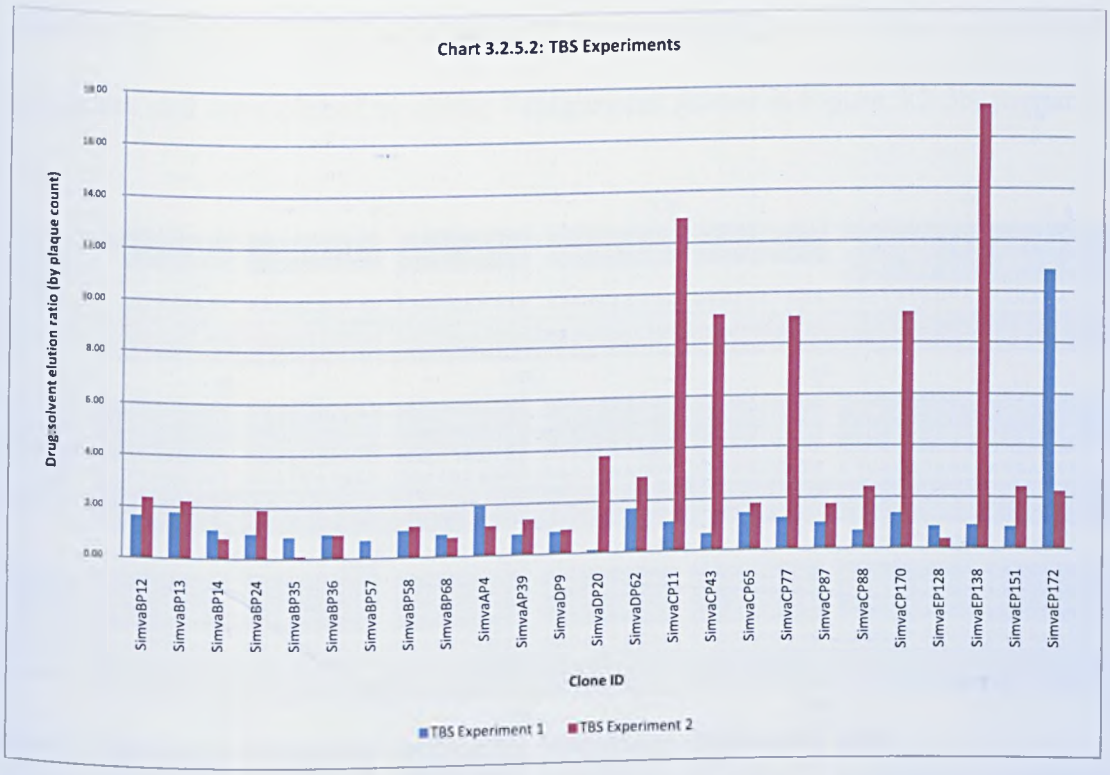


Chart 3.2.5.2: TBS Experiments
Drug:solvent elution ratio (by plaque count)
SimvaBP12
SimvaBP13
SimvaBP14
SimvaBP24
SimvaBP35
SimvaBP36
SimvaBP57
SimvaBP58
SimvaBP68
SimvaAP4
SimvaAP39
SimvaDP9
SimvaDP20
SimvaDP62
SimvaCP11
SimvaCP43
SimvaCP65
SimvaCP77
SimvaCP87
SimvaCP88
SimvaCP170
SimvaEP128
SimvaEP138
SimvaEP151
SimvaEP172
Clone ID
TBS Experiment 1
TBS Experiment 2

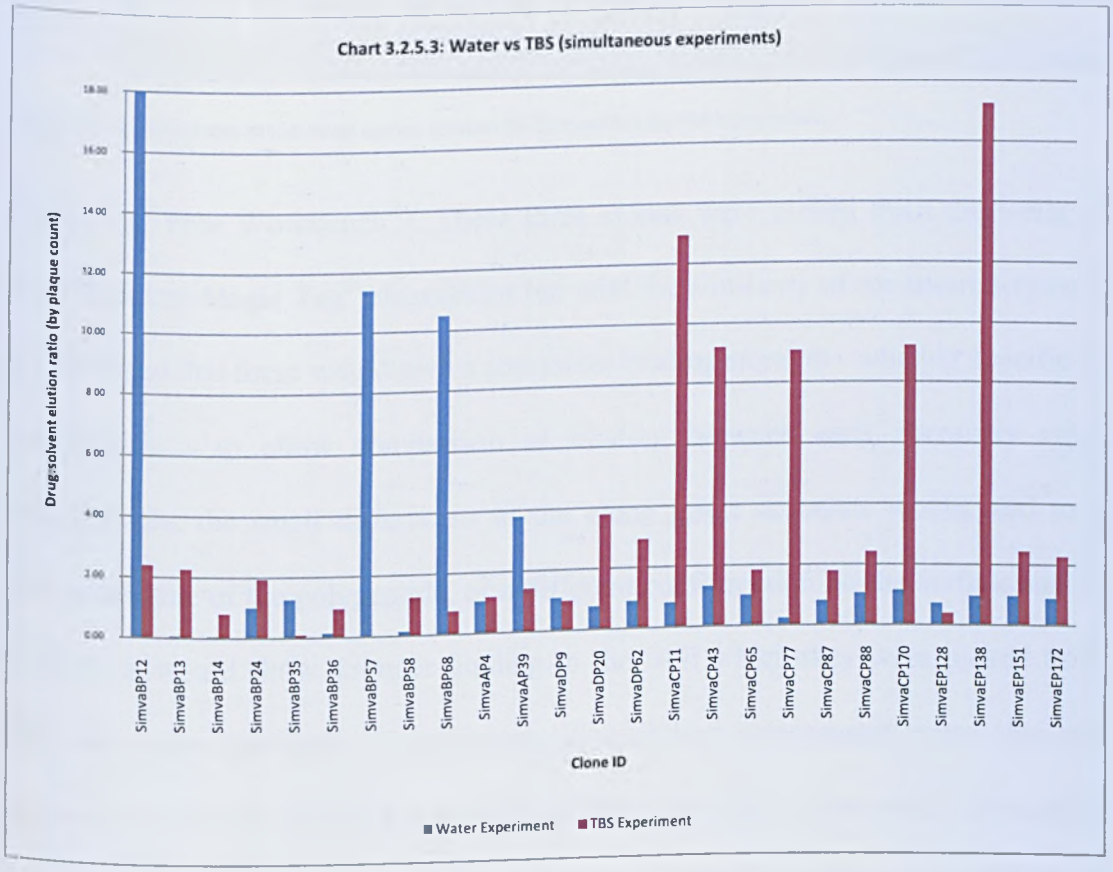


Chart 3.2.5.3: Water vs TBS (simultaneous experiments)
Drug:solvent elution ratio (by plaque count)
SimvaBP12
SimvaBP13
SimvaBP14
SimvaBP24
SimvaBP35
SimvaBP36
SimvaBP57
SimvaBP58
SimvaBP68
SimvaAP4
SimvaAP39
SimvaDP9
SimvaDP20
SimvaDP62
SimvaCP11
SimvaCP43
SimvaCP65
SimvaCP77
SimvaCP87
SimvaCP88
SimvaCP170
SimvaEP128
SimvaEP138
SimvaEP151
SimvaEP172
Clone ID
Water Experiment
TBS Experiment

Three of the clones were determined to be displaying polypeptides from very similar sequences and were placed in contig 7 (alignment shown in Figure 3.2.5b, prepared

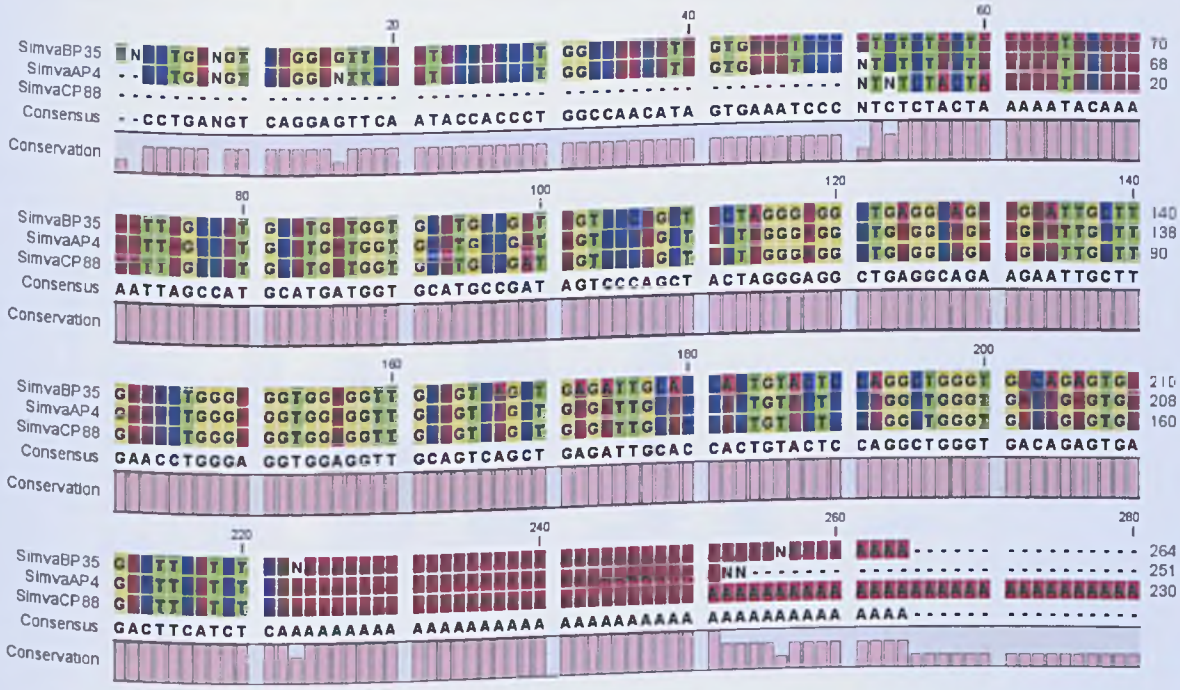
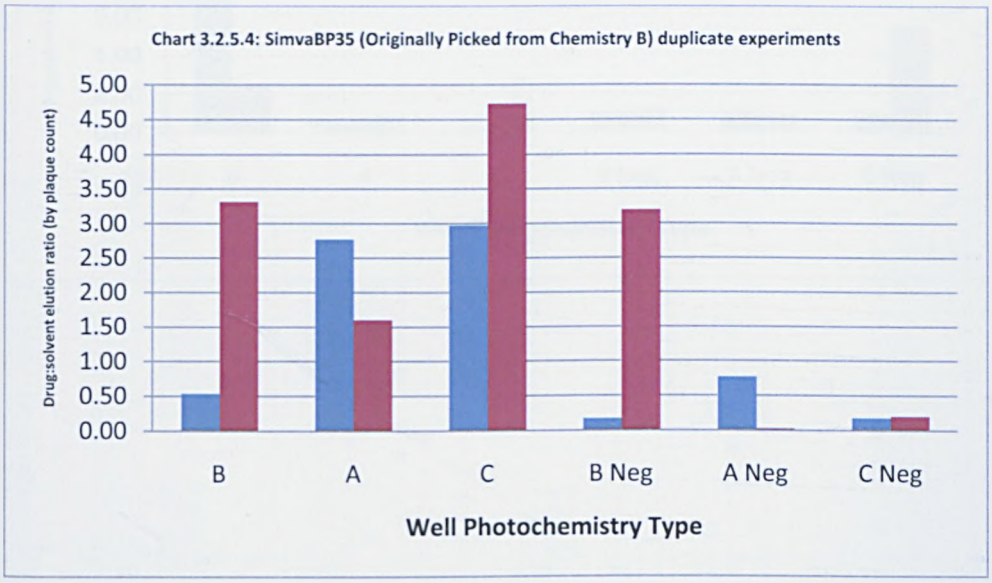


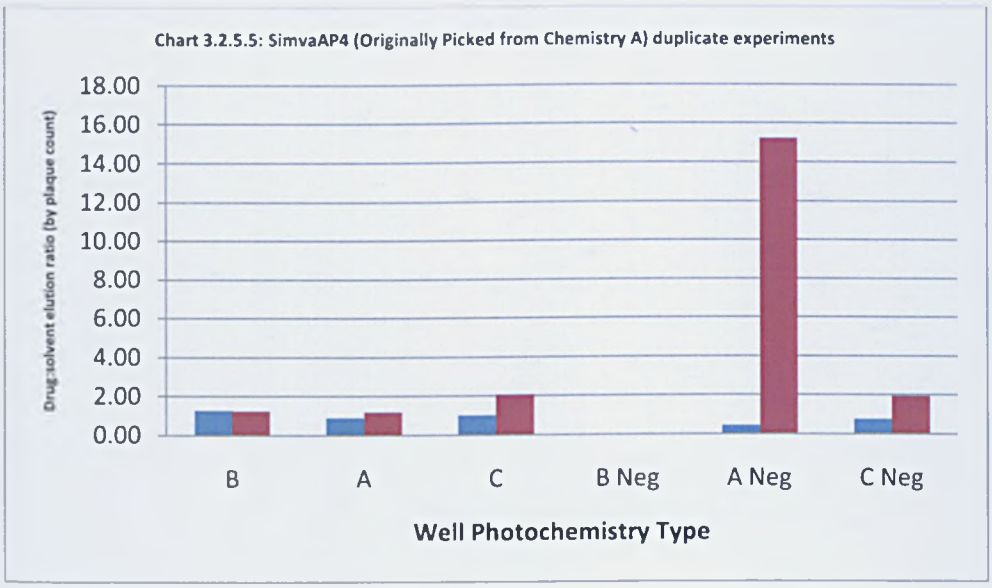
Figure 3.2.5b: Alignment of the three clones chosen for competitive elution comparison

using CLC Free Workbench®). These three clones were picked from the wells of three different Magic Tag® chemistries but with the similarity of the insert sequence it was hoped that these would give a consistent binding strength - whether specific or non-specific - to allow comparison of binding between each chemistry type. Alternatively, the small differences in the clone insert sequence would lead to a different frame of the polypeptide, or a different conformation on the surface and so each clone would show stronger binding to the well which they were picked from than any other chemistry. These competitive elution experiments were also run against control wells prepared as in the biopanning experiment, where no simvastatin had been immobilised. All of these experiments were carried out in duplicate to check for consistency of data, which had previously been a problem. The results for each clone are shown below. These continue to be startlingly inconsistent, although

interestingly the binding for the control wells of the chemistry from which each clone was picked was higher than the other control wells, and in one case higher than the wells containing immobilised simvastatin, but only in one of the two duplicate experiments. This suggests that this is not a reproducible result.



Charts 3.2.5.4-6: Three clones with very similar sequences were compared using their competitive elution scores (drug:solvent elution ratio determined by plaque count). These clones were compared across three different photochemistry types and the well negatives (no simvastatin immobilised), in duplicate experiments. As can be seen, there was very little consistency even between duplicates of one clone.



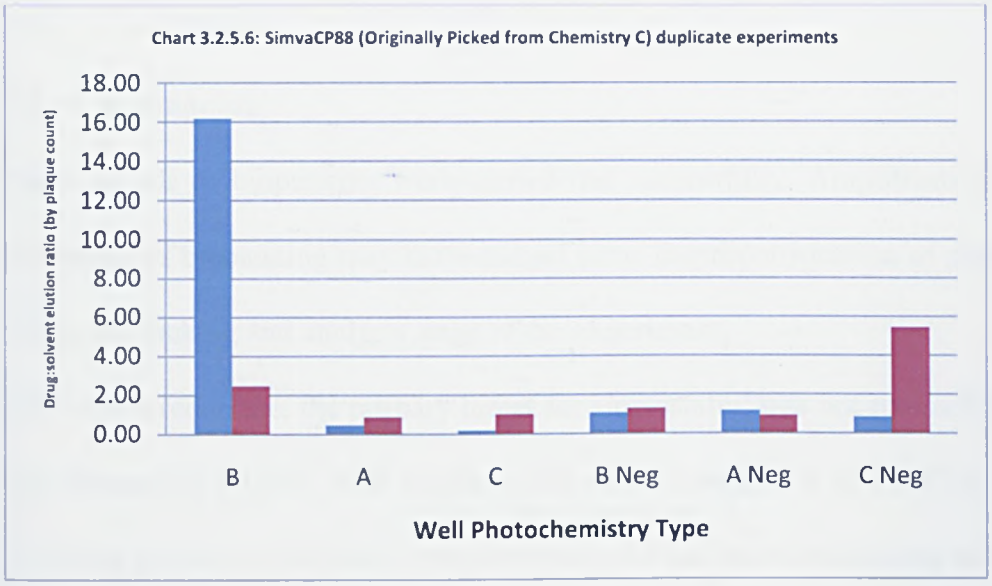


Chart 3.2.5.6: SimvaCP88 (Originally Picked from Chemistry C) duplicate experiments
Drug:solvent elution ratio (by plaque count)
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
0.00
B
A
C
B Neg
A Neg
C Neg
Well Photochemistry Type

3.3 Conclusions

3.3.1 Biopanning

Three rounds of biopanning were carried out successfully. Amplification after the last round of biopanning may have caused some disproportionation of phage clones before the picking and analysis stage of the experiment.

HMG-CoA reductase, the primary target for simvastatin, was not returned as a hit in our biopanning process. This could be due to the mRNA for HMG-CoA reductase not being present in the source vascular tissue, or due to overwhelming of the signal with other more prolific polypeptide binders in the phage library.

After initial bioinformatics, hits of interest include the ribosomal protein S7, disrupted in schizophrenia 1 (DISC1) protein, histone demethylase UTY, LYR motif-containing protein 4, signal-transducing adaptor protein 2. Hits of specific interest include myosin regulatory light chain 2, thromboxane A2 receptor and gap junction gamma-3. Further work will consist of validation for these interactions and analysis into their potential implications.

3.3.2 Further analysis of the biopanning outputs

From the TBS results plotted against the water elution data set (Chart 3.2.5.3), it can be seen that there is some strong variability between the water (blue) and the TBS (red) results. This could be due to differences in the interactions between the peptide and the small molecule, or in the folding and secondary structure of the peptide.

There was a lot of inconsistency from experiment to experiment even when the same solvent was used, which led to concern as to the validity of these experimental results.

There is also variability between the clones which contain very similar genetic material. This could be due to changes in folding, or the translation frame of the peptides.

The competitive elution technique has not produced any firm answers and therefore was determined not to be a valid technique at this point. Competitive elution was not used with further biopanning experiments. The clone results from the biopan are all of some interest, regardless of the strength of binding.

3.4 Methods and Materials

3.4.1 General Experimental

This experimental work was carried out under the HSE's GMO (contained use) guidelines at Tangent Reprofilling's labs at HRI, Wellesbourne.

Water used in cDNA and Phage Library preparation was MilliQ® (Millipore) filtered sterile water prepared on site.

Tris buffered saline (TBS, 0.05 M Tris, 0.15 M NaCl, pH 7.6 at 25 °C) with Tween 20 surfactant was prepared using TBS tablets (Sigma Aldrich, Cat# T5030). 13 tablets were used to prepare 195 mL MilliQ water. Two concentrations of Tween 20 were used, 0.5% and 0.1%, 975 µl and 195 µl respectively.

For preparation of agarose gels, it should be noted that the labs became ethidium bromide free, replacing ethidium bromide use with the non-toxic GelRed® (10,000X in water, Biotium, Cambridge Biosciences, Cat# BT41003). GelRed® is added to molten agarose solution (5 µl per 100 mL). GelRed® is not added to the running buffer. All other processes for preparation of agarose gel and loading of wells is identical to that previously described. GelRed® is visualised under UV light in an identical manner to ethidium bromide.

The amount of water in the sequencing PCR master mix was adjusted to account for 1 µl sequencing buffer.

Sequencing mastermix (per sample): 2 µl Big Dye® (Applied Biosystems, Cat# 4337454), 1 µl 5X sequencing buffer (Applied Biosystems, Cat# 4336697), 0.4 µl T7Select Up Primer, water to make up to 9 µl. 1 µl cleaned DNA solution added before submission to the HRI sequencing service.

3.4.2 Simvastatin Immobilisation

Simvastatin (5 mg, Sigma-Aldrich, Cat# S6196, ≥97% (HPLC) solid) was dissolved in DMSO (1 mL) as a working stock solution (5 mg/mL, 0.012 M, stored at 4 °C, allowed to return to r.t. before use). A 100-fold dilution of the stock drug solution was prepared (4 mL, 40 µl in 3.96 mL water, 0.12 mM), and this solution was added to three wells each of PCA, PCB, PCC, PCD, PCE (200 µl, 10 ng, 0.024 nmol per well). Water (200 µl per well) was added to three wells each of PCA, PCB, PCC, PCD, PCE to form the negative controls. Due to the sensitivity of the PCE photochemistry, preparation of these wells was carried out in a dark room before irradiation. All wells were irradiated for 10 minutes under a handheld 254nm UV lamp (220 W). All wells, drugged and control, were washed with water six times and then the plate was stored at 4 °C until use.

3.4.3 Biopanning

Phage libraries were amplified as a combined library (50 µl library 1 and 50 µl library 2) in 5 mL log phase *E. coli* solution at $O.D._{600} \approx 0.5$, with agitation for 3 h at 37 °C. The lysate was spun down and the clarified lysate was decanted into fresh tubes. Amplified human vascular phage library (200 µl per well) was added to round

1 drug wells and control wells and agitated for 45 minutes. Excess phage solution was removed and wells washed with 0.5% Tween® in TBS six times with 2 minutes agitation for each wash. Log phase *E. coli* was prepared:- stock *E. coli* strain BLT5615 (1 mL, OD_{600} ~2.0 in LB medium with ampicillin was added to LB medium (100 mL) with ampicillin (1 µl/mL) for 2 h at 37°C. At 1.5 h IPTG (500 µl, 100 mM) was added. 200 µl of the log phase *E. coli* was added per well, at OD_{600} ~0.5 and agitated for 10 minutes to elute the remaining bound phage.

The first round well contents were amplified in 5 mL log phase *E. coli* in labelled 50mL tubes. The tubes were incubated for 3 h at 37 °C with shaking during which time the solution cleared as the *E. coli* were lysed. 2 mL aliquots of amplified phage from each tube were separated from any *E. coli* detritus by centrifuge and the clarified lysates decanted into new labelled tubes.

Well-specific products from the first round of biopanning (200 µl per well) were then introduced to the next row of immobilized drug and water wells and the process repeated. On the third round of biopanning the well contents from the elution step were sampled and diluted to 10^{-5} and 10^{-7}, then mixed with stock *E. coli* culture and poured onto agarose/agar assay plates directly to avoid any disproportionation of the phage present. (For plating method, see Chapter 2.3.5)

Amplification was carried out and the amplified phage solutions were allowed to infect a fresh *E. coli* lawn at dilutions of 10^{-3}, 10^{-5} and 10^{-7}, giving adequate plaques for picking. From this plating, 24 plaques from positives wells (PCB, PCA, PCD, PCC and PCE, a total of 120 clones) and 12 plaques from water controls (a total of 60 clones) were picked for PCR and sequencing (PCR, agarose gel electrophoresis and sequencing methods described above and in the general experimental for chapter 2). Picked clones and PCR product were both stored at 4°C initially, then were

archived as glycerol stocks (1:9 80% glycerol solution to phage solution) and stored at -80°C.

3.4.4 Sequencing

The PCR product was introduced into a 96-well Millipore filter plate (Cat# LSKMPCR10) and filtered under vacuum. Water (100 µl) was added and the plate was agitated for 5 mins. Big Dye™ sequencing master mix was prepared as described in the general experimental and 9 µl was added to each well, followed by 1 µl DNA solution. This plate was submitted to the HRI sequencing service.

3.4.5 Determining Relative Binding Strengths

The competitive elution experiment is carried out as follows: For every clone under study, two wells of the appropriate Magic Tag strip were laid out.

The same concentration of stock drug solution (200 µl of 100 fold dilution in water from 5 mg/mL drug in DMSO) was added to both wells, and immobilised under 254nm UV light for 10 minutes. Wells were emptied and washed six times with water before use. Plates not used directly were stored at 4°C.

Newly amplified clones (200 µl per well, specific to the chemistry they were originally found on) were added to the wells and agitated for 30 minutes to allow binding to occur, then the plate was washed with Tween® in TBS (200 µl, 0.5%), six times with 2 minutes agitation each to remove any unbound phage.

Drug solution was added at 3x concentration (33-fold dilution, 200 µl per well) in one well per clone and Milli-Q® water in the duplicate well, and the plate was agitated for 30 minutes.

Well contents were removed into labelled tubes with no amplification. Well products were titred by plaque count assay to output a ratio of pfu/mL with which to compare clones. For plaque count assay, plates were put on to warm, and agarose heated to 50°C. Well contents were plated out at 10^{-2} (1 µl in 100 µl water) in IPTG (100 µl), *E. coli* stock solution (100 µl) and hot agarose (3 mL). Incubated for 3 h at 37°C. Plaques counted and numbers recorded.

3.5 References

1. J. Simons, *Fortune*, 2003, **147**.
2. R. H. Knopp, *N. Engl. J. Med.*, 1999, **341**, 498-511.
3. P. T. S. Ma, G. Gil, T. C. Sudhof, D. W. Bilheimer, J. L. Goldstein and M. S. Brown, *Proc. Natl. Acad. Sci. U. S. A.*, 1986, **83**, 8370-8374.
4. M. Jasiska, J. Owczarek and D. Orszulak-Michalak, *Pharmacol. Rep.*, 2007, **59**, 483-499.
5. R. P. Mason, M. F. Walter, C. A. Day and R. F. Jacob, *Am. J. Cardiol.*, 2005, **96**, 11F-23F.
6. T. Sakaeda, H. Fujino, C. Komoto, M. Kakumoto, J. S. Jin, K. Iwaki, K. Nishiguchi, T. Nakamura, N. Okamura and K. Okumura, *Pharm. Res.*, 2006, **23**, 506-512.
7. J. Armitage, *Lancet*, 2007, **370**, 1781-1790.
8. M. R. Ehrenstein, E. C. Jury and C. Mauri, *N. Engl. J. Med.*, 2005, **352**, 73-75.
9. G. Weitz-Schmidt, K. Welzenbach, V. Brinkmann, T. Kamata, J. Kallen, C. Bruns, S. Cottens, Y. Takada and U. Hommel, *Nat. Med.*, 2001, **7**, 687-692.

10. L. X. Zeng, H. S. Xu, T. L. Chew, E. Eng, M. M. Sadeghi, S. Adler, Y. S. Kanwar and F. R. Danesh, *Faseb J.*, 2005, **19**, 1845-+.

11. E. S. Istvan and J. Deisenhofer, *Science*, 2001, **292**, 1160-1164.

12. R. W. Wang, P. H. Kari, A. Y. H. Lu, P. E. Thomas, F. P. Guengerich and K. P. Vyas, *Arch. Biochem. Biophys.*, 1991, **290**, 355-361.

13. M. Igel, T. Sudhop and K. vonBergmann, *Eur. J. Clin. Pharmacol.*, 2001, **57**, 357-364.

14. T. Prueksaritanont, R. Subramanian, X. J. Fang, B. Ma, Y. Qiu, J. H. Lin, P. G. Pearson and T. A. Baillie, *Drug Metab. Dispos.*, 2002, **30**, 505-512.

15. *Pharmacogenomics*, 2010, **11**, 295-296.

16. I. B. Skottheim, A. Gedde-Dahl, S. Hejazifar, K. Hoel and A. Asbery, *Eur. J. Pharm. Sci.*, 2008, **33**, 317-325.

17. J. M. Backes, P. A. Howard, J. F. Ruisinger and P. M. Moriarty, *Ann. Pharmacother.*, 2009, **43**, 2012-2020.

18. S. J. Dilly, M. J. Bell, A. J. Clark, A. Marsh, R. M. Napier, M. J. Sergeant, A. J. Thompson and P. C. Taylor, *Chem. Commun.*, 2007, 2808-2810.

19. S. J. Dilly, M. P. Beecham, S. P. Brown, J. M. Griffin, A. J. Clark, C. D. Griffin, J. Marshall, R. M. Napier, P. C. Taylor and A. Marsh, *Langmuir*, 2006, **22**, 8144-8150.

20. T. Wind, B. Stausbol-Gron, S. Kjaer, L. Kahns, K. H. Jensen and B. F. C. Clark, *J. Immunol. Methods*, 1997, **209**, 75-83.

21. Novagen, 70014 T7Select® Packaging Kit http://www.merckbiosciences.co.uk/product/70014, Accessed 15/11/2010.

22. J. Kieleczawa, *DNA sequencing: optimizing the process and analysis*, Jones and Bartlett Publishers, 2005.

23. K. D. Pruitt, T. Tatusova and D. R. Maglott, *Nucleic Acids Res.*, 2007, **35**, D61-D65.

24. C. F. Barbas, D. R. Burton, J. K. Scott and G. J. Silverman, *Phage Display: A Laboratory Manual*, CSHL Press, 2004.

25. J. R. Jacobson, S. M. Dudek, K. G. Birukov, S. Q. Ye, D. N. Grigoryev, R. E. Girgis and J. G. N. Garcia, *Am. J. Respir. Cell Mol. Biol.*, 2004, **30**, 662-670.

26. A. Notarbartolo, G. Davi, M. Averna, C. M. Barbagallo, A. Ganci, C. Giammarresi, F. P. Laplaca and C. Patrono, *Arterioscler. Thromb. Vasc. Biol.*, 1995, **15**, 247-251.

27. C. J. Vaughan, M. B. Murphy and B. M. Buckley, *Lancet*, 1996, **348**, 1079-1082.

28. L. H. Wang, J. Z. Chen, Y. L. Sun, F. R. Zhang, J. H. Zhu, S. J. Hu and D. H. Wang, *Am. J. Hypertens.*, 2005, **18**, 1146-1153.

29. T. Beringhelli, I. Eberini, M. Galliano, A. Pedoto, M. Perduca, A. Sportiello, E. Fontana, H. L. Monaco and E. Gianazza, *Biochemistry*, 2002, **41**, 15415-15422.

30. T. Kodadek and K. Bachhawat-Sikder, *Mol. BioSyst.*, 2006, **2**, 25-35.

31. Y. Katakura, E. T. Lim, S. Tsujii, T. Omasa and K. Suga, *J. Ferment. Bioeng.*, 1998, **85**, 447-450.

Chapter 4 Biopanning Human Vascular Tissue Library against Four Small Molecules with Anti-inflammatory Activity

4.1 Introduction

4.1.1 Summary

This chapter describes the biopanning of four medicinally relevant small molecules against a T7Select® human vascular tissue library. A detailed investigation of the enrichment across three rounds of a selected biopanning experiment was also carried out.

4.1.2 Cyclooxygenase and its Role in Inflammation

Inflammation may be triggered by biochemical, mechanical or microbial damage, leading to the release of signalling molecules, which mediate further inflammation. These molecules, typically including prostaglandins, leukotrienes, prostacyclins and thromboxanes, lead to vasodilation, increased capillary permeability, monocyte accumulation, stimulation of sensory receptors, fever and contraction of smooth muscle.[1] Prostaglandins and thromboxanes are synthesised from phospholipids in the membrane of the endoplasmic reticulum, by the action of the enzyme cyclooxygenase (COX) on arachidonic acid.[2] There are three forms of the COX enzyme: COX-1, which is a housekeeping enzyme involved in gastric protection, COX-2 which is only found in inflammatory cells,[2] and COX-3 which is a transcript variant of COX-1 retaining one more intron.[3]

These inflammatory interactions are required for the stress response of cells, but excessive or prolonged inflammation can lead to chronic diseases such as arthritis and cardiovascular disease.[4]

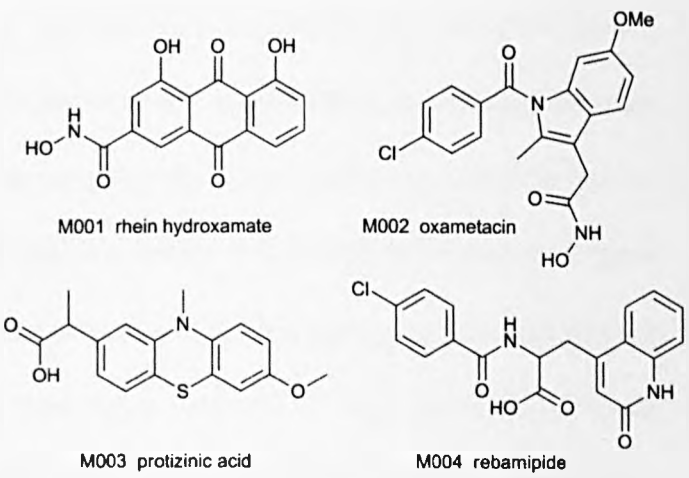
By inhibiting cyclooxygenase, hence blocking production of prostaglandins and thromboxane, an anti-inflammatory action can thus be obtained. Many non-steroidal anti-inflammatory drugs (NSAIDs) are non-selective COX inhibitors. Due to the necessary actions of COX-1, these NSAIDs have adverse effects including gastric intestinal upset, peptic ulcers, allergy and renal impairment.[2] Selective COX-2 inhibitors have been developed, but unfortunately these appear to have other serious side effects.[1]

COX inhibition is not the only route to anti-inflammatory action in non-steroidal drugs. Inhibition of other pro-inflammatory bioactives, such as the interleukin family of polypeptides, and oxidative species such as superoxides are also believed to provide anti-inflammatory therapeutic action.

4.1.3 Selected Small Molecules

Four molecules of commercial interest with known or suspected anti-inflammatory action *in vivo* were selected for biopanning against the human vascular tissue phage library.

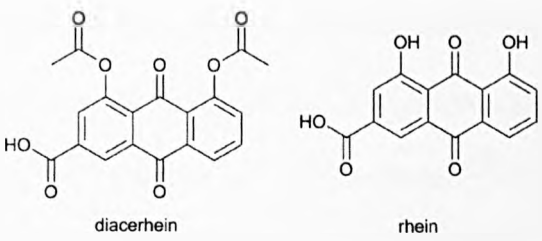
Rhein hydroxamate (M001) is a molecule in the anthraquinone family. Diacerein, which metabolises to rhein, and rhein itself show anti-inflammatory action through inhibition of production of interleukin-1, but unlike common NSAIDs, show stimulation of COX synthesis.[5-10]



Structures 4.1.3.1 - 4: Antiinflammatory molecules under study - rhein hydroxamate (M001), oxametacin (M002), protizinic acid (M003), rebamipide (M004)

The hydroxamate of rhein is currently under study.

Oxametacin(e)[11-13] (M002) and protizinic acid[14] (M003) are known non-steroidal inflammatory compounds, with analgesic and antipyretic action.



Structures 4.1.3.5 - 6: Diacerein and its metabolite rhein

Rebamipide[15-18] (M004) inhibits the production of superoxide species and suppresses myeloperoxidase activity, giving it its anti-inflammatory action.[17] It is also known to scavenge free radicals, providing oxidative protection as in the case of rhein, and has shown effectiveness against presbyacusis - the development of age-related hearing loss.[18]

These four compounds were chosen by the sponsor company, Tangent Reprofiling, for experimental study against the human vascular tissue phage display library.

4.1.4 Investigating Diversity Changes During Biopanning

Carrying out three rounds of biopanning allows amplification of the specific-binding event 'signal' against the background 'noise' of non-specific binding. This relies on non-specific binding being random and therefore less frequently observed than a selective binding event. This is often not the case. It also relies on the amplification rate for every phage clone to be identical given the same conditions, which is known not to be the case. To reduce false positive results it is therefore necessary to have control wells to identify hits which are selective to the background plates and solvent immobilised on the surface rather than being selective for the immobilised small molecule. Accepting these limitations, biopanning is carried out as described in the introduction and providing that the number of clones binding to the immobilised small molecule under study is maximised, the method is considered to be valid.

In order to allow this method to successfully amplify selectively bound clones it is important to minimise the clones not bound to the small molecule under study, for example clones binding to the polystyrene surface, or to the solvent-bound, inactivated or unfunctionalised Magic Tag® chemistries on the surface.[19] The washing procedure used for each round must therefore remove a large number of clones with low (non-selective) binding constants with respect to the surface.

The washing procedure used for the biopanning of phage libraries against Magic Tag® immobilised small molecules[20] is six washes using the buffer TBS containing 0.1 % surfactant Tween® 20.[21] After incubating the Magic Tag® well with the phage pool, the bulk of the phage solution is removed and the wells are filled with the washing solution. This solution is agitated in the wells and then removed to a waste container. New wash solution is introduced to the wells and the process repeated five further times.

In the biopanning experiment against M001, the standard washing procedure was compared to two alternative washing procedures. One of these was twelve washes with the same washing solution, and the other was based on the procedure from Smith[22] where the wells are washed with the media used to grow and store the *E. coli* cultures and phage, with an added surfactant. In the literature case, this medium was YB, whereas for us this was LB, made up containing 0.1 % Tween® 20. This has a very similar ionic strength to the solution used to add the phage library to the wells and the media in which the *E. coli* is added, and so was potentially less disruptive to phage binding.

The diversity of the starting phage pool, and the large reduction in phage diversity in the first round of biopanning, makes testing and comparing washing procedures difficult without sampling very large numbers of phage.

In order to give a better idea of what is happening to the diversity of the phage pool round by round, study of every well must be carried out individually. The quantity of phage retained in a biopanning well after washing can be determined by phage titre, which tells us a lot about relative strengths of washing procedures, as a harsher washing condition would be expected to give a lower eluted phage titre. This assay cannot show diversity or changes of diversity during the biopanning assay. As sequencing is relatively expensive (Oxford Sequencing: £600 per 96 well plate, plus PCR cleanup) an estimation of the insert length by gel electrophoresis can be used as a surrogate measure of diversity.

The mobility of nucleic acids in agarose gels is influenced by the agarose concentration (see table 4.1.4) and the molecular size and molecular conformation of the nucleic acid. Nucleic acids move at a rate that is inversely proportional to the logarithm of their molecular weight, allowing estimation of molecular weight by comparison with a nucleotide ladder.[23]

Table 4.1.4: Effect of agarose concentration change

Agarose (% w/v)	Effective Range (kbp)
0.3	5 - 50
0.5	2 - 25
0.7	0.8 - 10
1.2	0.4 - 5
1.5	0.2 - 3
2.0	0.1 - 2

A loading dye and glycerol are added to the DNA sample before loading into the prepared agarose gel. Bromophenol blue loading dye is anionic and moves slightly ahead of the DNA, allowing an ambient light visible marker for the DNA front, and the glycerol assists the DNA to sink to the bottom of the well. Dye added to the agarose gel in preparation (ethidium bromide or GelRed®) intercalates between the

DNA base pairs during electrophoresis.[24] Dye intercalated with DNA shows enhanced fluorescence over dye which is not intercalated, when illuminated with UV light.[23]

By running the gel against a DNA ladder of known length DNA pieces, the length of the sample DNA can be estimated (see figure 4.1.4). This is not an ideal measure of whether a clone is unique, as the prepared phage clones have an average insert length of 200-500 base pairs. If one phage clone was being enriched from the initial phage pool it might be expected that there would be a strong shift in the average insert length, or a narrowing of the standard deviation around the appropriate insert length for that phage. This narrowing might be considered indicative of enrichment of that clone.

Several clones have the same or very similar insert sizes within the error of insert length estimation. The insert size of the phage library population has a roughly normal distribution, with a mode insert length of estimated 350 base pairs. The population has a cut-off of 200 base pairs - the size of the empty T7Select® vector amplified by the 'up' primer. Bands estimated to be shorter than this by agarose gel electrophoresis may have amplified incompletely by PCR, or be a failed insertion product. Depending on the quantity of genetic material in the sample and the quality of the gel run, the primer dimer band at 20 base pairs can sometimes be seen.

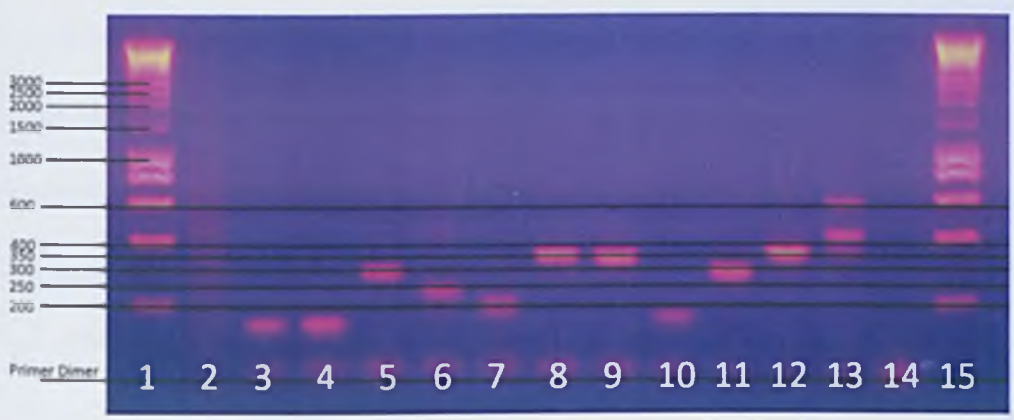


Figure 4.1.4: Estimating phage clone insert size using Bioline Hyperladder™ I DNA ladder

4.1.5 Selective vs non-selective sequencing

When sequencing from the third round of biopanning, as in Chapter 3, 24 clones are chosen from each positive wells and 12 clones are picked from each negative well. This gives an overview of the biopanning output, but quality and quantity of data can be limited by clones which are too short to be successfully sequenced, mixtures of clones picked from overlapping plaques, and a high number of empty vectors. To minimise this loss of information, it is possible to pick more clones than required, and observe the insert lengths of these clones after PCR by agarose gel electrophoresis. PCR products which are 200 base pairs or less are expected to be empty vectors or unsuccessful insert products. PCR products showing two bands on the gel will be difficult to sequence cleanly. If these clones are excluded from the sequencing process, the efficiency of information output is maximised. Selective sequencing was carried out on one of the four small molecules to determine whether this was a feasible process. Selective sequencing on some clones from the prior rounds was also carried out.

4.2 Results and Discussion

4.2.1 Summary

Small molecule M001 was biopanned against the Human Vascular Tissue Library using three different washing methods in order to assay the effect of washing methods on non-specific binding.[25] From each round the phage output was titred and a sample of phage clones from each of the tags and washing procedures were picked. The DNA insert of each picked clone was amplified by PCR and the insert length was estimated by agarose gel electrophoresis against a DNA ladder. Estimating the

length of the insert allows a faster and more cost effective observation of the changes in diversity, including a method for determining whether or not disproportionation has taken place. Observing the shift in insert lengths in comparison to a negative sample gives the clones which are being negatively selected (excluded) from the subsequent rounds and which are being either selectively or non-selectively bound to the small molecule immobilised on the surface and hence enriched. Three further blinded drugs - M002, M003 and M004 - were biopanned against the Human Vascular Tissue Library and sequences collected following the standard procedure as described in the previous chapter.

4.2.2 Validation of the Phage Library

As this biopanning experiment took place after several months of storage of the phage library, as glycerol stock at -80 °C, the library was amplified in *E. coli* to freshly express the displayed peptides and a sample of clones assayed by PCR and gel electrophoresis to ensure there was no disproportionation or loss of viability of the library.

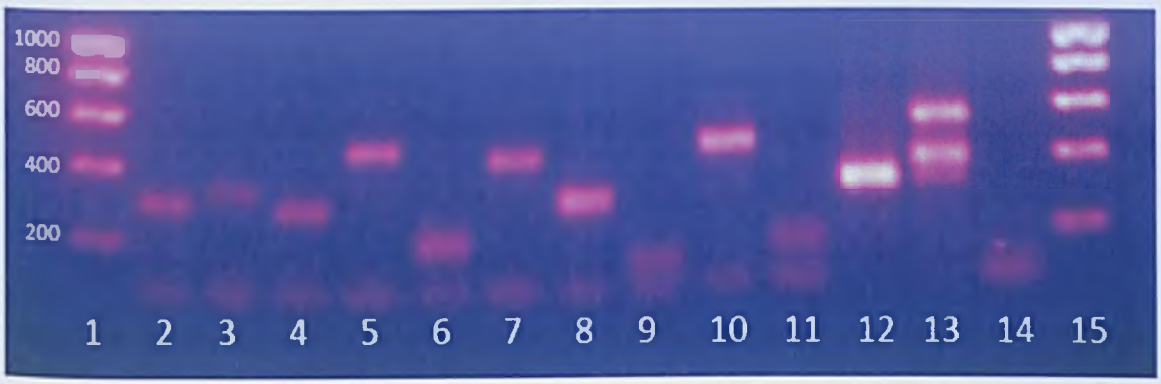


Figure 4.2.3: Sample clones from amplified library assessed by agarose gel electrophoresis. Lane 12 and 13 contain positive controls and lane 14 contains a negative control

Table 4.2.3: Sample clones from amplified library, insert length estimated		
Lane	Well identity	Size of Insert
1	A1	450
2	B1	450
3	C1	200
4	D1	550
5	E1	No Band observed
6	F1	350
7	G1	700
8	H1	450
9	A2	<200
10	B2	<200
11	C2	400
12	D2	400
13	E2	250
14	F2	300
15	G2	320
16	H2	380
17	A3	420
18	B3	200
19	C3	400
20	D3	300
21	E3	<200
22	F3	450
23	G3	200
24	TC1	350
25	TC2	(350) + 400 + 550
26	negative	

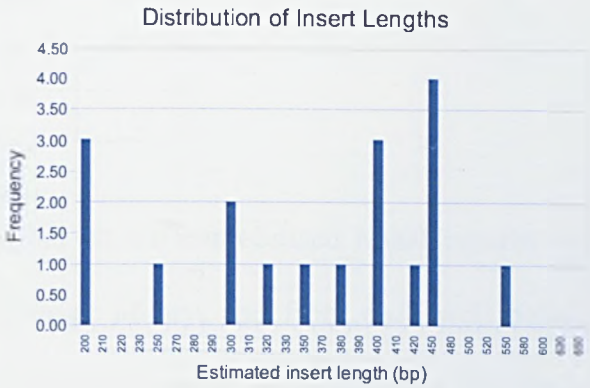


Figure 4.2.3b: Sample clones from amplified library, insert lengths estimated, distribution shown

4.2.3 Immobilisation of M001

Immobilisation of the small molecule M001 in water/DMSO solution onto five Magic Tag® chemistries, PCA, PCB, PCC, PCD, PCE, see Chapter 3.1.2 for identities of these chemistries, in sufficient wells for three rows of four rounds of biopanning, was carried out under 254nm UV light for ten minutes. A single row of Magic Tag® wells for each of the four rounds were prepared with water/DMSO immobilised under the same conditions as the M001 solution.[20]

The three M001 rows were labelled A-C, and the one water/DMSO control row was labelled D. After irradiation the plate was incubated for 50 minutes in the dark. No blocking agent was added. The plate was washed by submerging in a buffered surfactant solution and blot-drying three times.

4.2.4 M001 Washing Assay and Insert length Assay

On the plate the wells were laid out as follows:

Table 4.2.4: Washing types by row as laid out on Magic Tag® Plate	
Wells Row A	Standard washing method - 6x TBST
Wells Row B	Rigorous washing method - 12x TBST
Wells Row C	Alternative washing method - 6x LBT
Wells Row D	Control Wells - 6x TBST

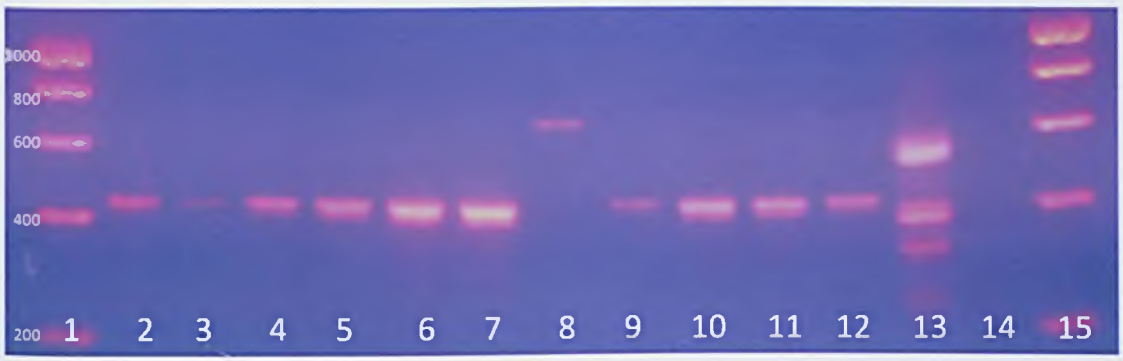
Four rounds of biopanning were carried out on the immobilised M001 against the newly amplified human vascular tissue phage library, and from each well before amplification for the next round a 10 µl sample of phage were collected into 90 µl water. To assay the effectiveness of the different washing procedures, the output of the wells was titred after every round, and analysis into the clones output from each round was also undertaken, see below.

If binding is taking place, and "non-specifically" bound phage are being successfully washed from the surface, the population of the new phage pool round-by-round should be enriching towards a small number of specifically binding clones. To test this theory, the 10 µl sample was diluted, and grown on an *E. coli* lawn. Growth of plaques was allowed to continue for three hours, or until the plaques were clearly defined and not overlapping.

Plaques on this *E. coli* lawn were then counted to titre the phage output and a small number picked into 100 µl water, the DNA insert amplified by PCR and the insert length estimated by agarose gel electrophoresis against a commercial DNA ladder (Bioline Hyperladder I, Cat# Bio-33025).

For the first two rounds 23 plaque picks per well were collected from the positive wells only. The control wells were titred, but not sampled. In the third round, 23 samples were picked from the positive wells, and 12 samples were picked from the

controls. When amplified after round three, the products from the third round disproportionated almost entirely to one clone, although this wasn't discovered until after round four of the biopanning procedure was attempted.



Figures 4.2.4: Agarose gel electrophoresis of a sample of clones from the disproportionated amplification of the third round of biopanning, positive controls in lanes 12 and 13 and negative control in lane 14

Disproportionation occurs when one clone within the population reaches a concentration at which it can out-compete all other clones for the resources available. This generally occurs when one clone has a growth advantage over the other clones in the pool. This one clone then takes over the phage population. There are some significant experimental markers for disproportionation. Lysis of the log phase *E. coli* stock when amplifying a disproportionating phage pool occurs at a much faster rate than the lysis carried out by a standard phage pool. The *E. coli* lysate typically appears dusty, with smaller *E. coli* residue, rather than stringy debris as had been seen previously. The titre of the phage output from the wells (10^7) is a factor higher than the expected 10^5-10^6, and after the picking of a small number of clones, their PCR and assay by agarose gel electrophoresis all of the bands are of the same length. It is not entirely clear what makes one phage clone capable of out-competing the other phage in this manner. Because of this disproportionation, the fourth round of biopanning was not considered to have given any results relevant to the washing assay.

4.2.5 Results of washing solution comparison by titre assay

Raw data is available in the appendix 8.5.1. PC chemistries defined in Chapter 3.1.2.

Table 4.2.5: Titre Changes Across Washing Types and Chemistries

Amplified library before biopanning			
Dilution Factor	Number of Plaques	Volume µl	pfu/mL x10^{10}
10^{-5}	lysis	90	N/A
10^{-6}	1672	90	1.86
10^{-7}	123	90	1.37
10^{-8}	15	100	1.67

ROUND 1	pfu/mL x10^{5}					Washing AVERAGE
	PCB	PCA	PCD	PCC	PCE	
Wash A	9.52	10.4	12.4	11.5	9.44	10.6
Wash B	2.71	5.72	10.3	7.13	5.97	6.37
Wash C	15.3	17.1	14.3	10.5	18.2	15.1
Wash D	21.4	38.3	11.4	12.2	24.0	21.5

ROUND 2	pfu/mL x10^{5}					Washing AVERAGE
	PCB	PCA	PCD	PCC	PCE	
Wash A	5.92	5.00	10.4	5.00	4.45	6.14
Wash B	2.91	6.31	7.00	4.70	5.00	5.18
Wash C	5.08	7.77	7.00	6.18	5.00	6.21
Wash D	1.63	1.45	0.06	2.15	0.81	1.22

ROUND 3	pfu/mL x10^{5}					Washing AVERAGE
	PCB	PCA	PCD	PCC	PCE	
Wash A	5.96	15.0	5.47	11.0	61.7	19.8
Wash B	4.92	5.27	4.55	1.17	40.0	11.2
Wash C	4.77	3.54	3.04	2.11	1.79	3.05
Wash D	3.11	2.16	4.40	16.8	2.73	5.84

MT Average (3 rounds)	6.94	9.83	7.52	7.54	14.9

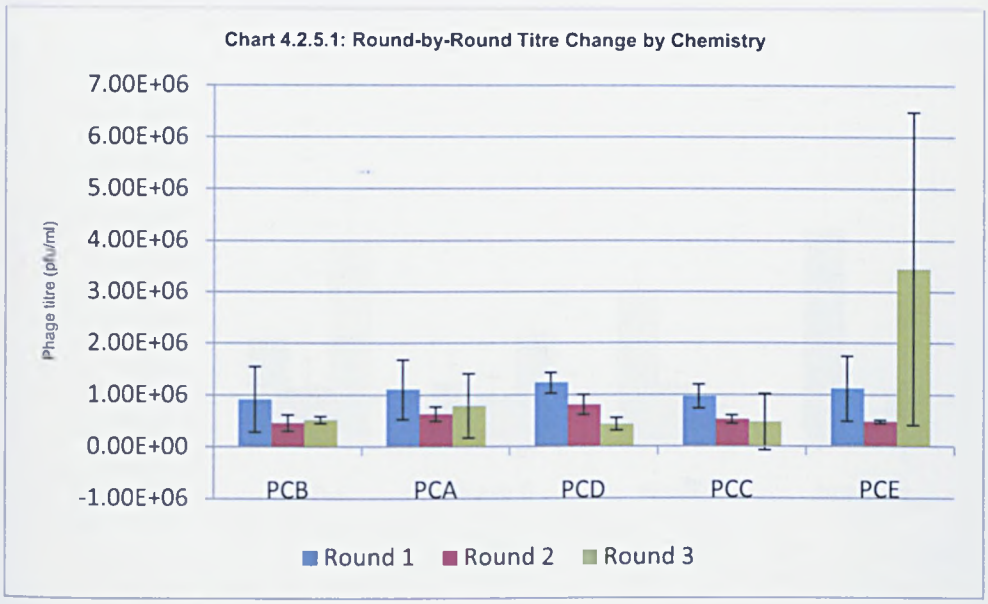
It was expected that, in terms of titre only, fewer phage clones would be eluted in the harsher washing conditions and in the negative wells, and more phage clones would be eluted under the milder LB washing conditions, with ionic strengths similar to the interaction conditions.

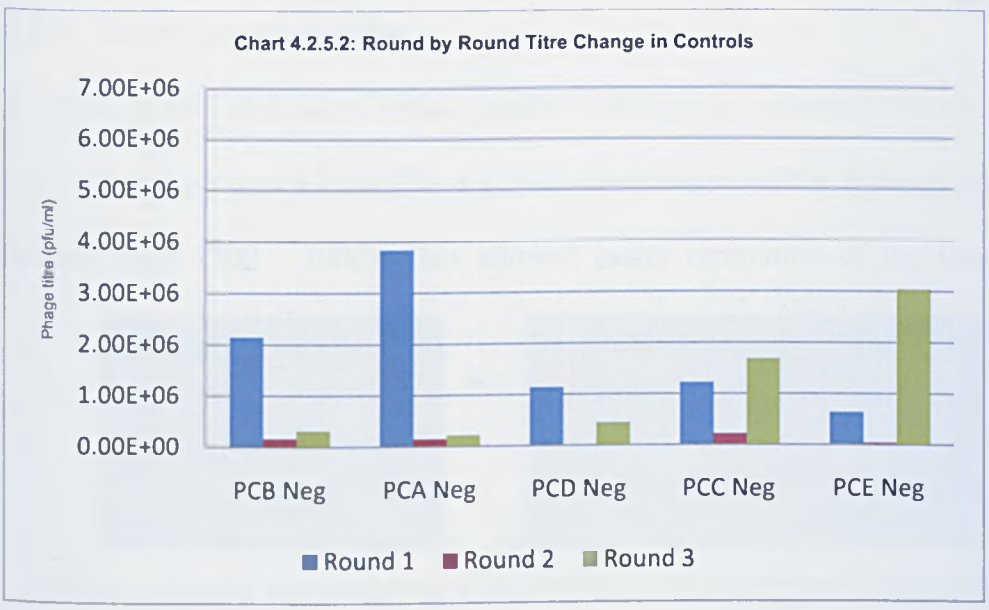
In every round the amplified titre "input" before biopanning was 10^{10} pfu/mL. After each round of biopanning the pfu/mL was 10^{5}-10^{6}. In the first round of biopanning the diversity was expected to drop dramatically, based on the possible interacting surface area of the well.

There will be a consistent background of non-specifically bound material based on the surface area with which the phage have to interact. Any changes in this titre would be observed in the control wells.

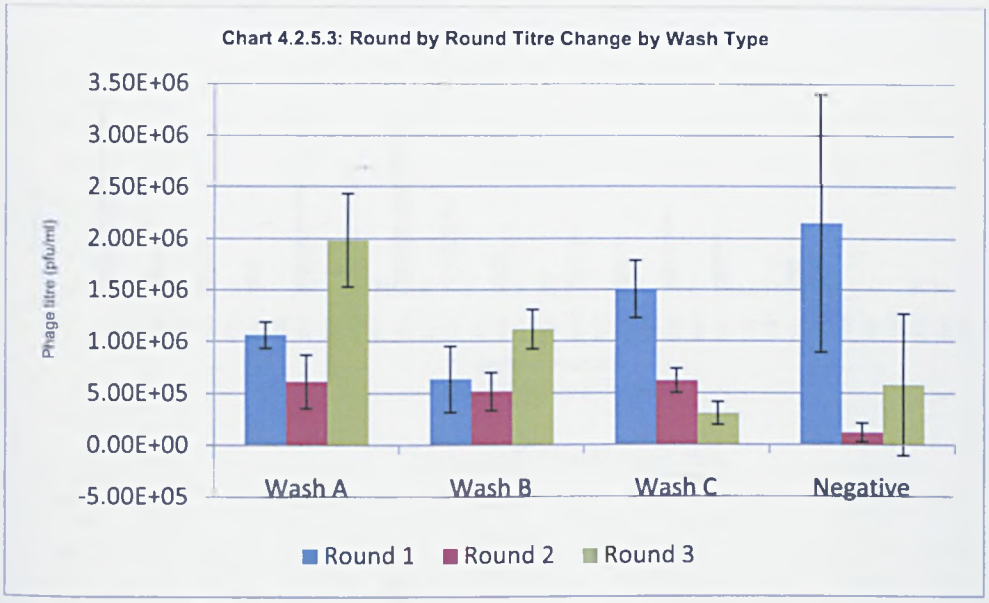
Round-by-round any enrichment of specifically bound phage in the positive (M001 immobilised) wells would lead to an increase in titre over the non-specific background. On this basis, the phage titres for the control wells would be expected to stay low and fairly consistent whereas the positive wells would be expected to slowly increase across the three rounds as enrichment takes place.

Titres were consistent across all five tags. This suggests that binding (whether specific or non-specific) is at a consistent level across all of the five chemistries.



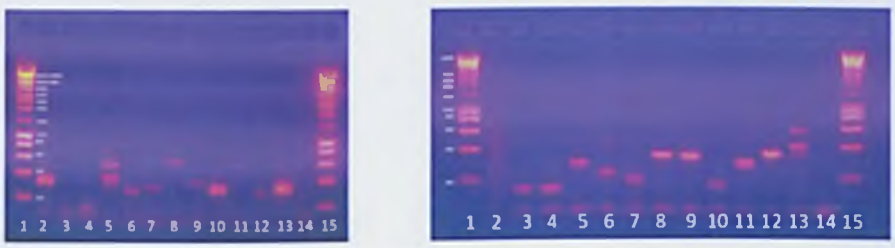


The control wells were not seen to be consistently low across the rounds, in the first round being higher or of equivalent titre to the wells containing immobilised small molecule. This suggests that this first round is very poorly enriched towards binding phage. In rounds two and three the titres of the wells washed with TBS (Wash A and Wash B) increased in comparison to the control wells, but the wells washed with LB (Wash C) did not, suggesting that the LB washing method was not stringent enough to allow enrichment of specifically binding phage against the background of non-specifically bound phage.



4.2.6 Insert Length Results

As most of the bands were initially under 1000 bp, the concentration of agarose in 1X TAE was increased from 1% to 1.5%, which increased the distance between the shorter bands (200 - 1000). This allowed easier estimation of the length of the



Figures 4.2.6a and b: Gel electrophoresis at two different agarose contentrations: a) 1% and b) 1.5%

shorter bands. A change of loading buffer (Bioline 5X DNA loading buffer to Novagen 6X DNA loading buffer) caused occasional band doubling in bands under 1000 base pairs. This led to some ambiguity as to the exact length of the insert. For this analysis of these numbers a midpoint was chosen. Uninserted vector sequences were expected to produce bands of 200 base pairs, but some bands were observed at lower than 200 base pairs in length.

The raw data for the recorded insert lengths is available in appendix 8.5.2.

Chart 4.2.6.1: Round 1 Insert Sizes

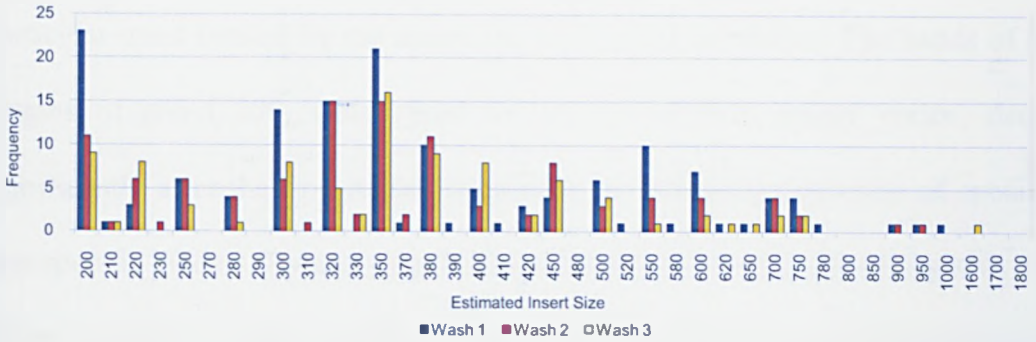


Chart 4.2.6.2: Round 2 Insert Sizes

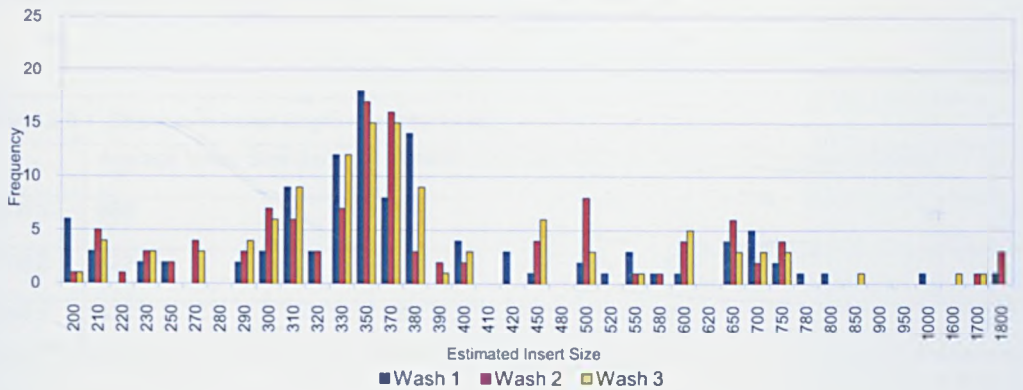


Chart 4.2.6.3: Round 3 Insert Sizes

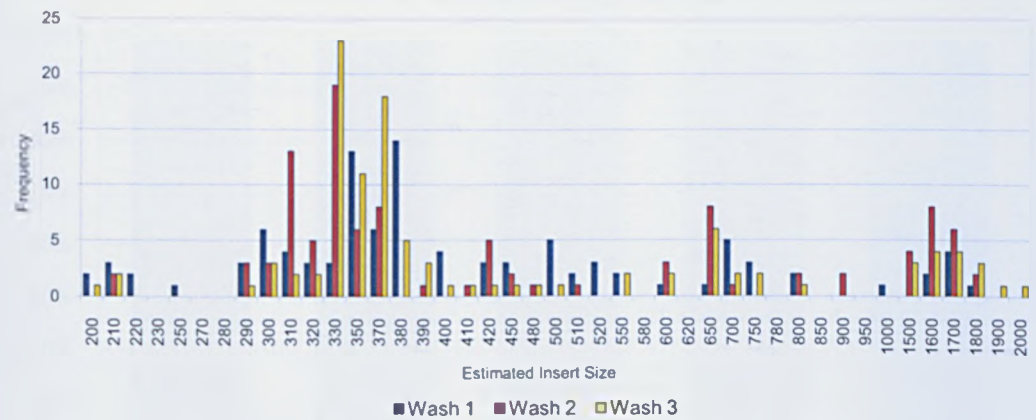
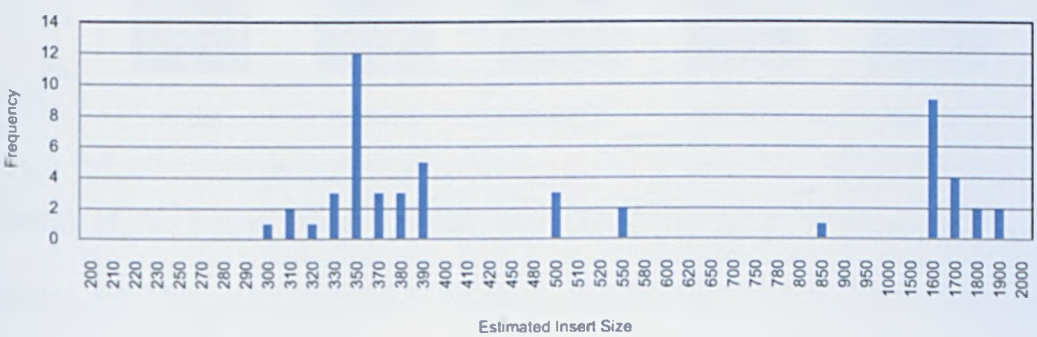


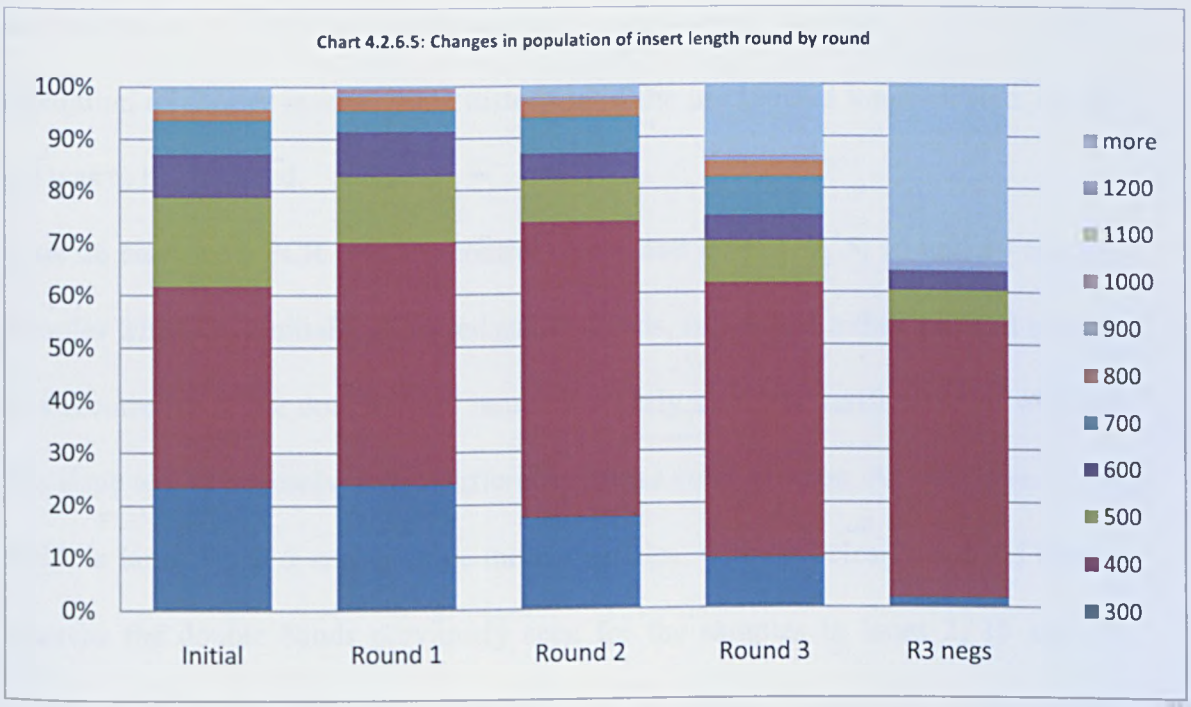
Chart 4.2.6.4: Round 3 Control Well Insert Sizes



Across rounds 1 - 3 there is an apparent narrowing of the distribution with no particular trend created by the magic tag or washing procedure. The bands of insert lengths of about 200, which tend to be derived from phage vector, decrease significantly after the first round, suggesting an increase in the ratio of specific to non-specific binders. By assaying the insert lengths from a sample of negative clones it was possible to observe qualitatively how the phage library population was tending to change from round to round.

Table 4.2.6.1: Changes in insert length round by round

	Average Insert Size (bp)	Number of bands >200 with insert successfully estimated
Round 1	386	294
Round 2	422	333
Round 3	575	316



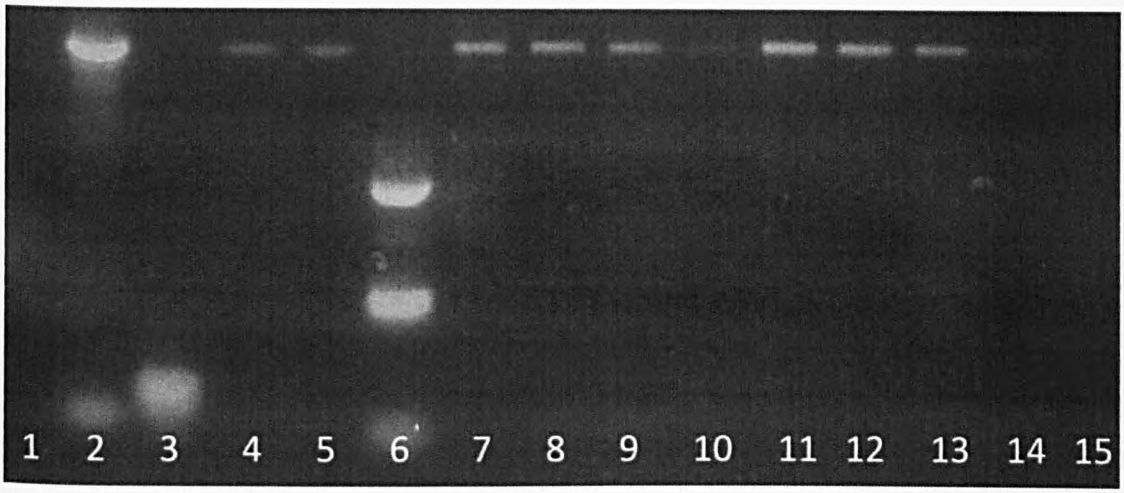
Several of the longer bands (>1000 base pairs) appeared distorted on the 1.5% agarose gel. This was suspected to be due to the different running rates on the higher concentration agarose gel and further distortion due to a large quantity of DNA present in the sample.

Electrophoresis was carried out on these longer inserts on a 1% agarose gel at low concentrations, and these samples appeared to run concurrently.

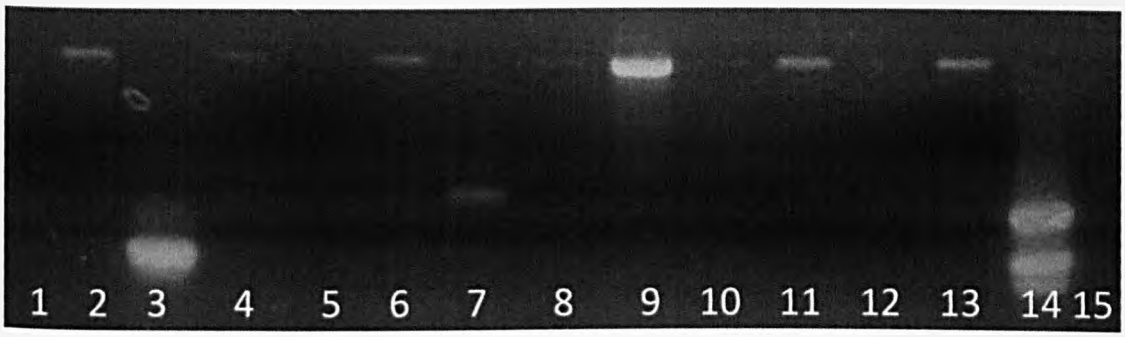
While there were initially estimated to be several different insert lengths between 1500 and 1900, all of these samples appeared at the same insert length in this side-by-side assay. Unfortunately no ladder was used in this gel and so the insert length for this group of samples could not be re-estimated. These gels are shown in black and white to emphasise the weaker bands in lanes 9 and 13 on the top of the gel, 7, 9 and 11 on the bottom of the gel. There was no band in lane 4 of the bottom image, and lane 13 of the same image contains a PCR positive control clone.

This result confirms that the inserts of length over 1000 base pairs tend to distort at higher concentrations on higher percentages of agarose during gel electrophoresis. This led to the inserts appearing at an inconsistent length. The use of 1.5% agarose concentrations for gel electrophoresis was initially used because it allowed better separation of shorter inserts, some distortion of the unexpected longer inserts on the gel was to be expected.

Lane 26 contains a PCR positive control clone, and lanes 1, 2, 5, 15 and 19 contain samples which had initially appeared as two bands, suggesting either a mixed sample or a recurrence of the double band issue previously seen, for verification of whether this issue would be resolved by a different agarose concentration. As can be seen, the DNA in lanes 1 and 5 appear to be mixed samples, with two clearly defined bands, whereas the double bands previously seen for the samples in lanes 2, 15 and 19 appear to have been resolved to one band by the change in agarose concentration, suggesting that this was the cause of this issue.



Figures 4.2.6b and c: Agarose gel electrophoresis of clone picks with determined to be >1000 base pairs or two distinct bands by initial gel analysis (see text)



After confirmation by selective sequencing that the identity of a large percentage of the bands appearing to show inserts greater than 1000 base pairs in length were the same, the mean insert lengths were recalculated with this clone removed and showed very little change round to round.

There appears to be very little change in distribution of insert length with chemistry on the surface, or with washing type, as can be seen in charts 4.2.6.6 and 4.2.6.7.

Chart 4.2.6.6: Round 3 Control Well Insert Sizes

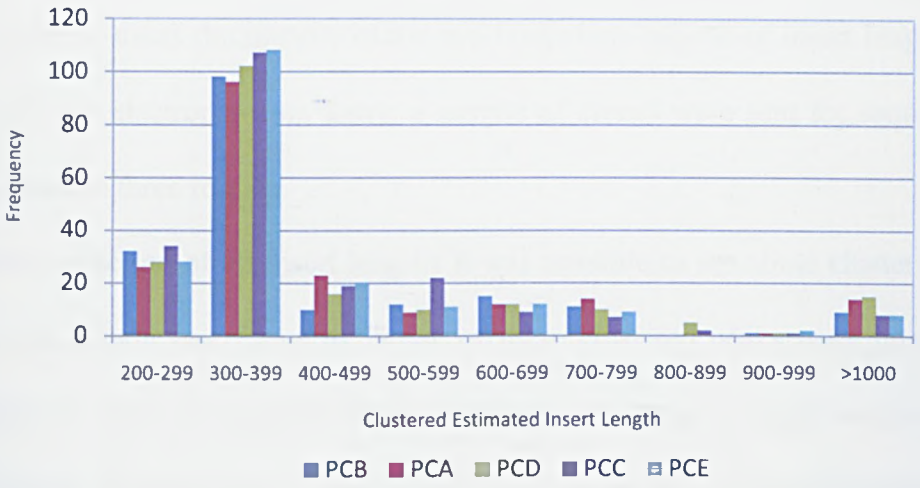
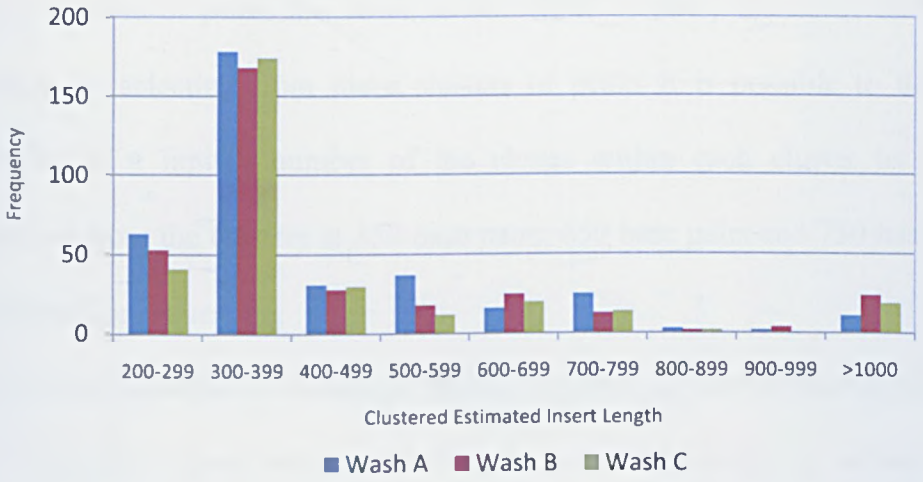


Chart 4.2.6.7: Round 3 Control Well Insert Sizes



4.2.7 Selective Sequencing

In order to assay the identity of the most populous regions of insert length identified by the gel electrophoresis assay, a sample of clones were sent for sequencing from each of the three rounds.

After collection of all insert lengths it was possible to see some clustering of peaks around certain insert lengths. Some of these clustered peaks were mirrored in the negatives taken from round three, and after subtracting the negatives peaks from the positives of round three it was possible to eliminate the clusters of peaks which were mostly due to non-specific interactions between the wells. It is not known whether these clusters of peaks are down to one clone or many clones of the same insert length, by selecting from these clusters of peaks it is possible to determine the identity of a limited number of the clones within each cluster by sequencing. Samples from the clusters at 350 base pairs, 650 base pairs and 750 base pairs were collected for sequencing.

It was also important to determine whether agarose gel electrophoresis could be used to 'weed out' clones which would give vector sequences or which would not successfully sequence. This would allow us to minimise the loss of information inherent in random picks and sequencing. By sequencing a sample of clones whose insert length appeared to be below 200 base pairs, these could be checked for empty vector contamination and for samples which gave samples too short for meaningful alignment.

A small sample of clones for which no band was visible on the electrophoresis gel were also sequenced, as it is possible these missing bands were due to the low concentration of the PCR product rather than a failed clone pick.

Clones were selected whose insert lengths were greater than 1000 base pairs. These clones were of particular interest, because long clones rarely survive the biopanning process due to their growth disadvantage over clones with shorter inserts.

A selection of clones which showed double bands were also sequenced, to determine whether the doubling of the bands was an artefact of the gel or due to contamination during the picking process leading to two different clones present in the pick.

Table 4.2.7: Phage clones submitted to sequencing with their estimated insert length

Insert Length	Number
>1000	46 (+ 17 in negatives)
750	8
650	8
350	35 (+ 12 in negatives)
<200	13 (+ 7 in negatives)
0 (no visible band)	1 (+ 2 in negatives)
Negatives From Round 3 (mixed lengths)	58
Mixed or double bands	20

The results from this sequencing were arranged into consensus regions using the DNAStar Program: Lasergene® SeqMan. This was also used to remove poor data regions and sequences attributed to the vector. This method identified 26 contiguous sequences, 15 of which were determined to be unique and not appearing in the negative sequences. These contigs and the original sequences are available in Appendix 8.5.

Most (81%) of the greater-than-1000 base pair insert sizes fell into M001 Contig 5, which appears in the positives and the negatives, and were sequenced to over 900

base pairs, 94% of the M001 Contig 5 sequences were over 900 base pairs long by sequencing.

After sequencing, the sequence contribution from the phage vector is removed from each end of the sequence. Consequently the insert sequence is typically 100 base pairs shorter than the band on the gel would have predicted.

By sequencing 87% of the inserts with lengths estimated at 350 base pairs were between 500 and 250 base pairs long, averaging 315 base pairs. After sequencing and trimming, this sequence would be expected to be roughly 250 base pairs long. By sequencing 67% of the inserts with lengths estimated at 650 base pairs were between 800 and 500 base pairs long, averaging 467 base pairs. By sequencing only 20% of the inserts with lengths estimated at 750 base pairs were between 1,000 and 600 base pairs long, averaging 532 base pairs. From these results it suggests that it is easier to estimate the length of smaller inserts than that of longer inserts.

Every vector sequence identified in the M001 sequencing experiment showed gel electrophoresis insert lengths below 200 base pairs. Of the 20 clones estimated at shorter than 200 base pairs 11 did not sequence and 7 were vector sequences with no insert.

Associations between insert lengths and certain contigs have been seen, and with a 90% success rate it is possible to identify phage picks which are empty vectors or will not successfully sequence, which suggests that selecting sequences by firstly analysing the round output by gel electrophoresis and then choosing a wide range of insert lengths and avoiding choosing any picks shorter than 200 base pairs may be an effective, if very time consuming, way of determining which clones to sequence to maximise the information gained in sequencing.

4.3 M002-4 Biopanning Results and discussion

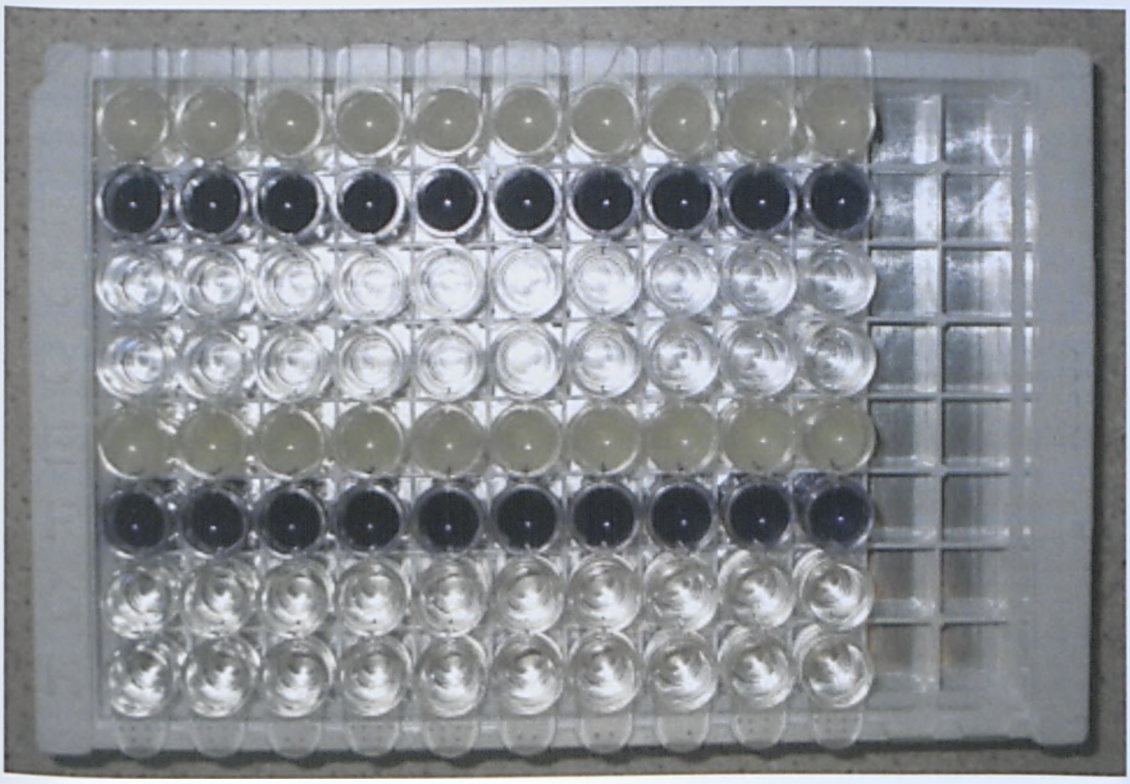
4.3.1 M002-4 Immobilisation

Immobilisation of the small molecules M002, 3 and 4 in water/DMSO solution onto five Magic Tag® chemistries was carried out as described previously.[20]

After irradiation at 254 nm for 10 minutes, the solution of M003 was seen to have changed colour from a grey milky solution to a strongly coloured purple solution. This could be due to a change in the structure brought about by UV irradiation. Further investigation of the behaviour of the small molecule under these UV conditions will be required to ensure the biopanning has been carried out against the correct molecule.

After irradiation the plate was incubated for 50 minutes in the dark. No blocking agent was added. The plate was washed by submerging in a buffered surfactant solution and blot-drying three times.

Figures 4.3.1: Image of the Magic Tag® plate for experiment 4.3.1, showing the dark coloration of M003 solution after irradiation



4.3.2 Biopanning Against Human Vascular Tissue Library

Three rounds of biopanning were carried out against this Magic Tag® plate. After the third round of biopanning, 24 plaques from positive wells and 12 plaques from control wells were picked, and the DNA inserts isolated. The DNA was sent for sequencing. Sequence data is available in appendices 8.6.1, 8.7.1, 8.8.1 and 8.9.1.

4.3.3 Bioinformatics

Returned sequences were processed as described in Chapter 3.2.4. Results are available in Appendices 8.6.2, 8.7.2, 8.8.2, (electronic copies only) 8.6.3, 8.7.3, 8.8.3). Contigs are described below, with the Magic Tag® chemistry each was found on. Chemistry codes are defined in Chapter 3.1.2. An N following the chemistry code dictates that the contig was found on the control well for that chemistry.

Table 4.3.3.1: M001 Contigs, chemistries and wash types

M001 Contig	Chemistries	Round/Wash	Count
M001 Contig 1	PCA, PCB, PCBN, PCC, PCE	R1B, R2B, R3A, R3C	8
M001 Contig 2	PCAN, PCB, PCE	R1B, R2B, R3A, R3C	7
M001 Contig 3	PCD, PCE, PCAN, PCDN	R1A, R1B	5
M001 Contig 4	PCBN		2
M001 Contig 5	PCA, PCB, PCC, PCD, PCE, PCAN, PCBN, PCCN, PCEN	R2A, R2B, R2C, R3A, R3B, R3C	54
M001 Contig 6	PCB, PCE	R1C, R2C	3
M001 Contig 7	PCA, PCE, PCCN	R1B, R3A, R3B	4
M001 Contig 8	Lib, PCC	Lib, R2B	3
M001 Contig 9	PCA, PCC, PCD, PCE, PCAN, PCBN, PCCN, PCDN, PCEN	R1B, R2A, R2B, R2C, R3A, R3B, R3C	39
M001 Contig 10	PCDN, PCEN		2
M001 Contig 11	PCA, PCB, PCD, PCE	R1B, R1C, R2A, R2B, R2C, R3B, R3C	10
M001 Contig 12	PCD, PCE, PCAN, PCEN	R1A, R1B, R2A	5
M001 Contig 13	PCE	R3B	1
M001 Contig 14	PCA	R3B	1
M001 Contig 15	PCC	R2A	1
M001 Contig 16	PCA	R1C	1
M001 Contig 17	PCEN		1
M001 Contig 18	PCCN		1
M001 Contig 19	PCDN		1
M001 Contig 20	D	R2C	1
M001 Contig 21	B	R2B	1
M001 Contig 22	D	R2B	1
M001 Contig 23	D	R3A	1
M001 Contig 24	C	R2B	1
M001 Contig 25	C	R3A	1
M001 Contig 26	B	R3A	1

Table 4.3.3.2: M002 Contigs and chemistries

M002 Contig	Chemistries	Count
M002 contig 1	PCA, PCAN, PCBN, PCDN	11
M002 contig 2	PCE, PCEN	2
M002 contig 3	PCB, PCCN	4
M002 contig 4	PCA, PCB, PCC, PCD, PCE, PCAN, PCBN, PCCN, PCDN, PCEN	35
M002 contig 5	PCC, PCE, PCBN, PCDN	6
M002 contig 6	PCA, PCC, PCBN, PCCN	11
M002 contig 7	PCAN, PCBN	3
M002 contig 8	PCB, PCEN	3
M002 contig 9	PCE, PCAN, PCBN, PCCN, PCEN	5
M002 contig 10	PCB, PCD	2
M002 contig 11	PCB, PCBN	4
M002 contig 12	PCC, PCBN, PCDN	4
M002 contig 13	PCB	1
M002 contig 14	PCD	1
M002 contig 15	PCC	1
M002 contig 16	PCC	1
M002 contig 17	PCE	1
M002 contig 18	PCE	1
M002 contig 19	PCAN	1
M002 contig 20	PCAN	1
M002 contig 21	PCAN	1
M002 contig 22	EN	1

Table 4.3.3.3: M003 Contigs and chemistries		
M003 Contig	Chemistries	Count
M003 contig 1	PCA, PCB, PCC, PCBN, PCDN	12
M003 contig 2	PCD, PCEN	3
M003 contig 3	PCBN, PCCN	3
M003 contig 4	PCA, PCCN	7
M003 contig 5	PCD	2
M003 contig 6	PCC	2
M003 contig 7	PCA, PCB, PCC, PCD, PCE, PCAN, PCBN, PCCN, PCDN, PCEN	36
M003 contig 8	PCD, PCE, PCEN	6
M003 contig 9	PCE	3
M003 contig 10	PCC, PCE, PCAN	7
M003 contig 11	PCA, PCC, PCD, PCE, PCAN, PCBN	12
M003 contig 12	PCD, PCBN	4
M003 contig 13	PCA, PCE	2
M003 contig 14	PCB, PCD, PCE, PCBN, PCDN	7
M003 contig 15	PCA, PCB, PCC, PCD, PCE, PCAN, PCBN, PCCN, PCEN	11
M003 contig 16	PCA, PCB, PCD, PCE, PCAN, PCBN, PCDN	9
M003 contig 17	PCB	1
M003 contig 18	PCB	1
M003 contig 19	PCA	1
M003 contig 20	PCD	1
M003 contig 21	PCC	1
M003 contig 22	PCAN	1
M003 contig 23	PCAN	1
M003 contig 24	PCEN	1

Table 4.3.3.4: M004 Contigs and chemistries		
M004 Contig	Chemistries	Count
M004 contig 1	PCBN	2
M004 contig 2	PCA, PCE, PCCN	8
M004 contig 3	PCA, PCB, PCC, PCD, PCE, PCAN, PCBN, PCCN, PCDN, PCEN	41
M004 contig 4	PCA, PCB, PCD, PCE, PCAN, PCBN, PCCN, PCEN	13
M004 contig 5	PCA, PCD, PCE, PCAN, PCBN	14
M004 contig 6	PCC	7
M004 contig 7	PCC, PCEN	8
M004 contig 8	PCE, PCBN, PCCN	6
M004 contig 9	PCA, PCC, PCBN, PCDN	5
M004 contig 10	PCA, PCB, PCAN	3
M004 contig 11	PCA, PCC, PCBN, PCDN	7
M004 contig 12	PCB, PCAN	2
M004 contig 13	PCA, PCEN	3
M004 contig 14	PCA, PCB, PCD, PCE, PCAN, PCBN, PCDN	10
M004 contig 15	PCB	1
M004 contig 16	PCD	1
M004 contig 17	PCE	1
M004 contig 18	PCE	1
M004 contig 19	PCAN	1
M004 contig 20	PCEN	1

The remaining contigs, those which did not contain negative clones as well as positive clones, were then searched against the RefSeq protein sequence database, using the BLASTX algorithm.

The top annotated protein BLAST hit (not including 'predicted' or 'analogue' protein hits) is listed below. The BLAST parameters and full BLAST output data is available in Appendix 8.5.6, 8.6.4, 8.7.4, 8.8.4 (electronic only).

M001

```
Database: Homo sapiens RefSeq protein

Query= M001 Contig 8
Length=337
>ref|NP_001002.1| 40S ribosomal protein S7 [Homo sapiens]
Length=194
 Score =  176 bits (446),  Expect = 1e-44
 Identities = 88/88 (100%), Positives = 88/88 (100%), Gaps = 0/88 (0%)
 Frame = +3

Query  12   KSRTKNKQKRPRSRTLTAVHDAILEDLVFPSEIVGKRIRVKLDGSRLIKVHLDKAQQNNV  191
            KSRTKNKQKRPRSRTLTAVHDAILEDLVFPSEIVGKRIRVKLDGSRLIKVHLDKAQQNNV
Sbjct  107  KSRTKNKQKRPRSRTLTAVHDAILEDLVFPSEIVGKRIRVKLDGSRLIKVHLDKAQQNNV  166

Query  192  EHKVETFSGVYKKLTGKDVNFEFPEFQL  275
            EHKVETFSGVYKKLTGKDVNFEFPEFQL
Sbjct  167  EHKVETFSGVYKKLTGKDVNFEFPEFQL  194

Query= M001 Contig 11
Length=526
>ref|NP_000423.2| myosin regulatory light chain 2, ventricular/cardiac muscle isoform
[Homo sapiens]
Length=166
 Score =  184 bits (468),  Expect = 3e-47
 Identities = 90/90 (100%), Positives = 90/90 (100%), Gaps = 0/90 (0%)
 Frame = +1

Query  4    INFTVFLTMFGEKLKGADPEETILNAFKVFDPEGKGVLKADYVREMLTTQAERFSKEEVD  183
            INFTVFLTMFGEKLKGADPEETILNAFKVFDPEGKGVLKADYVREMLTTQAERFSKEEVD
Sbjct  77   INFTVFLTMFGEKLKGADPEETILNAFKVFDPEGKGVLKADYVREMLTTQAERFSKEEVD  136

Query  184  QMFAAFPPDVTGNLDYKNLVHIITHGEEKD  273
            QMFAAFPPDVTGNLDYKNLVHIITHGEEKD
Sbjct  137  QMFAAFPPDVTGNLDYKNLVHIITHGEEKD  166

Query= M001 Contig 13
Length=100
>ref|NP_001018114.1|  fumarylacetoacetate  hydrolase  domain-containing  protein  1
isoform
1 [Homo sapiens]
Length=248
 Score = 37.4 bits (85),  Expect = 0.007
 Identities = 16/23 (70%), Positives = 18/23 (79%), Gaps = 0/23 (0%)
 Frame = -3

Query  71   LRQNLALLPRLECSSTISAHCKL  3
            LRQ L L P+LECSS I+AHC L
Sbjct  212  LRQGLTLSPKLECSSAITAHCSL  234

Query= M001 Contig 15
Length=437
>ref|NP_060190.2| signal-transducing adaptor protein 2 isoform 1 [Homo sapiens]
Length=449
 Score = 61.6 bits (148),  Expect = 4e-10
 Identities = 32/41 (79%), Positives = 33/41 (81%), Gaps = 1/41 (2%)
 Frame = +2

Query  29   VETGFHPVCQAGLELLTSGSPPTSSSQSAGWITGVSHHIRP  151
            VE GFH V QAGLELLTS  PPTS+SQSAG ITGVSHH  P
Sbjct  358  VEKGFHHVAQAGLELLTSSDPPTSASQSAG-ITGVSHHTWP  397

Query= M001 Contig 20
Length=286
>ref|NP_001018114.1|  fumarylacetoacetate  hydrolase  domain-containing  protein  1
isoform
1 [Homo sapiens]
Length=248
 Score = 37.4 bits (85),  Expect = 0.007
 Identities = 16/23 (70%), Positives = 18/23 (79%), Gaps = 0/23 (0%)
 Frame = -2

Query  72   LRQNLALLPRLECSSTISAHCKL  4
            LRQ L L P+LECSS I+AHC L
Sbjct  212  LRQGLTLSPKLECSSAITAHCSL  234
```

```
Query= M001 Contig 22
Length=420
>ref|NP_001166925.1| pyruvate dehydrogenase E1 component subunit alpha, somatic form,
mitochondrial isoform 2 precursor [Homo sapiens]
Length=428
 Score = 49.3 bits (116),  Expect = 2e-06
 Identities = 22/32 (69%), Positives = 24/32 (75%), Gaps = 0/32 (0%)
 Frame = -2

Query  221  RQGLAMLPGLVLSSRLKRSSRLGLPNCWDWRH  126
            R GLA LP LV  SRLK+SS LGLP CWD+ H
Sbjct  20   RHGLATLPSLVSISRLKQSSHLGLPKCWDYSH  51

Query= M001 Contig 23
Length=568
>ref|NP_060190.2| signal-transducing adaptor protein 2 isoform 1 [Homo sapiens]
Length=449
 Score = 67.0 bits (162),  Expect = 8e-12
 Identities = 31/40 (78%), Positives = 33/40 (83%), Gaps = 0/40 (0%)
 Frame = -3

Query  533  VDMGFHNVGQAGLELLTSDDPPVSASQSVGITGVSHHAQP  414
            V+ GFH+V QAGLELLTS DPP SASQS GITGVSHH  P
Sbjct  358  VEKGFHHVAQAGLELLTSSDPPTSASQSAGITGVSHHTWP  397

Query= M001 Contig 26
Length=338
>ref|NP_853516.1| gap junction gamma-3 protein [Homo sapiens]
Length=279
 Score = 92.0 bits (227),  Expect = 2e-19
 Identities = 45/48 (94%), Positives = 47/48 (98%), Gaps = 0/48 (0%)
 Frame = -3

Query  144  MCGRFLRRLLAEESRRSTPVGRLLLPVLLGFRLVLLAASGPGVYGESR  1
            MCGRFLRRLLAEESRRSTPVGRLLLPVLLGFRLVLLAASGPGVYG+ +
Sbjct  1    MCGRFLRRLLAEESRRSTPVGRLLLPVLLGFRLVLLAASGPGVYGDEQ  48

Query= M001 Contig 26
Length=338
>ref|NP_853516.1| gap junction gamma-3 protein [Homo sapiens]
Length=279
 Score = 92.0 bits (227),  Expect = 2e-19
 Identities = 45/48 (94%), Positives = 47/48 (98%), Gaps = 0/48 (0%)
 Frame = -3

Query  144  MCGRFLRRLLAEESRRSTPVGRLLLPVLLGFRLVLLAASGPGVYGESR  1
            MCGRFLRRLLAEESRRSTPVGRLLLPVLLGFRLVLLAASGPGVYG+ +
Sbjct  1    MCGRFLRRLLAEESRRSTPVGRLLLPVLLGFRLVLLAASGPGVYGDEQ  48
```

M003

```
Database: Homo sapiens RefSeq protein

Query= M003 contig 9
Length=321
>ref|NP_872601.1| histone demethylase UTY isoform 1 [Homo sapiens]
Length=1079
 Score = 67.0 bits (162),  Expect = 4e-12
 Identities = 42/84 (50%), Positives = 45/84 (54%), Gaps = 18/84 (21%)
 Frame = -1

Query  294   QAGV*WCALGSLHPPPSGFKRFFCLSL----------------LVEMGFCHFAQAGLE  169
             +AG+ WC L SL PPP GFKRF  LSL                 VE GF H  QA LE
Sbjct  995   RAGMQWCDLSSLQPPPPGFKRFSHLSLPNSWNYRHLPSCPTNFCIFVETGFHHVGQACLE  1054

Query  168   LVTSSDPHPSALQSAGNTGVSHRA  97
             L+TS     SA QSAG TGVSH A
Sbjct  1055  LLTSGGLLASASQSAGITGVSHHA  1078
```

M004

```
Database: Homo sapiens RefSeq protein

Query= M004 contig 16
Length=390
>ref|NP_001166925.1| pyruvate dehydrogenase E1 component subunit alpha, somatic form,
mitochondrial isoform 2 precursor [Homo sapiens]
Length=428
 Score = 49.3 bits (116),  Expect = 3e-06
 Identities = 22/32 (69%), Positives = 24/32 (75%), Gaps = 0/32 (0%)
 Frame = -3

Query  232  RQGLAMLPGLVLSSRLKRSSRLGLPNCWDWRH  137
            R GLA LP LV  SRLK+SS LGLP CWD+ H
Sbjct  20   RHGLATLPSLVSISRLKQSSHLGLPKCWDYSH  51
```

4.4 Conclusions

4.4.1 M001 Washing Assay

By phage titre we have shown that a 0.5% LB Tween® washing solution was not sufficiently harsh to lend itself to removing non-specifically bound clones and allowing the enrichment of specifically bound clones. The harsher washing conditions (12 washes using 0.5% TBS Tween®) did not improve the titre by any notable amount over the current washing method of 6 washes using 0.5% TBS Tween®, may in fact have led to a reduction in enrichment.

4.4.2 M001 Insert Length Assay and Selective Sequencing Method

Over the three rounds of biopanning the appearance of one clone with a long insert size (>1000 base pairs), shown in the negatives to be non-specifically binding, appeared and enriched from round to round. By sequencing it was confirmed that this very long insert size group was represented almost entirely by one clone (M001 Contig 5), which was highly represented in the negatives. The enrichment of this insert resulted in the increase in the mean insert length round by round as estimated by gel electrophoresis, but when the influence of this clone was removed the mean insert size was very stable at 409 ± 10 base pairs from the initial library through all three rounds.

There was evidence of some clustering of insert lengths around 650-750 base pairs, and around 450 base pairs in length. The latter trend was represented in the negatives also. The initial library is highly represented in the 300-400 base pairs insert length, so any changes or trends in these insert lengths are hard to discern from the results.

That the negative wells were overtaken by the long clone (M001 Contig 5) and the positive wells retained a portion of other clones suggests that the presence of the molecule M001 on the surface was either reducing the binding of the clone, or selective binding was competing with this clone for interacting surface area. This is encouraging when looking at any clones sequenced from this biopanning assay.

Running agarose gels of clone picks before sequencing would allow clone picks which would show unmodified phage vector or not give sequence information to be identified before sequencing, and be removed from further analysis. Based on this method, the choice picks from M001 assay would give no vector assignments and 85.8% useable sequence output, whereas M002 with random picks gave 16.7% vectors and 47.5% useable output, M003 gave 14.2% and 75.8% respectively and M004 gave 8.3% and 76.7%.

4.4.3 Biopanning results

Three rounds of biopanning were carried out on all four molecules. Selective sequencing was carried out on M001 and random sample sequencing was carried out on M002, M003 and M004.

After initial bioinformatics, hits of interest for the molecule M001, rhein hydroxymate, include the ribosomal protein S7, myosin regulatory light chain 2, fumarylacetoacetate hydrolase domain containing protein and signal transducing adaptor protein 2. For the molecule M003, protizinic acid, one hit of interest was

returned, histone demethylase UTY. Hits of specific interest against these molecules include pyruvate dehydrogenase E1 component against compounds M001 and M004, rebamipide, and gap junction gamma-3 protein against M001. Further work will consist of validation for these interactions and analysis into their potential implications.

There were no resulting polypeptide interactions with nucleotides which returned annotated sequences for M002, oxametacin.

4.5 Methods and Materials

4.5.1 General Experimental

This experimental work was carried out under the HSE's GMO (contained use) guidelines at PepTCell's labs in Cherwell Innovation Centre, Bicester. Water used in these reactions was DI RO water, prepared on site. For all uses where the prepared solution or media was not autoclaved prior to use, the water used was autoclaved prior to use.

LB was prepared as follows: 20 LB capsules in 800 mL DI RO water were warmed to 30°C with stirring until completely dissolved. This mixture was autoclaved and after cooling to below 40°C ampicillin (800 uL, 50 mg/mL) was added. LB was allowed to return to room temperature and stored at 4°C.

0.5% LBTween® was prepared as above, with 1.1 mL Tween® 20 added before autoclaving of the LB solution. No ampicillin was added to the LBTween® solution.

0.5% TBSTween® was prepared as follows: in 800 mL DI RO water, 6.4 g tris and 9.3 g NaCl were dissolved. The pH of the buffer was adjusted to 7.53 using HCl. 1.07 mL Tween® 20 was added and the solution was autoclaved, allowed to room temperature and stored at 4°C.

To measure the optical density of the culture solutions, 1 mL *E. coli* solution was placed in a capped cuvette and absorbance was measured at 600 nm against a fresh media control. Measurements were carried out on a Biomate 3 UV/vis Spectrometer.

Novagen Taq Polymerase PCR mastermixes were prepared as previously described in chapter 2.3.1. PCR cycle programs were consistent with these methods, run on a Techne Genius PCR machine, or a MWG AG Biotech Primus HT PCR machine.

Hybripol (Bioline, Cat #BIO-21080) PCR mastermix (per sample): 5 µL 10X reaction buffer, 1.5 µL MgCl2 solution (50 mM, final concentration 15 mM), 1 µL T7Select® Up Primer, 1 µL T7Select® Down Primer, 1 µL dNTP mix (10 mM each dNTP), 0.25 µL Hybripol DNA polymerase, made up to 49 µL with water. Same cycle program used as previous.

From commercial primer samples, using the stated mass of primer, stock solutions of 50 nmol/mL were prepared and stored at -20°C. From these stock solutions, 100 µL aliquots of 5 nmol/mL were prepared and stored at -20°C.

For agarose gel electrophoresis, two different sizes and concentrations of gels were prepared as follows – TAE (120 mL or 80 mL, 1X) and agarose (1% or 1.5% g/mL) were melted in a microwave. GelRed® (0.05% v/v, 10,000X in water, Biotium, Cambridge Biosciences, Cat# BT41003) was added and the gel poured into a gel frame and allowed to set with combs in place to prepare the running lanes. After the gel was set the comb was removed and the PCR product (8 µl) was loaded with 20% by volume Crystal 5X DNA loading buffer blue (Bioline, Cat# BIO-37045). 5 µl Hyperladder™ I (Bioline, Cat# Bio-33025) was added to the lanes on either side of the samples being assayed. Where two combs were used, two lanes of Hyperladder™ I were added on both rows. The gel was run in the electrophoresis

module at 100-120 V for 45-90 mins. Gel was visualised under UV light and photographed.

4.5.2 Library validation

Stored 80% glycerol solution of human vascular tissue phage library (100 µL) was amplified in log phase *E. coli* (5 mL) at 37°C with agitation for 3 hrs. The lysate was spun down and the clarified lysate supernatant was decanted into a fresh tube. Agarose-agar plates were prepared at dilutions of 10^{-5}, 10^{-6}, 10^{-7} and 10^{-8}. 16 plaques were picked from these plates, PCR carried out and 1% agarose gels run.

4.5.3 Immobilisation of M001

Immobilisation of the drug onto Magic Tag strips 1, 2, 4, 5 and 6 was carried out at a drug concentration of 7.4 mg/mL suspension in 8% DMSO in water. A solution of 8% DMSO in water was prepared for the negative control. These solutions were prepared as described below:

Table 4.5.3: M001 immobilisation

M001	89.2	mg
	897	µl DMSO
	11.121	mL water
	12.018	mL total volume
	60	wells
	200	µL/well
	299.3	g/mol

% DMSO	drug conc.	drug conc.	drug conc.	drug conc.
8	7.4	25	1.5	5
	mg/mL	µmol/mL	g/well	µmol/well

Three positive rows (A-C) and one negative row (D) per round, over four rounds were prepared. After irradiation at 254 nm for 10 minutes the plate was incubated for

50 minutes in the dark. No blocking agent was added. The plate was washed by submerging in 0.05% Tween® in PBS and blot-drying three times.

4.5.4 Immobilisation of M002, M003 and M004

Immobilisation of M002 onto Magic Tag® strips 1, 2, 4, 5 and 6 was carried out at a drug concentration of 0.008 g/mL in 8% DMSO in water. Immobilisation of M003 onto Magic Tag® strips was carried out at a drug concentration of 0.008 g/mL in 8% DMSO in water. Immobilisation of M004 onto Magic Tag® strips was carried out at a drug concentration of 0.0004 g/mL in 8% DMSO in water (due to solubility issues). All three of these compounds were poorly soluble in the solvent mixture, giving predominantly suspended compounds. Using the consistent concentration of DMSO allowed the use of one negative control for all three compounds.

Table 4.5.4: M002/3/4 Immobilisation

M002

72	mg
720	µL DMSO
8.280	mL water
9.000	mL total volume
45	Wells
200	µL/well
373	g/mol

% DMSO	drug conc.	drug conc.	drug conc.	drug conc.
8.0	8.0	21	1.6	4.29
	mg/mL	µmol/mL	mg/well	µmol/well

M003

72	mg
720	µL DMSO
8.280	mL water
9.000	mL total volume
45	Wells
200	µL/well
315	g/mol

% DMSO	drug conc.	drug conc.	drug conc.	drug conc.
8.0	8.0	25	1.6	5.08
	mg/mL	µmol/mL	mg/well	µmol/well

M004

3.6	mg
720	µL DMSO
8.280	mL water
9.000	mL total volume
45	Wells
200	µL/well
371	g/mol

% DMSO	drug conc.	drug conc.	drug conc.	drug conc.
8.0	0.4	1	0.08	0.22
	mg/mL	µmol/mL	mg/well	µmol/well

Three positive rows (A - M002, B - M003, C - M004) and one negative row (D) per round, over four rounds were prepared. The plate was irradiated for ten minutes, after which the solution of M003 had turned a strong purple colour. After irradiation the plate was incubated for 50 minutes in the dark. No blocking agent was added. The plate was washed following the general procedure.

4.5.5 Biopanning of M001 with three washing solutions

Amplified and validated phage library (200 µl per well) was added to round 1 drug wells and control wells and agitated for 45 minutes. Excess phage solution was removed and wells washed as follows: Row A, washed with 0.5% Tween® in TBS six times with 2 minutes agitation for each wash; Row B, washed with 0.5% Tween® in TBS twelve times with 2 minutes agitation for each wash; Row C, washed with 0.5% Tween® in LB six times with 2 minutes agitation for each wash; controls in Row D, washed with 0.5% Tween® in TBS six times with 2 minutes agitation for each wash. Log phase *E. coli* was prepared as described in chapter 3.3.2 and 200 µl of the log phase *E. coli* was added per well, at OD_{600} ~0.5 and agitated for 10 minutes to elute the remaining bound phage. 10 µl of this solution was removed to 90 µl water, agarose/agar plates were prepared at a dilution of 10^{-2} and 10^{-3} as has been described previously and the remaining solution was stored at 4 °C. These plates were titred and 32 plaques were picked from the agarose surface from the positive (drug immobilised) wells only. Some plates were stored overnight and picked the following day without apparent detriment to the pick quality. The insert DNA from these picks was amplified by PCR and the insert length determined by gel electrophoresis run as described in 4.4.1.

The first round well contents were amplified in 5 mL log phase *E. coli* in labelled 50mL tubes. The tubes were incubated for 3 h at 37 °C with shaking during which time the solution cleared as the *E. coli* were lysed. 2 mL aliquots of amplified phage from each tube were separated from any *E. coli* detritus by centrifuge and the clarified lysates decanted into new labelled tubes. Well-specific products from the first round of biopanning (200 µl per well) were then introduced to the next row of immobilized drug and control wells and the process repeated for rounds two, three and four. If more than a week was allowed to elapse between rounds of biopanning the phage were amplified again just prior to use.

From the third round of biopanning 12 phage picks from every control well were collected and PCR and agarose gel electrophoresis carried out as above.

After the fourth round of biopanning the tubes for amplification of the well outputs remained cloudy after 4.5 hours and the agarose/agar plates prepared from the outputs showed plaques equivalent to a titre of 10^7 pfu/ml in only 1.5 hours. This was due to disproportionation. Agarose/agar plates from a 10^{-6} dilution of two different amplified round 3 phage well outputs were prepared, a sample of plaques picked, PCR carried out and agarose gel electrophoresis run to demonstrate disproportionation.

Picked clones and PCR product were both stored at 4°C initially, then picked clones were archived as glycerol stocks (1:9 80% glycerol solution to phage solution) and stored at -80°C.

4.5.6 PCR and gel analysis of M001 Biopanning outputs

Plaque pick insert DNA was amplified by PCR (1 µl pick solution in 49 µl mastermix) (see general experimental 3.3.1 for PCR master mix and cycle program) and stored at 4°C. Two positive controls and a water negative control were used to validate each plate of PCR reactions. These PCR products were stored at 4°C.

4.5.7 Sequencing analysis of M001 Biopanning outputs

PCR products for chosen samples were transferred from the PCR plate into new 96-well plates which were sealed with foil covers. Where there was only a limited volume of PCR product remaining - through evaporation and gel electrophoresis analysis - 20 µl DI RO autoclaved water was added to the PCR well to dissolve any remaining product in the well before transferring to the new plate.

This PCR plate was sent directly for PCR clean-up and sequencing to Oxford Sequencing Service. The plate was transported in a cold box with ice packs.

4.5.8 Biopanning of M002, M003, M004

Amplified and validated phage library (200 µl per well) was added to round 1 drug wells and control wells and agitated for 45 minutes. Excess phage solution was removed and wells washed with 0.5% Tween® in TBS six times with 2 minutes agitation for each wash. Log phase *E. coli* was prepared as described in chapter 3.3.2 and 200 µl of the log phase *E. coli* was added per well, at OD_{600} ~0.5 and agitated for 10 minutes to elute the remaining bound phage. The first round well contents were amplified in 5 mL log phase *E. coli* in labelled 50mL tubes. The tubes were incubated for 3 h at 37 °C with agitation during which time the solution cleared as the *E. coli* were lysed. 2 mL aliquots of amplified phage from each tube were

separated from any *E. coli* detritus by centrifuge and the clarified lysates decanted into new labelled tubes.

Well-specific products from the first round of biopanning (200 µl per well) were then introduced to the next row of immobilized drug and control wells and the process repeated. On the third round of biopanning the well contents from the elution step were sampled and diluted to 10-2 and 10-3 followed by agarose/agar plate assay. (For plating method, see Chapter 2.3.5)

From these plates, 24 plaques from positives wells (PCB, PCA, PCD, PCC and PCE, 120 clones for each drugs, 360 clones total) and 12 plaques from water controls (a total of 60 clones) were picked for PCR and sequencing (PCR, agarose gel electrophoresis and sequencing methods described previously. Picked clones and PCR product were both stored at 4°C initially. Picked clones were archived as glycerol stocks (1:9 80% glycerol solution to phage solution) and stored at -80°C.

4.5.9 Sequencing analysis of M002, M003 and M004

Sequencing for all picked samples were prepared as in 4.4.7 and sent directly for PCR clean-up and sequencing to Oxford Sequencing Service.

4.6 References

1. H. McGavock, *How drugs work: basic pharmacology for healthcare professionals*, Radcliffe Publishing, 2005.
2. P. Winstanley and T. Walley, *Medical pharmacology: a clinical core text for integrated curricula with self-assessment*, Churchill Livingstone, 2002.
3. H. E. Vonkeman and M. A. F. J. van de Laar, *Semin. Arthritis Rheum.*, 2010, **39**, 294-312.
4. P. Libby, *Nature*, 2002, **420**, 868-874.
5. J.-H. Lee, J. M. Kim and C. Kim, *J. Ethnopharmacol.*, 2003, **84**, 5-9.
6. A. F. Mendes, M. M. Caramona, A. P. de Carvalho and M. C. Lopes, *Pharmacol. Toxicol.*, 2002, **91**, 22-28.
7. B. Dozin, M. Malpeli, L. Camardella, R. Cancedda and A. Pietrangelo, *Matrix Biol.*, 2002, **21**, 449-459.
8. C. Sanchez, M. Mathy-Hartert, M. A. Deberg, H. Ficheux, J. Y. L. Reginster and Y. E. Henrotin, *Biochem. Pharmacol.*, 2003, **65**, 377-388.
9. J. P. Pelletier, F. Mineau, J. C. Fernandes, N. Duval and J. Martel-Pelletier, *J. Rheumatol.*, 1998, **25**, 2417-2424.
10. S. Franchimicheli, L. Lavacchi, C. A. Friedmann and L. Zilletti, *J. Pharm. Pharmacol.*, 1983, **35**, 262-264.
11. J. Broekhuysen, F. Deger, J. Douchamps, H. Ducarne, A. Herchuelz and J. Polderman, *Eur. J. Clin. Pharmacol.*, 1983, **24**, 671-673.
12. J. S. Franzone, T. Natale and R. Cirillo, *Farmaco-Edizione Scientifica*, 1980, **35**, 498-503.
13. J. Polderman and M. Colon, *J. Int. Med. Res.*, 1980, **8**, 156-162.

14. K. D. Stewart, S. Loren, L. Frey, E. Otis, V. Klinghofer and K. I. Hulkower, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 529-534.

15. T. Arakawa, K. Higuchi, Y. Fujiwara, T. Watanabe, K. Tominaga, E. Sasaki, N. Oshitani, T. Yoshikawa and A. S. Tarnawski, *Dig. Dis. Sci.*, 2005, **50**, S3-S11.

16. Y. Niwa, M. Nakamura, N. Ohmiya, O. Maeda, T. Ando, A. Itoh, Y. Hirooka and H. Goto, *J. Gastroenterol.*, 2008, **43**, 270-276.

17. R. M. Genta, *Aliment. Pharmacol. Ther.*, 2003, **18**, 8-13.

18. S. E. Heman-Ackah, S. K. Juhn, T. C. Huang and T. S. Wiedmann, *Otolaryngology-Head and Neck Surgery*, 2010, **143**, 429-434.

19. S. J. Dilly, M. P. Beecham, S. P. Brown, J. M. Griffin, A. J. Clark, C. D. Griffin, J. Marshall, R. M. Napier, P. C. Taylor and A. Marsh, *Langmuir*, 2006, **22**, 8144-8150.

20. S. J. Dilly, M. J. Bell, A. J. Clark, A. Marsh, R. M. Napier, M. J. Sergeant, A. J. Thompson and P. C. Taylor, *Chem. Commun.*, 2007, 2808-2810.

21. S. Zampieri, A. Ghirardello, A. Doria, M. Tonello, R. Bendo, K. Rossini and P. F. Gambari, *J. Immunol. Methods*, 2000, **239**, 1-11.

22. G. P. Smith, *Science*, 1985, **228**, 1315-1317.

23. R. Boyer, *Modern Experimental Biochemistry*, Addison Wesley Longman, Inc, San Francisco, 2000.

24. V. A. Gault and N. H. McClenaghan, *Understanding Bioanalytical Chemistry, Principles and Applications*, Wiley-Blackwell, Chichester, 2009.

25. F. D'Mello and C. R. Howard, *J. Immunol. Methods*, 2001, **247**, 191-203.

26. K. D. Pruitt, T. Tatusova and D. R. Maglott, *Nucleic Acids Res.*, 2007, **35**, D61-D65.

Chapter 5 Cleavable Linker Chemistry

5.1 Introduction

5.1.1 Aims

Interactions of a drug molecule through unexpected (off-target) pathways may be beneficial, giving new pharmacological action, or may lead to toxicity.[1] One means of investigating such activity is by seeking new cellular molecular targets through immobilisation of the drug molecule onto a bead or plate surface[2] and probing it with polypeptides from cellular lysate[3] or a displayed library.[4] A new photoimmobilisation strategy[5] uses functionality including benzophenones,[6] azides,[7] and diazirines[8-10] to ensure that most orientations of the immobilised drug, or other small molecule bioactive, are presented to potential targets. This approach, as embodied in the Magic Tag® kit, has been shown to facilitate discovery of novel interactions from genomic polypeptide libraries.[5, 11]

In order to evaluate the molecular structure in relation to these binding events, we need to know the orientation of the drug molecule on the surface. In this chapter, I present the first acid labile supported photoactive chemistries intended to allow the covalently bound drug molecule to be recovered from the support whilst remaining attached to the photoactive functionality. This approach has recently been shown to be viable using a reductively cleavable disulfide derivatised with a photoactive diazirine that is related, although not identical to that used in our current work.[12] In addition we note the success of the bifunctional photoaffinity tag[13] and solution-based tag-and-click approaches.[14] We anticipate that based upon the appearance of dominant regioisomers from cleavage of the photoactive group – drug molecule

adducts, proposals might be made for structure-activity relationships from observed drug molecule – polypeptide interactions in phage biopanning screens.

Rather than use a 96-well plate format as preferred previously for biological assays, our new cleavable analogues were prepared initially on polystyrene 'Wang' resin beads.[15] The advantage of using beads rather than plates is that it is possible to characterize the solid phase product by elemental analysis, infra-red and other spectroscopic measurements.[16, 17] Cleavage from the resin has previously[18, 19] been possible in sufficient quantity to enable analysis of the products by standard solution phase techniques.

5.1.2 Small Molecule affinity binding

Small molecule pharmaceuticals often have complex structures, and by binding a small molecule to a surface in order to carry out a screening or selection process, part of this structure is obscured or concealed; indeed the chemical or pharmacological nature of the molecule may be dramatically altered. This, together with the nature of the surface, limits the accessibility of the small molecule to the binding site and controls which parts of the small molecule are available for binding.

The Magic Tag® system is designed to maximise regiochemical diversity by using five different chemistries, and photoactivated immobilisation which is generally non-specific in its method of reaction. Some specificity has been identified in certain functional groups.[20, 21] Different arrangements of the small molecule on the surface will allow it to interact with different binding sites in the protein/phage library.

The low specificity of immobilisation means that it is difficult to assess which part of the molecule has interacted with the identified polypeptide target. If this target might result in a detrimental side effect it would be useful to be able to identify the

part of the small molecule responsible for binding, so that the molecule might be modified to reduce the likelihood of this binding event occurring. To do this it would be useful to be able to identify the mode of binding of the small molecule on the surface. It would also be beneficial to observe the complete interaction site in complex with the small molecule,[12] the methods for which are not covered in this thesis.

5.1.3 Photoactive chemistries and references

The five different chemistries used in the Magic Tag® photoimmobilisation method utilise three different photochemistries - benzophenones, aryl azides and diazirines. These chemistries and their interactions have been heavily reviewed elsewhere,[6, 8, 22] and a brief introduction is given here.

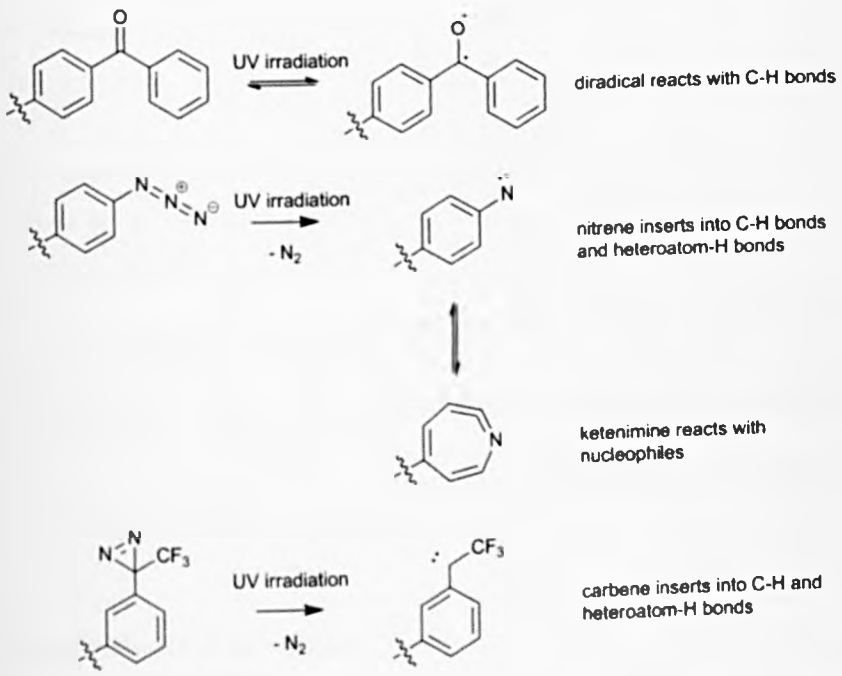


Figure 5.1.3a: Photoactive chemistries used, and their active forms after UV irradiation

Benzophenones are irradiated to form the triplet state active moiety, which tends to behave as a di-radical, abstracting one hydrogen, followed by radical termination or recombination with the remainder of the molecule or anything which may be closer. Through interaction with oxygen, these photoactivated

chemistries can undergo intersystem crossing to become singlet state, leading to a reaction where the oxygen of the benzophenone ketone inserts into a bond.

The aryl azide chemistries and the diazirine chemistry form nitrenes and carbenes respectively upon UV activation. These are formed as the singlet state nitrogen or carbon, and are able to insert directly into a bond, and as with the benzophenones can intersystem cross, without the assistance of oxygen, to form the triplet state which is capable of abstracting a proton followed by radical recombination. Aryl nitrenes are prone to deactivation by ring expansion to the ketenimine 7-membered ring, which is only sensitive to nucleophilic attack.

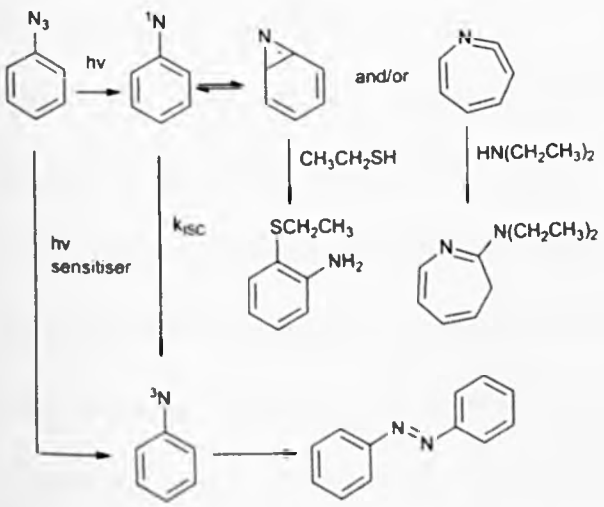


Figure 5.1.3b: Details on the reactivity of the azide moiety after UV activation to a nitrene - figure from W. T. Borden, N. P. Gritsan, C. M. Hadad, W. L. Karney, C. R. Kemnitz and M. S. Platz, Acc. Chem. Res., 2000, 33, 765-771.

It is obvious from the literature[21-23] that rearrangement and undirected interactions of these different photoprobes lead to decreases in yield and the formation of the desired photoadducts is a challenge to achieve in good yield.

5.1.4 Magic Tag® Plates

Magic Tag® Plates are prepared from amino propyl derivatised polystyrene 96-well plates (Corning Aminopolystyrene Stripwell™ plates, Costar Cat #2388).[5] A short polyethylene glycol chain linker is introduced between the plate surface and the photochemical tag to improve biocompatibility and reduce non-specific binding. All

steps of the plate preparation were carried out at room temperature. These plates were validated using direct observation of an immobilised dye molecule (unpublished work), as well as antigen-antibody immobilisation methods[5, 11] and phage display selection methods.[11]

5.1.5 Solid Phase Techniques on resin Beads

Organic synthesis on the solid phase, whether towards supported reagents, combinatorial libraries of chemicals or preparation of easily purified compounds, has been intensively researched. Books by Seneci[24] and Dorwald[25] provide an in depth background to the use of resins, surfaces and beads to prepare functional chemicals.

The use of solid phase to prepare supported compounds allows purification of prepared resin by filtration and washing in a solid phase extraction cartridge with a polyethylene frit. Washing resins with sequential solvents providing good swelling properties followed by poor swelling properties allows the removal of all material not bound to the surface of the resin. Swelling is an important property in the preparation and use of resin beads, as permeation of the reagents into the beads to maximise the yield and avoid the accumulation of previous intermediates requires good swelling.[25]

When a library is constructed on beads, beads must swell efficiently in both organic and aqueous solvents, must have good mechanical stability, suitable density of functional sites and low non-specific protein-binding properties. If one wishes to screen against fluorescently labelled proteins, then beads must have low intrinsic fluorescence at the emission maximum of the label.[26]

The problem of non-specific adsorption can be avoided by pre-coating the surface with a material that is resistant to protein adsorption; such materials are typically

non-ionic and very hydrophilic.[27] Polyethylene glycol groups can be used for this purpose,[28-30] and the Tentagel™ resin (Rapp Polymere GmbH) is a grafted copolymer of a low crosslinked polystyrene polymer onto which polyethylene glycol is grafted. However, Tentagel™ beads show considerable intrinsic fluorescence and batch variation in fluorescence which would be detrimental to any fluorescence assay.[26]

The addition of cleavage chemistries to the construction of the surface-supported moiety allows removal of the prepared molecule from the surface. This gives the opportunity for solution phase analysis after each reaction step. A substantial review of available cleavage methods from solid support has been reported by Guillier.[16] Cleavage also allows the estimation of loading as a percentage by weight of the resin.

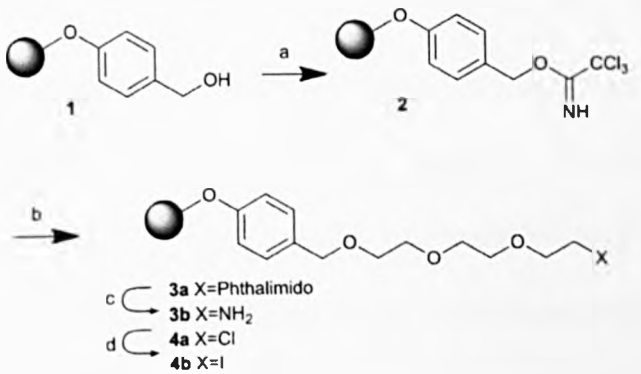
Rather than use a 96-well plate format as preferred for biological assays, the new cleavable analogues were prepared on polystyrene 'Wang' resin beads.[15] Cleavage from Wang resin has previously[18] been possible in sufficient quantity to enable analysis of the products by standard solution phase techniques.

5.2 Results and Discussion

5.2.1 Wang Resin

Several differently functionalised triethylene glycol (TEG) linkers were added to commercial 'Wang' resin **1**, thereby forming acid-, or mild oxidatively labile[31] ether functionality (Scheme 1). TEG has been shown to resist non-specific binding of peptides[28, 29] including those displayed on phage in particular.[30] This linker is also expected to allow greater mobility of the covalently derivatised drug molecule at an interface when assaying biological activity.

Wang resin was activated to nucleophilic displacement through formation of the trichloroacetimidate moiety 2.[32, 33] This was then displaced by an alcohol functionality in the presence of Lewis acid to give TEG phthalimide **3a** and TEG chloride **4a**.



Scheme 1 *Reagents and conditions*: (a) CCl_3CN, DBU, DCM, 0 °C, 1 h; (b) $H(OCH_2CH_2)_3X$, $BF_3{\cdot}OEt_2$, DCM, cyclohexane, r.t., 15 mins; (c) $H_2NNH_2{\cdot}H_2O$, THF, EtOH, 65 °C, 2 h; (d) acetone, NaI, 55°C, 48 h in dark.

Preparation of **3a** and its subsequent transformation to **3b** under Gabriel conditions (Scheme 1), together with preparation of **4a** and transformation to **4b** under Finkelstein conditions (less than 100% conversion) gave an acceptable mass yield of cleaved material using 5% trifluoroacetic acid (TFA) in dichloromethane. These yields correspond to an average of 66 mg/g resin recovery, equivalent to an average 0.3 mmol/g loading from the initial 0.9 mmol/g Wang resin. The treatment with trifluoroacetic acid under anhydrous conditions (*vide infra*), meant that the product alcohols appeared primarily as the corresponding trifluoroacetate esters, a useful ^{1}H NMR spectroscopy indicator since the β-methylene protons generally appear at 4.4 ppm ($CDCl_3$, 400 MHz), well separated from the other ethylene glycol peaks at 3.6 - 4.0 ppm.

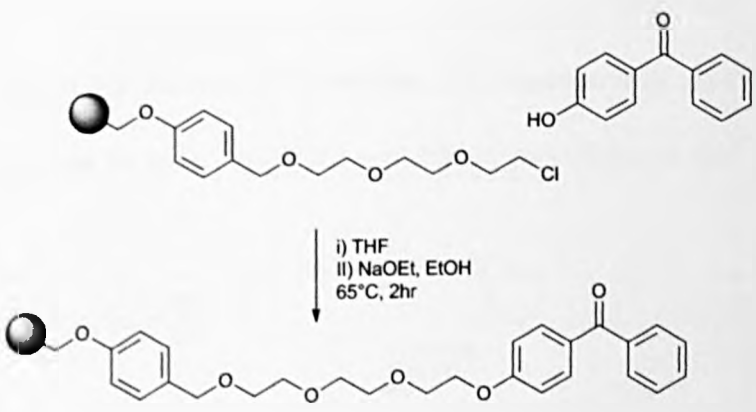
5.2.2 Photochemical preparation

Two benzophenone chemistries **5**, **6** and two azide chemistries **7**, **8** were subsequently prepared, although the methods used on polystyrene 96-well plates did not translate efficiently to the Wang-functionalised polystyrene beads. Plates

do not tolerate solvents with good swelling properties, whereas the polystyrene beads require them[25] to allow permeation of the reactants. In particular, substitution of the chloro-TEG linker by photochemistry **6** mediated by sodium ethoxide[34] was found to be inefficient on the beads, but transformation to the corresponding iodide and exposure to potassium carbonate at high temperature enabled the reaction to proceed with satisfactory conversion.

Preparation of **6** - the ether bound benzophenone - on chloride functionalised triethyleneglycol Wang resin was initially unsuccessful under the experimental conditions used for the plate surface. The plate surface reaction is carried out using a solution of in situ prepared sodium ethoxide in ethanol. The reaction of the hydroxybenzophenone with the chloride functionalised surface takes place at room temperature over four hours.[5]

The preparation of the chloride functionalised triethyleneglycol Wang resin was confirmed by NMR spectroscopy and mass spectrometry. However, the preparation of the ether bound benzophenone showed very low yield after cleavage from the surface (less than 10%).



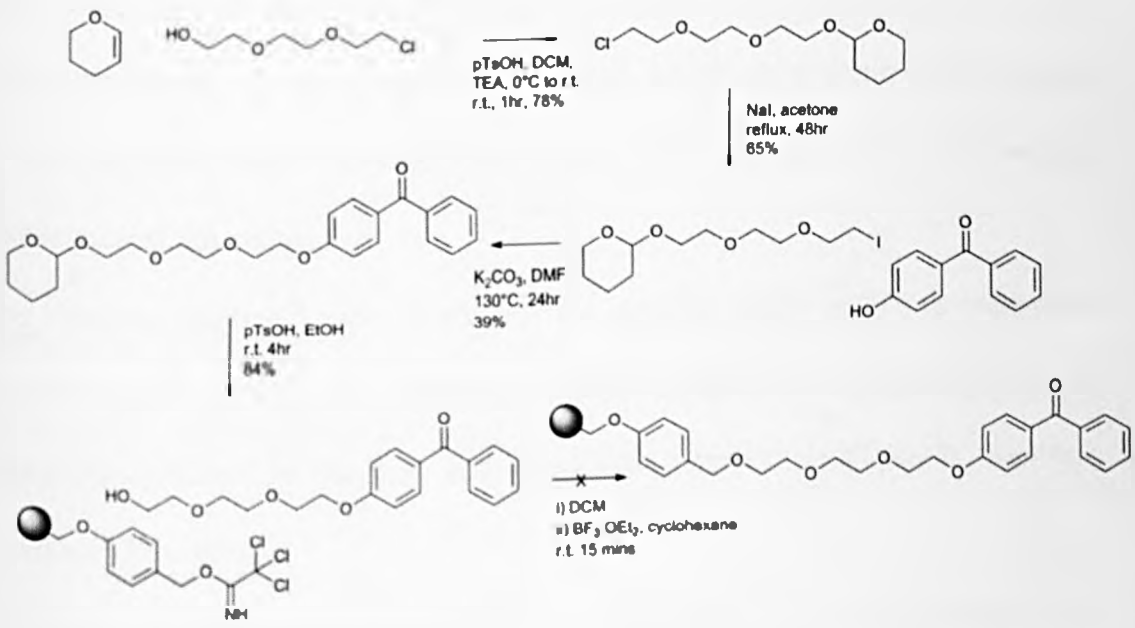
Scheme 5.2.2.1: Low yielding sodium ethoxide preparation of ether benzophenone photoactive chemistry

To try to improve this yield, several different techniques were attempted.

Initially a conversion from chloride to the more reactive iodide using a Finkelstein reaction was attempted (Scheme 1, **4a** to **4b**). This proceeded in 60-80% yield (by cleavage), with the remaining 20-40% being clean chloride functionalised resin.

This iodide product was then reacted with the hydroxybenzophenone following the initial procedure. Both the original chloride and the iodide resins were also reacted with the use of K_2CO_3 in DMF[35] and NaH in DMF (65°C, 24h). The products of these reactions were observed by cleavage and NMR/MS analysis, and showed equivalent yield to what had been seen before, <10%.

In an attempt to immobilise the correct group on the surface in good yield, the preparation of the tetrahydropyranoxyethoxyethoxyethoxybenzophenone in solution was carried out on the iodide functionalised THP protected triethylene glycol. For experimental details see Appendix 8.10. [36] This preparation was attempted initially with sodium ethoxide in ethanol, which showed low yields (>30%) due to substantial deprotection of the THP group and elimination of the iodide. The reaction was also attempted with potassium carbonate as base,[35] following a related procedure in the



Scheme 5.2.2.2: Attempted solution preparation for the ether benzophenone photoactive functionality for direct immobilisation

literature. This latter procedure gave better yield (39%) with considerably less elimination on first attempt.

The remaining protected solution phase product was deprotected with para-toluenesulphonic acid, followed by purification by column chromatography (83%). The aim was to immobilise this hydroxyl functionalised molecule onto the trichloroacetimidate functionalised surface using Hanessian chemistry.[32]

The immobilisation was carried out using trifluoroborane diethyletherate, and has previously in this project gone to good yield (Scheme 1, 33% loading by cleavage), allowing immobilisation of PEG-chloride and PEG-phthalimide groups on the Wang resin surface. Upon cleavage this reaction showed no product on the surface.

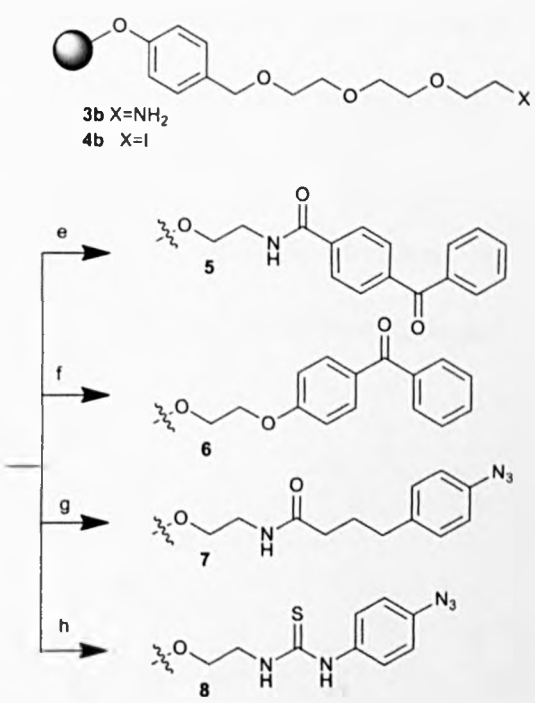
The reaction of a triethylene glycol molecule with the trichloroacetimidate gives a terminal hydroxyl group which can be activated using a toluenesulphonyl group (Botros et al.[37]) and exchanged for a hydroxybenzophenone molecule (Fürstner et al.[38]). This reaction was carried out alongside the previous preparation.

This reaction was successful, but it appears that the reaction of the activated Wang resin with the triethylene glycol molecule has produced a mixture of singly attached triethylene glycol and doubly attached, which gives some contamination to the product upon cleavage, but does not interfere with the reaction (experimental not shown).

The literature reaction[35] used to produce the solution phase protected triethylene glycol benzophenone was attempted on iodide functionalised Wang resin using potassium carbonate as the base and considerably higher temperatures than had previously been used.

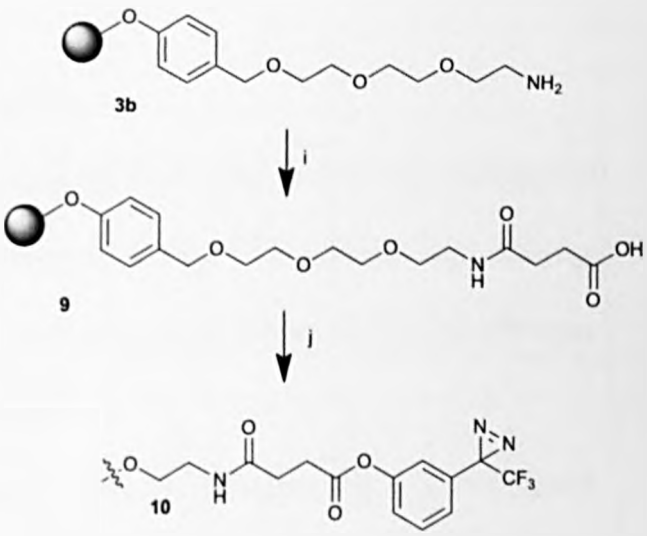
Despite some initial experimental difficulty using such high temperatures with the resin, this process was successful. The reaction was initially attempted with the

chloride resin, but while it was predominantly the expected product, it showed a mixture of products on the surface under NMR and MS. The iodide gave the correct product on the surface with some (6:1) starting material observed by fluorine NMR spectroscopy. This final product was used in further interaction studies.



Scheme 2 *Reagents and conditions, (in dark)*: (e) **3b**, benzoylbenzoic acid, EDC, DMAP, DMF, 65 °C, 48 h; (f) **4b**, hydroxybenzophenone, K_2CO_3, DMF, 130 °C, 48 h; (g) **3b**, azidophenylbutyric acid, EDC, DMAP, DMF, 65 °C, 48 h; (h) **3b**, azidophenylisothiocyanate, DCM, r.t., 24 h.

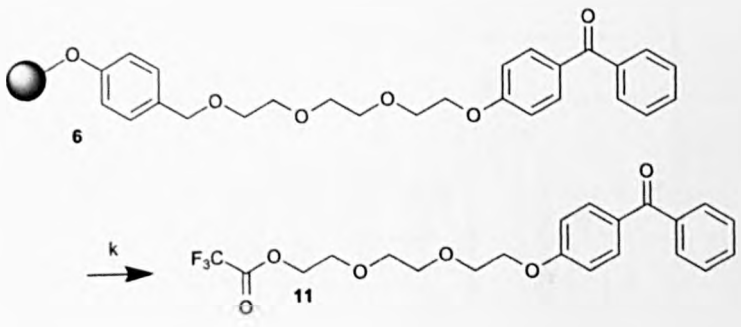
The remaining photochemistry used in the original Magic Tag® set was the 3-(3-trifluoromethyl-3*H*-diazirin-3-yl)phenyl group. The method for preparing the diazirine is based on the method in the 1994 paper by Hatanaka et al.[39] The diazirine photochemistry is highly light sensitive, and the final two steps of this synthesis are carried out in the dark to avoid premature activation to the carbene. (Experimental in Appendix 8.10).



Scheme 3 *Reagents and conditions:* (i) succinic anhydride, DIPEA, MeCN, DCM, r.t., 20 h; (j) 3-(3-hydroxyphenyl)-3-(trifluoromethyl)-3*H*-diazirine, EDC, DMAP, DCM, 30 °C, 48 h, in dark.

Addition of succinic anhydride to the TEG linker **3b** gave amidosuccinic acid **9**, allowing successful coupling of the diazirine with EDC to give **10**.[39] This was also carried out in the dark to protect the diazirine functionality.

Acid-mediated removal of the TEG-linked photochemistries (for example Scheme 4) provided trifluoroacetate esters such as **11** with suitable properties for unambiguous identification by NMR spectroscopy and LC-MS. This was the case for all photochemistries except supported thiourea azide **8**, produced by coupling to the isocyanate, which gave lower than expected conversion (Scheme 2 - h) and whilst acid-mediated cleavage gave a satisfactory accurate mass measurement the NMR spectrum was ambiguous and showed contamination by side-products not obeserved for **5-7**.



Scheme 4 *Reagents and conditions:* (k) 5 % TFA in DCM, r.t., 1 h.

5.2.3 Analysis of results on the solid phase

On the solid phase, analysis of products prepared on wang resin beads consisted of KBr disk IR, gel phase NMR spectroscopy (^{13}C and ^{19}F), nitrogen and chlorine elemental analysis and colorimetric tests including the Kaiser test for primary amines and base catalysed fluorescein test for halides.

Using KBr pellets allowed a greater number of resolved peaks to be observed out of the background of the resin, allowing observation of the characteristic peaks for each of the materials (see comparison figs 5.2.3.6 and 7 below). IR spectroscopy allowed a fast analysis of the beads after reaction, as the characteristic peak from the starting material of each product disappears and new characteristic peaks appear. Due to intermolecular interactions, band positions in solid state spectra are often different

from those of the corresponding solution spectra. This is particularly true of those functional groups which take part in hydrogen bonding.[40]

Shown below is the reference KBr disk IR of polystyrene (Figure 5.2.3.1).[41]

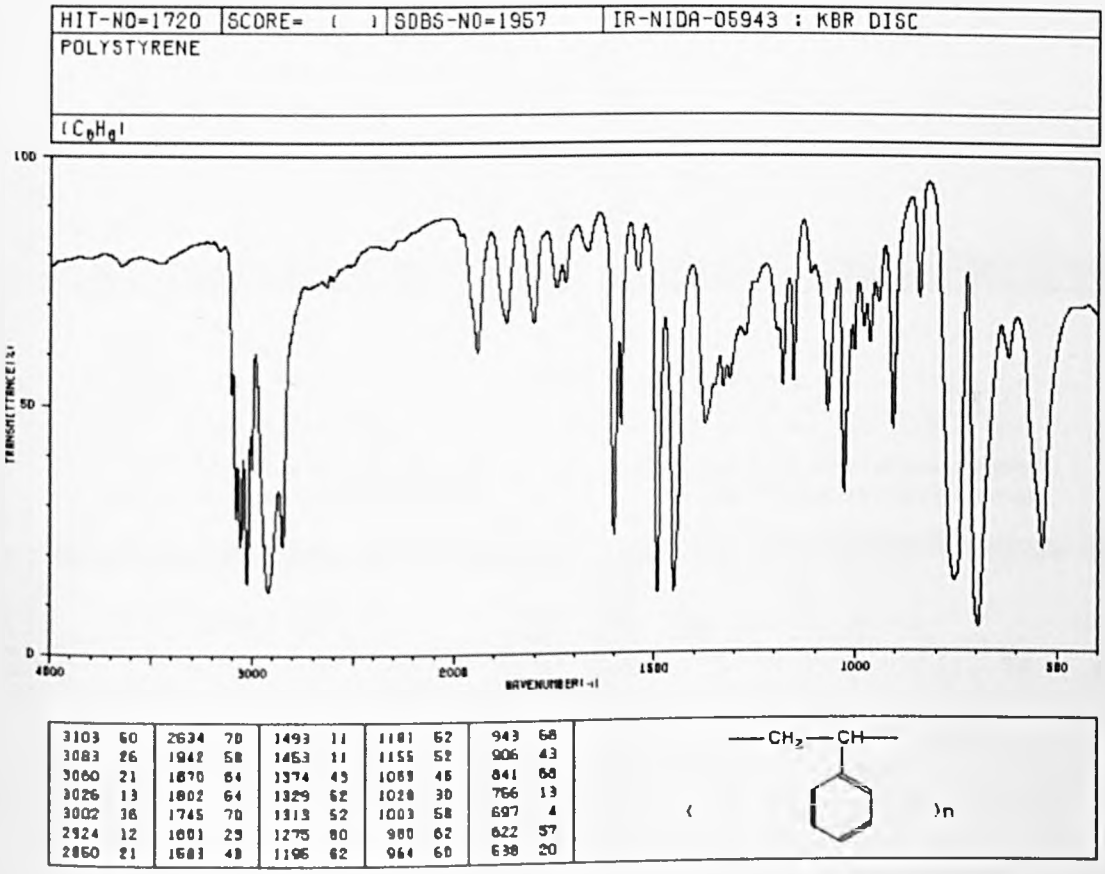
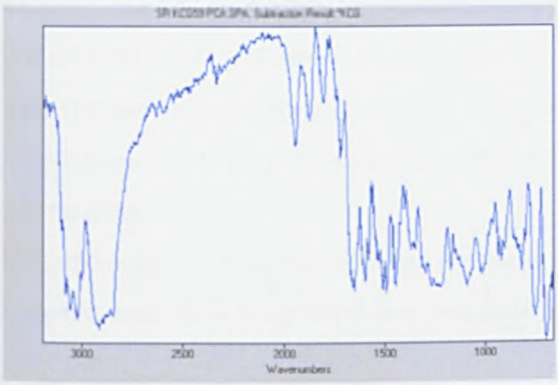


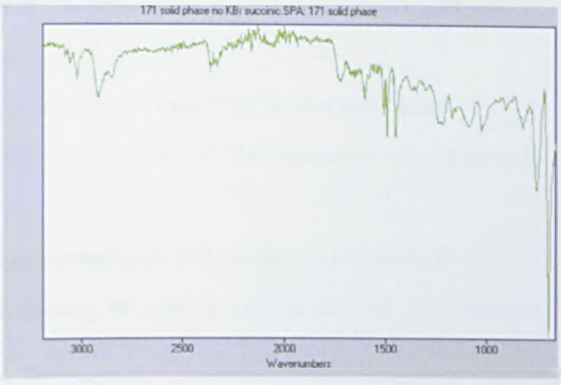
Figure 5.2.3.1: Reference KBr disk IR spectrum of polystyrene (SDBS reference)

The KBr disks IR spectra for the products look very similar at first glance. Shown below is a KBr disk IR spectrum for one of the five photochemistries **5** (Figure 5.2.3.2), the spectrum of the butanoic acid compound without the use of the KBr pellet (Figure 5.2.3.3), and two spectra taken of photochemistry **8**, Figure 5.2.3.4 taken with low grade KBr and Figure 5.2.3.5 taken with higher quality KBr.

Figure 5.2.3.2: Amide benzophenone **5** KBr disk IR spectrum

Figure 5.2.3.3: Amide succinic acid **9** solid resin IR spectrum





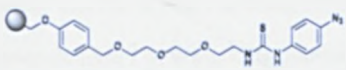
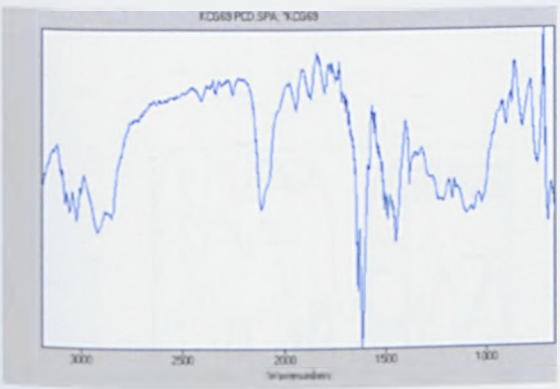
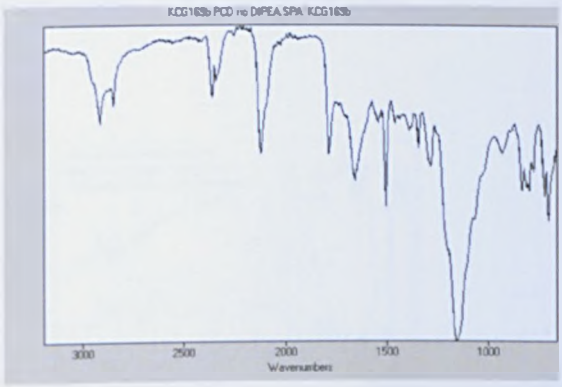


Figure 5.2.3.4: Thiourea phenyl azide **8** KBr disk IR spectrum (low quality KBr)

Figure 5.2.3.5: Thiourea phenyl azide **8** KBr disk IR spectrum (high quality KBr)





Changes in IR spectra were recorded as follows (see Figure 5.2.3.6 and Figure 5.2.3.7):

Raw Wang resin 1 funtionalised to trichloroacetimidate 2: Appearance of peaks at 1663 (ketimine)

Wang trichloroacetimidate 2 to Wang-TEG-phthalimide compound 3a: Disappearance of peak at 1663, appearance of peaks at 1773 and 1708 (-CO-N-CO- 5 membered ring) and 1109-1080 (ether)

Wang-TEG-phthalimide 3a to Wang-TEG-amine 3b: Disappearance of peaks at 1773 and 1708, no other peaks appear

Wang-TEG-amine 3b to Wang-TEG-amide benzophenone 5: Apperance of peaks at 1657 (benzophenone C=O and amide I) and 1535 (amide II)

Wang-TEG-amine 3b to Wang-TEG-amide butyl phenylazide 7: Appearance of peaks at 2110 (azide), 1655 (amide I)

Wang-TEG-amine 3b to Wang-TEG-thiourea phenylazide 8: Appearance of peaks at 2123 (azide) and 1657 (thiourea RNHCSNHR)

Wang-TEG-amine 3b to Wang-TEG-amidesuccinamic acid 9: Appearance of peaks at 1725 (CO_2H) and 1661 (amide I)

Wang-TEG-amide succinamic acid 9 to Wang-TEG-succinamic acid 3-(3-trifluoromethyl-3*H*-diazirin-3-yl)-phenyl ester 10:

Wang trichloroacetimidate 2 to Wang-TEG-chloride 4a: Disappearance of peak at 1663, appearance of peaks 1130-1080 (ether)

Wang-TEG-chloride 4a to Wang-TEG-iodide 4b: No visible changes by IR

Wang-TEG-iodide 4b to Wang-TEG-ether benzophenone 6: Appearance of peak at 1650 (benzophenone C=O)

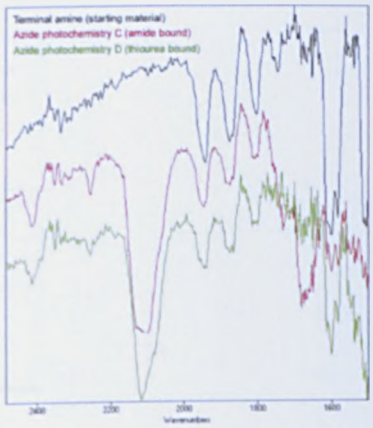


Figure 5.2.3.6: Infrared analysis (KBr disk) of the reaction scheme from **1** to **5**

Figure 5.2.3.7: Infrared analysis (KBr disk) of the azide chemistry ($2100cm^{-1}$) in starting material **3b** and **7** and **8**

Gel phase carbon NMR spectroscopy allowed direct analysis of the functionalised resins before cleavage. By comparing the cleaved proton and carbon 1D and 2D NMR spectra to the starting material gel phase ^{13}C NMR spectra it was possible to assign many of the peaks from the gel phase ^{13}C NMR spectrum. To obtain gel phase ^{13}C NMR spectra from the solid phase product, the dry beads were placed in the NMR tube and chloroform was added until all beads were fully swollen, floating at the top of the solvent. By controlling the volume of solvent, the height of the beads was controlled to allow optimum positioning for the NMR probe. To discern individual peaks, NMR spectra were run on a 500MHz NMR machine, with over 12,288 scans at room temperature (298K). These NMRs were carried out by the Warwick NMR Service.

Figure 5.2.3.8: Gel phage carbon NMR spectra showing the starting material Wang resin and the added functionalities

Fluorine NMR spectroscopy was primarily used as a diagnostic test for immobilisation (see chapter 5.2.1), but was also used to validate the chemistry of photochemistry **10** on the surface.

Elemental analysis was used as a crude test for the presence or absence of nitrogen and chlorine.[42, 43] Due to batch-to-batch changes, it would be very difficult to

calculate loading based on the elemental analysis of the beads. Results from these assays are shown below.

Table 5.2.3.1 - 2: Elemental analysis showing change in the percentage of nitrogen and chlorine in reaction steps

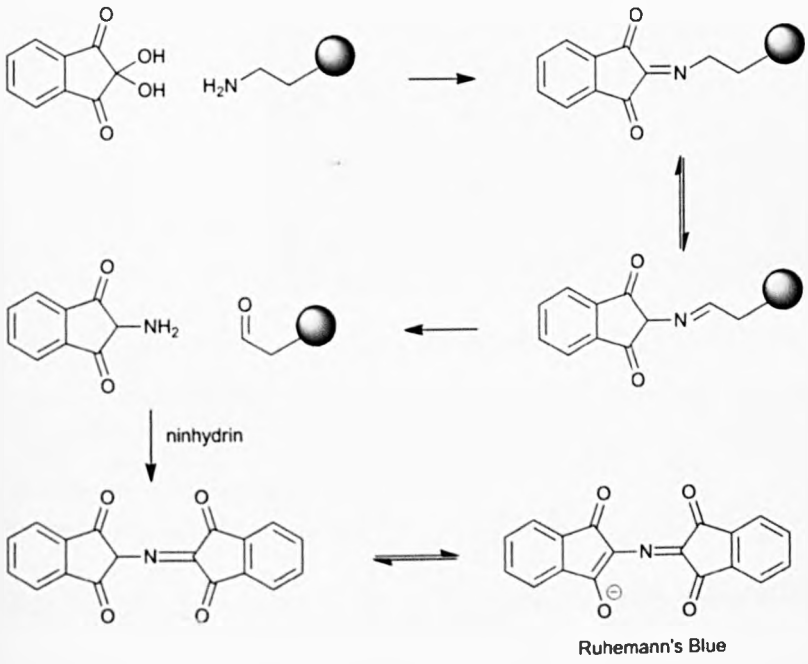
Compound	N (% wt/wt)
1	0.0
2	1.2
3a	0.7
3b	0.5
4a	0.1
4b	0.2

Compound	Cl (% wt/wt)
1	0.1
2	9.1
3a	0.8
3b	0.6
4a	2.8
4b	2.4

In 2004, Gagginni et al published a paper on colorimetric tools for solid phase organic synthesis.[44] Two of these methods - the test for primary amines (Kaiser test, or ninhydrin test), and the base-catalysed fluorescein assay were used on the beads to assay functionalisation.

Kaiser test:

0.2M ninhydrin in ethanol and 4M phenol in ethanol solutions were prepared, and four drops of each were mixed with roughly 100 mg beads under test, followed by four drops of pyridine. This was heated to 100°C for two minutes, and the colour noted and recorded.



Scheme 5.2.3.1: Kaiser test, showing the interaction of ninhydrin with primary amines to give the highly coloured ruhemann's blue

A positive result in the Kaiser test is a blue solution and blue beads, as some of the anionic chromophore Ruhemann's Blue is retained by the resin. This test was carried out on the phthalimide and amine chemistries, as well as the products of the amine resin, photochemistries **5**, **7** and **8** to determine completeness of reaction.

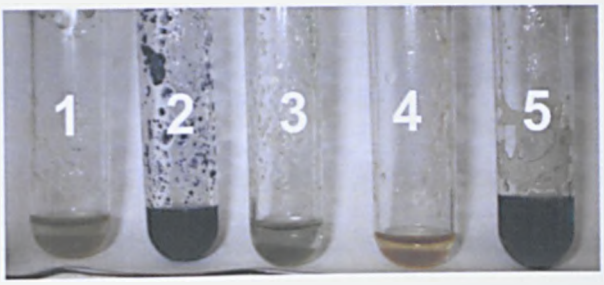
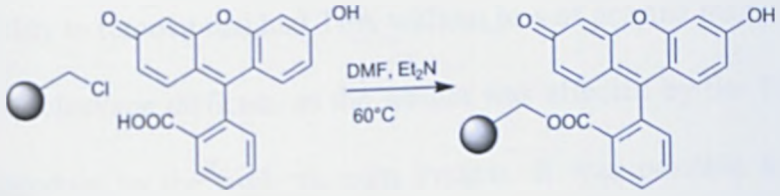


Figure 5.2.3.9: Kaiser test on functionalised beads. Positive result is blue beads and blue solution (see tube 2)

From left to right, the reactions are: 1) Wang resin immobilised ethoxyethoxyethylphthalimide **3a**, 2) Wang resin immobilised ethoxyethoxyethylamine **3b**, 3) Wang resin immobilised ethoxyethoxyethylamide benzophenone **5**, 4) Wang resin immobilised ethoxyethoxyethylamidebutylphenyl-4-azide **7**, 5) Wang resin immobilised ethoxyethoxyethylthioureaphenyl-4-azide **8**

As can be seen from Figure 5.2.3.9, the amine gave a positive result (blue beads and blue solution), but the thiourea phenylazide also gave a blue solution. The beads remained a pale yellow colour in this assay. This may be due to a limited reduction of the azide chemistry to primary amine, or due to incomplete reaction of the amine to form the thiourea.

Base-catalysed Fluorescein test:



Scheme 5.2.3.2: Base catalysed fluorescein test

To 100 mg beads in 1 ml dimethylformamide was added 1 drop triethylamine and 1 mg fluorescein. The reaction was heated at 60 °C for 15 minutes. The beads were washed and fluorescence was measured against a control sample.

This reaction was carried out on the Wang resin immobilised ethoxyethoxyethylchloride beads and the Wang resin immobilised

ethoxyethoxyethylphthalimide beads as comparison. This gave a positive result, but only by a small margin.

Table 5.2.3.3: Change in fluorescence reading after base catalysed fluorescein test		
Base-catalysed Fluorescein test	Relative fluorescence -Phth group (control)	Relative fluorescence -Chloride group
Before Assay	2549	3367
After Assay	3474	4936
Difference	**925**	**1569**

5.2.4 Analysis of results in the solution phase

Cleavage was carried out using 5% TFA in DCM, resulting in the TFA ester on the cleaved alcohol. The cleavage method was carried out twice on the beads, to ensure the maximum quantity of product is removed from the surface. Removal of residual TFA can be carried out using THF azeotrope under high vacuum, but this was found to lead to some loss of product and so this has been avoided where possible.

The inability to remove residual TFA without loss of product made calculation of the loading by cleavage difficult, as the weight was affected by the TFA present or by loss of product to the high vacuum system. It was possible to calculate some loadings by this method, but they are expected to be underestimates due to possible loss of material.

All products of cleavage from the solid phase were analysed using standard solution phase techniques. NMR spectra of materials removed from the resin by acid-mediated cleavage were not as clean as solution phase spectra, but had the advantage of diagnostic β-methylene protons for the TFA ester, which appeared at 4.4 ppm ($CDCl_3$, 400 MHz), well separated from other ethylene glycol peaks at 3.6 - 4.0 ppm.

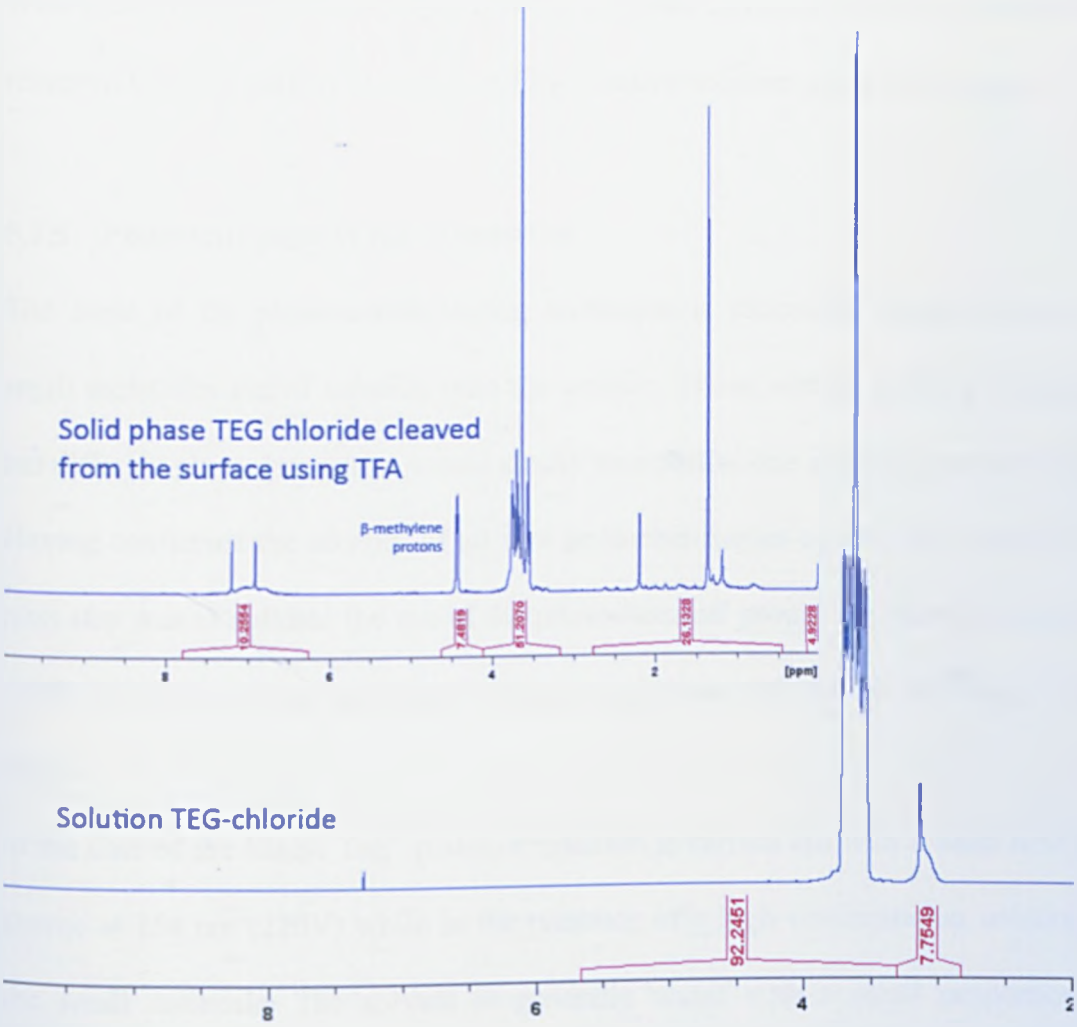


Figure 5.2.4.1: Proton NMR spectra comparing the solution phage triethyleneglycolchloride and 2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethoxychloride after cleavage of triethyleneglycolchloride from the Wang resin surface

The isothiocyanate chemistry **8** was successfully identified by IR spectroscopy on the surface, and also by mass spectrometry and IR spectroscopy after cleavage from the surface, but could not be observed by NMR spectroscopy. There was evidence of starting material peaks, but not in sufficient quantities to suggest that the reaction may not have occurred.

The reason for this result has not yet been entirely identified, but is thought to be due to a combination of the azide group 'biting back' onto the resin after activation, which would explain the observed low loading results, and cleavage of the thiourea bond under treatment with TFA.

With these results we have shown that the prepared molecule could be successfully removed from the surface and analysed by standard solution phase techniques.

5.2.5 Photoactivation of the chemistries

The basis of the photoimmobilisation technique is successful immobilisation of small molecules out of solution onto the surface. These will be varied in structure, but different photochemistries would ideally immobilise one structure preferentially. Having confirmed the identity of all five photochemistries on the resin surface, the next step was to validate the use of the photochemical groups for immobilisation of small molecules, in an analogous manner to that carried out on the Magic Tag® plates.

In the case of the Magic Tag® plates, irradiation is carried out with a hand held UV source at 254 nm (220V) while in the presence of a high concentration solution of the small molecule. The solvent is generally water with a small proportion of DMSO.

Molecule immobilisation on the beads can be assessed by ^{19}F NMR spectroscopy in the instance of fluorine-containing compounds such as flurbiprofen, fluorescence in the case of molecules such as fluorescein or commercial antibody assays in the case of known antigens such as digoxin, and biotin. These assays can be used to explore immobilisation, and further removal of material by TFA cleavage would allow analysis of functional group selectivity upon UV irradiation.

Organic solvents tend to have good swelling properties which are not well tolerated with polystyrene plates. However, small bioactives can have very different solubilities in aqueous solvents. One option for laying the beads out for irradiation is to use a 96-well plate – maximising the surface area of beads exposed to the

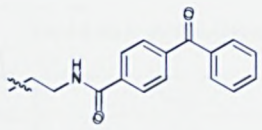
irradiating light. This requires solvents compatible with the plastic surface, which can limit the reaction conditions. However, in the case of reactions assayed by fluorescence, the products can be assayed directly from the plate, which can be advantageous as it avoids transferring the beads between vessels.

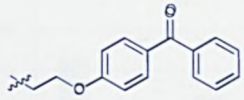
Irradiation of beads was optimised to maximise the exposure of the beads. Irradiation of beads in SPE tubes allowed vigorous agitation, but the positioning of the UV lamp made it difficult to get even exposure across each reaction. By spreading the beads over a number of wells in a 96-well filter plate the surface exposed to the UV was maximised, and the shallow depth of the wells meant that the UV source could be closer to the beads themselves. A disadvantage with using a 96-well plate was that there was limited movement in the beads, even with vigorous agitation, as the use of plates requires poorly swelling solvents to avoid damaging the plate. To increase the exposed bead surface further a shallow watchglass was used in the photoreactor oven with a small amount of a highly concentrated solution of the small molecule. No agitation was used in this method, but the base of the oven was rotated to give even UV exposure to each of the reactions.

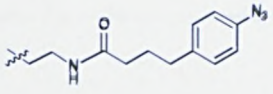
Two different UV irradiation methods and sources were used, a hand held TLC lamp and a 16-bulb Luzchem ICH-2 photoreactor oven. Wavelengths of 254 nm and 365 nm were used with the hand held lamp, 220 V, one 6 W bulb, and the photoreactor oven irradiated at a range of wavelengths between 316 and 400 nm (UVA), sixteen 8 W bulbs. Under the hand held lamp the beads were irradiated in a 96-well plate, either after soaking in the small molecule solution with the solution removed before irradiation or whilst in the presence of the aqueous solution. The swelling properties of aqueous solutions are very poor and so the beads were more consistently and evenly irradiated when the solution had been removed.

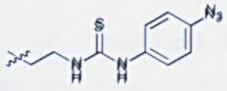
Table 5.2.5.1: Immobilisation Experiments and their Initial Assay Type				
Resin Chemistry	Molecule	Solvent	Irradiation type	Assay Type
3a, 3b, 4a **5, 6, 8**	Fluorescein	THF or DCM	Hand held 254 nm, 10 mins	Fluorescence
5, 6, 7, 8	Biotin	H_2O	Hand held 254 nm, 10 mins and 1 hr	Antibiotin-FITC
2, 3a, 3b, 4a, 4b, 5, **6, 7, 8, 9, 10, 12**	Digoxin	PBS	Hand held 254 nm, 30 mins dry	DI-22 FITC
5, 6, 7, 8	Flurbiprofen	DMSO	Hand held 254 nm, 10 mins	^{19}F gel phase
5, 6, 7, 8, 10	Biotin	DMSO	Photoreactor 316-400 nm, 100 mins	Antibiotin-FITC
5, 6, 7, 8, 10	Flurbiprofen	DMF or DMSO	Photoreactor 316-400 nm, 10 mins, 100 mins, 17 hrs	^{19}F gel phase

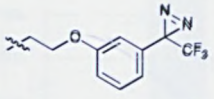
Table 5.2.5.2: Immobilisation Results Collected and Summarised		
Chemistry	Small molecule	Result
3a, 3b, 4a 5, 6, 8	Fluorescein	**3a**, **3b** and **4a** beads also assayed (0.5 ml 1 mg/mL THF only). Control against no fluorescein present. Positive result apparently on **5**, but not on **6** or **8**. No visible additional peaks in proton NMR or MS peaks for immobilisation reactions.
5, 6, 7, 8	Biotin	Control against no antibody present. Immobilisation visible in **5**, **6** and **7** at conc. of 1:500 antibody solution, poor at higher or lower antibody conc.. No immobilisation in **8**. No visible additional peaks in proton NMR or MS peaks for immobilisation reactions.
2, 3a, 3b, 4a, 4b, 5, 6, 7, 8, 9, 10, 12	Digoxin	Most bead types assayed. Control against no digoxin present. Very inconsistent results across all bead types. Control against covalently prepared digoxin bead **12** allowed more comparative assay, continued to be highly inconsistent. No visible additional peaks in proton NMR or MS peaks for immobilisation reactions.
5, 6, 7, 8	Flurbiprofen	MS and NMRs show no change from starting materials. No NMR evidence for product **8**
5, 6, 7, 8, 10	Biotin	**5**, **6**, **7** show no change by ^{1}H NMR, **7** and **8** show some loss of N_2 by MS, **10** shows change by NMR and MS but no **10**+biotin peak in MS
5, 6, 7, 8, 10	Flurbiprofen	**5**, **6**, **7** show no change by ^{1}H NMR, **7** and **8** show some loss of N_2 by MS, **10** shows change by NMR and MS but no **10**+flurbiprofen peak in MS

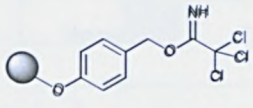
5 = 

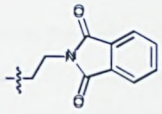
6 = 

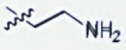
7 = 

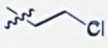
8 = 

10 = 

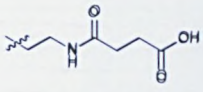
2 = 

3a = 

3b = 

4a = 

4b = 

9 = 

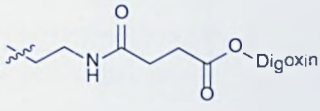
12 (appendix 8.10) = 

Chart 5.2.5.1: UV Immobilisation of Fluorescein on functionalised Wang beads

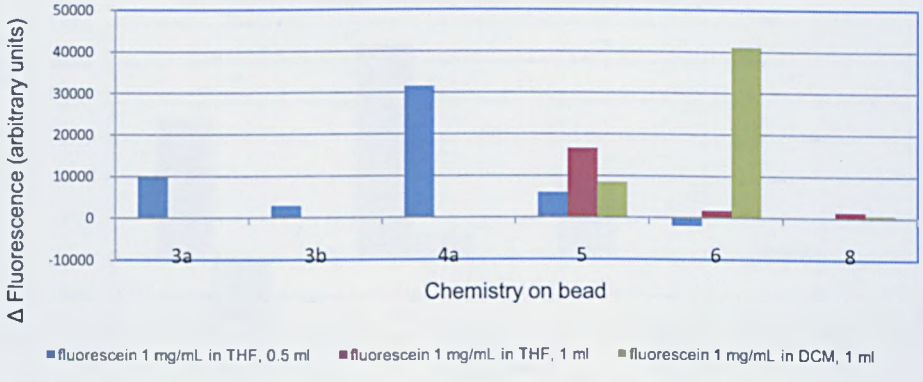


Chart 5.2.5.2: UV Immobilisation of Fluorescein on functionalised Wang beads (Repeat)

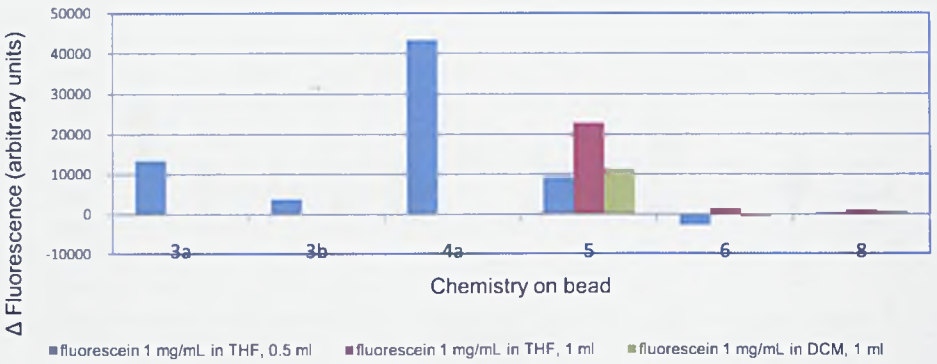


Chart 5.2.5.3: Antibiotin-FITC assay against immobilised biotin (10 mins)

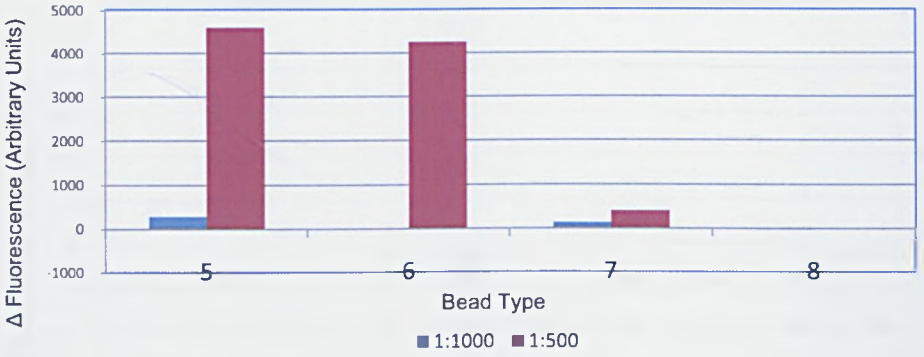


Chart 5.2.5.4: Antibiotin-FITC assay against immobilised biotin (1 hour)

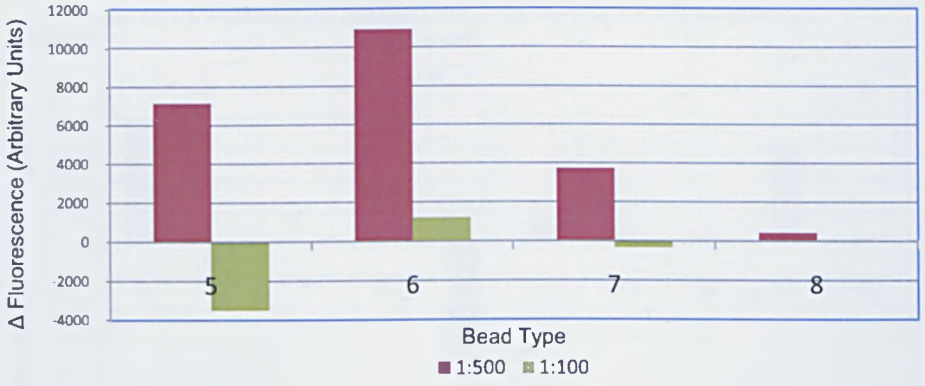
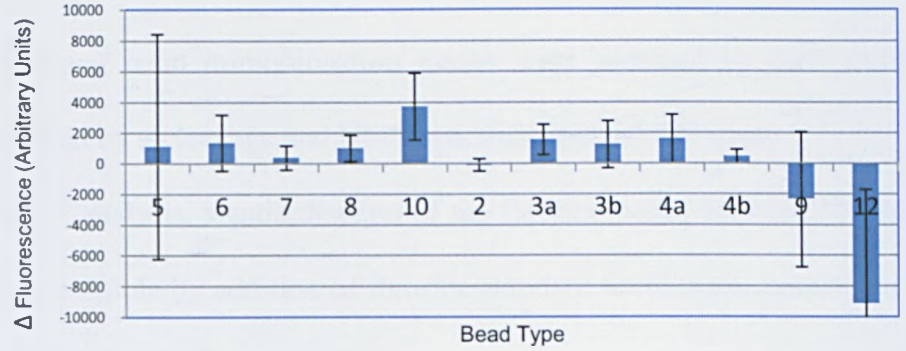


Chart 5.2.5.5: Antidigoxin-FITC probing Digoxin-immobilised Wang Beads



Repeat

Chart 5.2.5.6: Antidigoxin-FITC probing Digoxin-immobilised Wang Beads

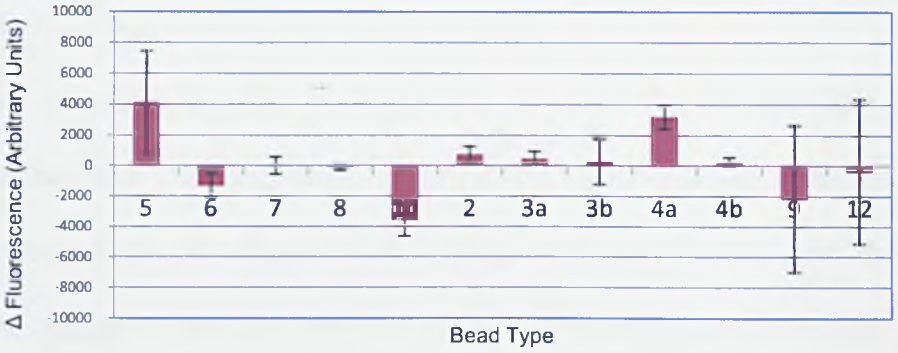
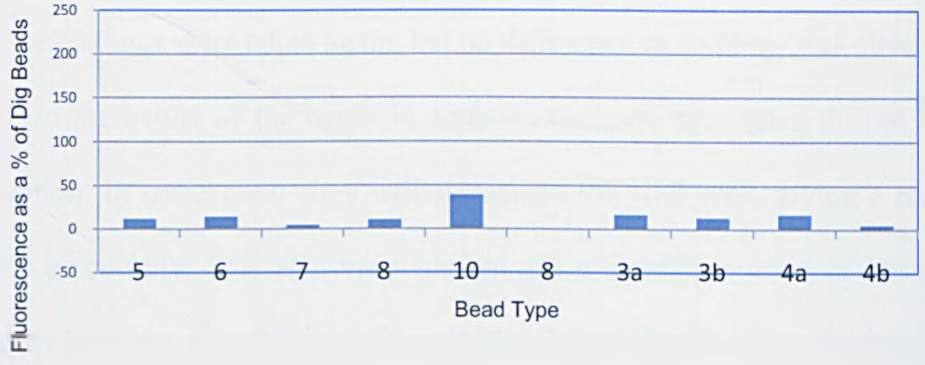
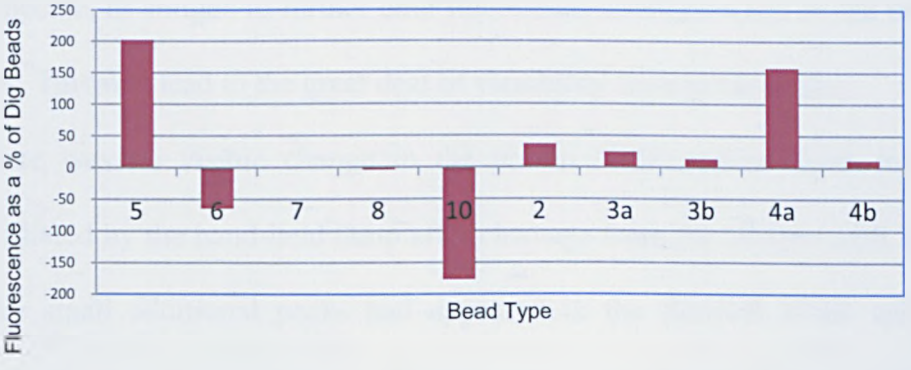


Chart 5.2.5.7: Fluorescence of Antidigoxin-FITC probed Beads as a % of Prepared Digoxin Beads



Repeat

Chart 5.2.5.8: Fluorescence of Antidigoxin-FITC probed Beads as a % of Prepared Digoxin Beads



All Wang resin immobilisation assays were assessed by their initial assay type, followed by a cleavage and NMR spectroscopy and MS assay.

For ^{19}F analysis, standardisation of the fluorine peaks between NMR machines was made possible by addition of fluorine standard trichlorofluoromethane (0 ppm). The

sample-to-sample variation in this measurement was determined to average 0.1 ppm after standardisation.

The fluorescence plate reader has no definition file for the filter plates used to process these samples, which could potentially have led to incorrect readings. The Nunclon plate, for which there is a plate definition file, matches the filter plate in size exactly but does not have the white base caused by the filter piece.

In an attempt to determine whether this issue would add error into the readings, the beads were transferred into a Nunclon plate after functionalisation in filter plates. These readings were taken again, but no difference in readings was observed.

Due to the nature of the beads in aqueous medium, they often do not lay flat after filtration. In some cases they will sit against the well wall, giving a reading of the filter piece only. It is also very hard to get a consistent mass of beads in a well, especially when transferring the beads dry. This will affect the output from any well, as if there are fewer beads there is less surface functionalised with the small molecule, or antigen to further bind fluorescent antibodies and so the reading will be less. This will lead to the great deal of variability seen in the wells.

There was no visible change in the proton NMR spectra from the compounds irradiated by the hand-held lamp after cleavage from the surface with TFA and only very small additional peaks had appeared in the fluorine NMR spectrum. Mass spectra were also collected for these cleaved products, and showed primarily starting materials, with no peaks for starting materials plus bound small molecule (or solvent), or starting material minus N_2 in case of azide or diazirine compounds.

From the compounds irradiated in the photoreactor oven it was hoped that increasing the power of the irradiation, as well as increasing the wavelength to bring it into line

with literature irradiation wavelengths,[6, 9, 13, 22, 45] would lead to more successful immobilisation of the small molecule.

By NMR spectroscopy resins **5**, **6**, **7** and **8** show small or no changes after irradiation. By mass spectrometry it was possible to identify some changes in the azide and diazirine resins, as well as some peaks which may suggest some immobilisation of solvent molecules. These have not been validated, and no peaks have been seen to account for the immobilisation of the desired molecules.

Tables 5.2.5.3 to .9, Photo-reactor Oven Immobilisation MS results

Reaction Conditions	Short ID
DMSO 10 mins	1
DMSO 100 mins	2
DMF 10 mins	3
11 mg/mL biotin in DMSO 100 mins	4
0.016 mg/mL fluorescein DMF 10 mins	5
1.6 mg/mL fluorescein DMF 10 mins	6
15 mg/mL flurbiprofen in DMF 10 mins	7
15 mg/mL flurbiprofen in DMSO 10 mins	8
15 mg/mL flurbiprofen in DMSO 100 mins	9
High conc. flurbiprofen in DMSO 17 hours	10

Resin **5**	Base Peaks (m/z)	Other Important Peaks (m/z)
Reaction 2	305 (unknown), 380 (SM Na^+)	358 (SM H^+)
Reaction 4	305 (unknown), 380 (SM Na^+)	358 (SM H^+)
Reaction 10	305 (unknown)	358 (SM H^+), 380 (SM Na^+)

Resin **6**	Base Peak (m/z)	Other Important Peaks (m/z)
Reaction 3	353 (SM Na^+)	331 (SM H^+), 427 (SM TFA H^+), 449 (SM TFA Na^+)
Reaction 1	369, 465 (unknown), 427 (SM TFA H^+)	331 (SM H^+), 353 (SM Na^+), 449 (SM TFA Na^+)
Reaction 2	331 (SM H^+), 427 (SM TFA H^+)	353 (SM Na^+), 449 (SM TFA Na^+)
Reaction 6	353 (SM Na^+)	331 (SM H^+), 427 (SM TFA H^+), 449 (SM TFA Na^+)
Reaction 5	353 (SM Na^+)	331 (SM H^+), 427 (SM TFA H^+), 449 (SM TFA Na^+)
Reaction 7	353 (SM Na^+)	331 (SM H^+), 427 (SM TFA H^+), 449 (SM TFA Na^+)
Reaction 10	369 (unknown)	331 (SM H^+), 353 (SM Na^+)

Resin **7**	Base Peaks (m/z)	Other Important Peaks (m/z)
Reaction 2	433 (SM TFA H^+)	337 (SM H^+), 359 (SM Na^+), 407 (TFA loss of N_2 gain of H_2 H^+)
Reaction 4	433 (SM TFA H^+)	337 (SM H^+), 359 (SM Na^+), 407 (TFA loss of N_2 gain of H_2 H^+)
Reaction 10	311 (loss of N_2 gain of H_2 H^+)	337 (SM H^+), 333 (loss of N_2 gain of H_2 Na^+)

Resin **8**	Base Peaks (m/z)	Other Important Peaks (m/z)
Reaction 3	302 (Phth Na^+), 617 (unknown)	246 (TFA NH_2 H^+), 298 (Loss of N_2 H^+), 326 (SM H^+), 348 (SM Na^+), 394 (TFA Loss of N_2 H^+), 422 (SM TFA H^+), 451 (unknown), 547 (unknown)
Reaction 1	326 (SM H^+), 422 (SM TFA H^+)	246 (TFA NH_2 H^+), 280 (Phth H^+), 302 (Phth Na^+), 348 (SM Na^+), 376 (TFA Phth H^+), 394 (TFA loss of N_2 H^+)
Reaction 2	326 (SM H^+), 422 (SM TFA H^+), 547 (unknown)	246 (TFA NH_2 H^+), 376 (TFA Phth H^+), 394 (TFA Loss of N_2 H^+), 610 (unknown), 621 (unknown), 643 (unknown)
Reaction 6	302 (Phth Na^+), 451 (unknown)	246 (TFA NH_2 H^+), 298 (Loss of N_2 H^+), 326 (SM H^+), 348 (SM Na^+), 422 (SM TFA H^+), 617 (unknown)
Reaction 5	302 (Phth Na^+), 617 (unknown)	326 (SM H^+), 348 (SM Na^+), 422 (SM TFA H^+)
Reaction 4	426 (unknown)	302 (Phth Na^+)
Reaction 7	422 (SM TFA H^+), 246 (TFA NH_2 H^+)	298 (Loss of N_2 H^+), 302 (Phth Na^+), 326 (SM H^+), 348 (SM Na^+), 394 (TFA loss of N_2 H^+), 451 (unknown), 547 (unknown)
Reaction 8	326 (SM H^+), 432 (unknown)	246 (TFA NH_2 H^+), 298 (Loss of N_2 H^+), 422 (SM TFA H^+), 617 (unknown)
Reaction 9	326 (SM H^+), 432 (unknown)	246 (TFA NH_2 H^+), 280 (Phth H^+), 298 (Loss of N_2 H^+), 302 (Phth Na^+), 422 (SM TFA H^+), 617 (unknown)
Reaction 10	298 (Loss of N_2 H^+), 617 (unknown)	280 (Phth H^+), 326 (SM H^+), 432 (unknown)

Resin **10**	Base Peaks (m/z)	Other Important Peaks (m/z)
Reaction 10	272 (Succinic acid Na^+), 424 (loss of N_2 gain of H_2O H^+)	246 (TFA NH_2 H^+), 250 (Succinic acid H^+), 328 (unknown), 346 (TFA Succinic acid H^+), 405 (loss of N_2), 440 (unknown), 446 (loss of N_2 gain of MeCN H^+), 501 (TFA loss of N_2)
Reaction 4	520 (TFA loss of N_2 gain of H_2O H^+), 536 (unknown)	328 (unknown), 424 (loss of N_2 gain of H_2O H^+), 501 (TFA loss of N_2), 542 (TFA loss of N_2 gain of MeCN H^+), 573 (dimer succinic acid H^+), 613 (unknown), 687 (unknown)
Reaction 2	520 (TFA loss of N_2 gain of H_2O H^+)	328 (unknown), 536 (unknown)

The resin chemistries swell sufficiently in DMF and DMSO to allow even distribution of the beads across a watchglass surface. This provided a larger surface area for irradiation, and minimised the concealed resin surface. This also allowed much better solvation of the small molecule in the experiment, greater than 100 mg/mL flurbiprofen concentration in one experiment.

On resin **6** there were two peaks (m/z 369 and 465) which suggested a mass of 16 added to the resin after irradiation.

The azide chemistries both showed a loss of N_2 at all irradiation times, with resin **7** showing the incorporation of H_2 and resin **8** not showing any incorporation. The latter is likely due to a ring expansion to the ketenimine, which is much less sensitive to addition.

Resin **10**, the diazirine chemistry, showed peaks which suggested the incorporation of water and acetonitrile, presumably from the mass spectrometry experiment. By proton NMR spectroscopy it was possible to observe (Figure 5.2.5.1) the disappearance of the starting material aryl peaks as the diazirine chemistry is activated, and the appearance of many different smaller aryl peaks in this region. However, there was no evidence of incorporation of any of the selected small molecules by MS.

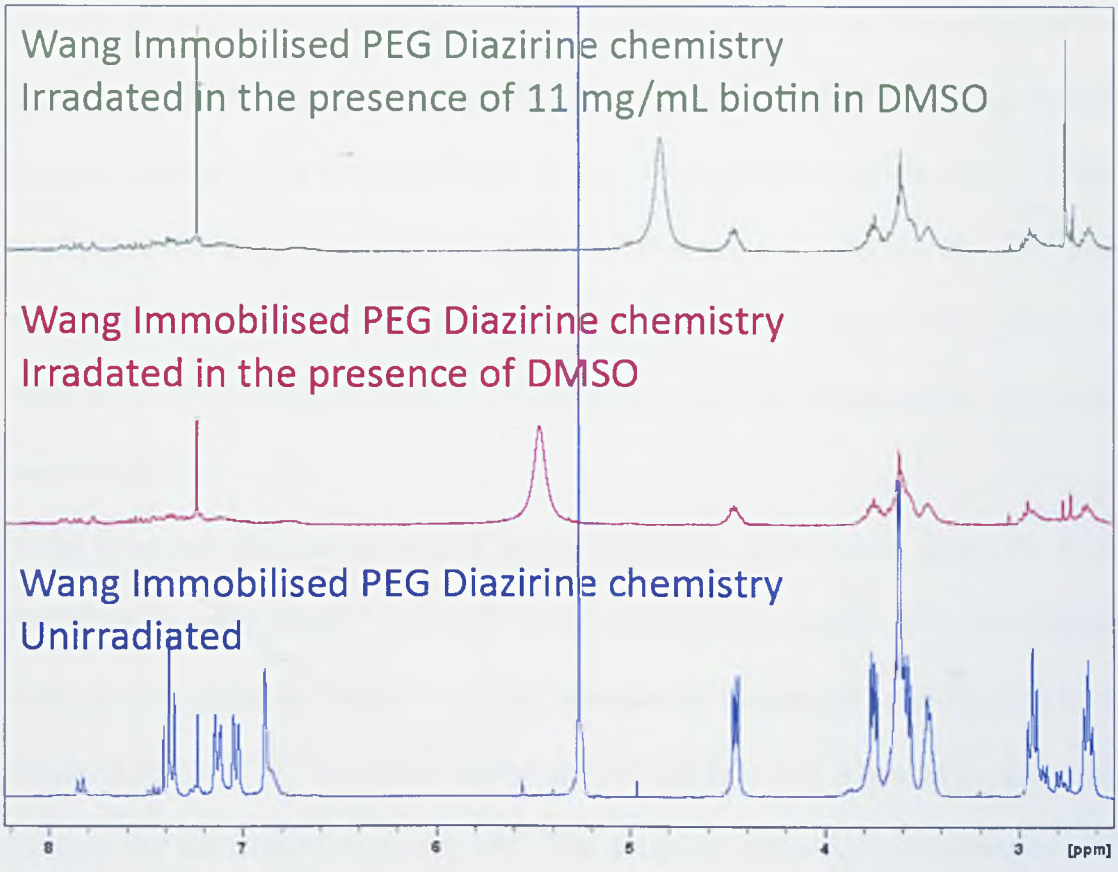


Figure 5.2.5.1: Proton NMR spectra showing the effect of irradiation at 316-400 nm in the photoreactor oven for 100 minutes on the aryl peaks of the trifluoromethyldiazirine chemistry **10** (blue) in the presence of solvent only (red) and in the presence of a solution of the small molecule biotin (green)

There could be many reasons for the lack of expected activity of these compounds. There is the possibility that the reactor oven is not functioning to its full level. Specifications say that the oven should output 25-30 Wm^{-2}, but readings of less than 0.28 Wm^{-2} were read inside the oven. It should also be noted that the Magic Tag® plates are successfully irradiated with a handheld lamp giving considerably less power.[5]

The concentration of the small molecule may still not be high enough to lead to a detectable quantity of immobilisation, even when complete saturation of the solvent is used, in the instance of the high concentration flurbiprofen experiment. Solubility is a big issue in all interactions with compounds of low solubility in the reaction solvent,[21] and some photoactive chemistries have been seen to be unreactive to certain small molecules.[21]

It is also possible that the formation of photoadducts is not in high enough yield for detection by NMR spectrometry or by MS when only a sample of the reaction product is assayed. It is known that these reactions proceed in very low yields,[22] with aryl azides showing yields of 30%, diazirines showing yields as low as 2-8% and benzophenones, while showing yields up to 70% in the literature, being prone to returning to the starting material if a suitable hydrogen for abstraction is not found immediately.[6]

In the literature, the nitrene created by the photolysis of the azide group (**7**, **8**) is known to be short lived,[46] especially in the presence of oxygen. The aryl nitrene tends to rearrange to form the seven membered ketenimine moiety which is relatively stable.[8, 47, 48] It is also known that the azide is not activated to the same levels as the diazirine in chemistry **10**.[14] The diazirines are prone to elimination and hydrolysis, leading to the loss of the immobilised small molecule[9] and the reaction with the benzophenone has been seen to be reversible under certain conditions,[20] and is deactivated by the presence of oxygen.[23]

However, small molecules have been successfully immobilised on these photoactive chemistries in the past.[20, 22, 49]

5.3 Conclusions

5.3.1 Preparation of cleavable chemistries

Five photochemical functionalities were prepared on acid-labile polystyrene resins designed for non-chemoselective immobilisation of small bioactive molecules. Characterisation was readily accomplished by acid mediated cleavage to generate the corresponding trifluroacetate esters. The thiourea chemistry, **8**, appeared to be successfully prepared from MS and IR evidence, however this

could not be verified by NMR specroscopy. This may be due to cleavage of the thiourea bond under TFA cleavage conditions, or activation of the aryl azide leading to 'back-biting' onto the base resin, resulting in the observed reduction in output yield.

5.3.2 Irradiation and activation of chemistry

Preliminary investigations of the response to UV photochemical immobilisation with test ligands, suggests the Wang-resin supported surfaces are less susceptible to the original conditions used for derivatisation of the Magic Tag® plates and comparable Tentagel® beads. More intensive irradiation conditions allowed photocleavage of N_2 from **7**, **8** and **10**, as visible in the mass spectra, and in the NMR spectrum for **10**, but none of the expected photoadduct was formed as observed by MS. Fluorescence experiments, either by direct measurement of an immobilised fluorescein molecule or by immunosorbant assay with an FITC-conjugate antibody, gave highly variable results, even between exact replicate experiments. This is likely due to experimental design.

5.4 Solution Phase Experimental

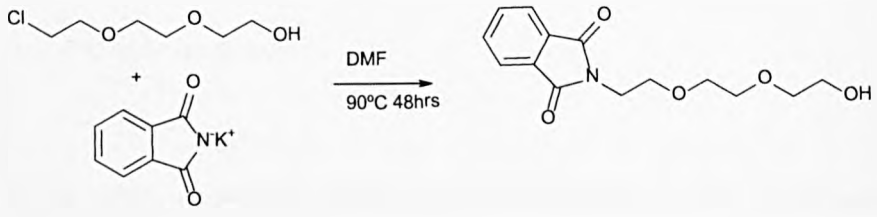
5.4.1 General Experimental

All chemicals were used as supplied unless otherwise stated. NMR spectra were obtained on Bruker Avance 300, 400 or 500 MHz spectrometers in deuterochloroform unless otherwise stated and are referenced to residual chloroform at 7.26 ppm. Shift is stated in units of ppm. Assignments were made by 2D HMBC/HMQC experiments. ^{19}F chemical shifts are quoted from the directly observed signals referenced to internal $CFCl_3$ at 0 ppm. All IR spectra were

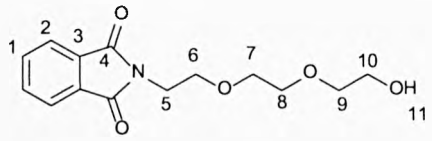
collected on a Nicolet Avatar 320 FT-IR using an attenuated total reflectance attachment, or a pellet pressed from 15-30 mg beads in 250 mg KBr. Elemental analyses were performed by Warwick Analytical Service, the analytical division of Exeter Analytical (UK) Ltd. Mass spectra were acquired on a Bruker Esquire 2000 ESI mass spectrometer.

Phthalimidoethoxyethoxyethanol,[50] trichloroacetimidate functionalised Wang resin,[51] and 3-(3-trifluoromethyl-3*H*-diazirin-3-yl)phenol[39] were prepared following previously published procedures.

5.4.2 Preparation of 2-(2-[2-phthalimidoethoxy]ethoxy)ethanol[50]



A slurry of potassium phthalimide (6.68 g, 0.036 mol) in DMF (40 mL) was stirred at room temperature as 2-(2-[2-chloroethoxy]ethoxy)ethanol (5.08 g, 0.030 mol) was added. The reaction mixture was stirred at 90°C for 48 hrs and then allowed to cool to room temperature. Water (80 mL) was added and the white suspension yielded yellow solution. This solution was extracted with DCM (80 mL), and washed with water (2x100 mL) and then dried over $MgSO_4$. The DCM and any remaining DMF were removed under reduced pressure and heating. The remaining oil was left on high vacuum over night. White crystals of KCl drop out overnight. A small amount of chloroform was added, the solution was filtered and the filtrate was concentrated by reduced pressure, giving the product 2-(2-[2-phthalimidoethoxy]ethoxy)ethanol (6.73 g, 80 %) as a yellow oil.



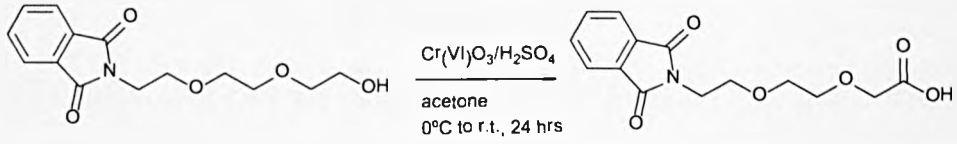
^{1}H NMR ($CDCl_3$) δ 2.15 (br s, 1H, 11), 3.54 (dd, $^{3}J_{H\text{-}H}$ 4.9 Hz, 4.3 Hz, 2H, 10), 3.58-3.61 (m, 2H, 9), 3.63-3.66 (m, 4H, 8 + 7), 3.77 (t, $^{3}J_{H\text{-}H}$ 5.7 Hz, 2H, 6), 3.91 (t, $^{3}J_{H\text{-}H}$ 5.7 Hz, 2H, 5), 7.74 (dd, $^{3}J_{H\text{-}H}$ 5.5 Hz, $^{4}J_{H\text{-}H}$ 3.1 Hz, 2H, 1), 7.84 (dd, $^{3}J_{H\text{-}H}$ 5.5 Hz, $^{4}J_{H\text{-}H}$ 3.1 Hz, 2H, 2)

^{13}C NMR ($CDCl_3$) δ 37.16 (5), 61.60 (8), 67.84 (6), 69.95 (7), 70.29 (9), 72.47 (10), 123.20 (2), 132.02 (3), 133.96 (1), 168.28 (4)

m/z (%): 280.0 (100)

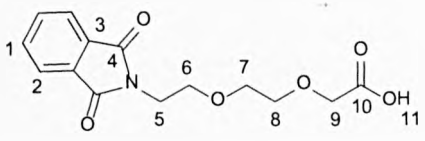
ν_{max}: 3472 (alcohol), 2867 (C-H, saturated), 1704, 1392 (conjugated ketone), 1100-1024 (aliphatic ether)

5.4.4 Preparation of 2-(2-[2-phthalimidoethoxy]ethoxy)ethanoic acid[5]



To a solution of chromic acid (chromium (VI) oxide – 2.66 g, 27 mmol in sulphuric acid – 40 mL of 1.5 M) was added dropwise at 0-5°C with stirring 2-(2-[2-phthalimidoethoxy]ethoxy)ethanol (2.04 g, 7 mmol) in acetone (50 mL). After full addition the reaction mixture was allowed to return to room temperature and was stirred for 20 hrs. The reaction mixture was filtered to remove any insoluble chromium salts and concentrated on vacuum. The aqueous layer was extracted with 3x150 mL DCM, dried with magnesium sulphate and filtered. The solvent was removed on vacuum. The acetone was removed under reduced pressure to give product 2-(2-[2-chloroethoxy]ethoxy)ethanoic acid + ester (1.75 g, 83 %; from

proton NMR, approximately 69% product) as a white solid. No further purification was carried out.

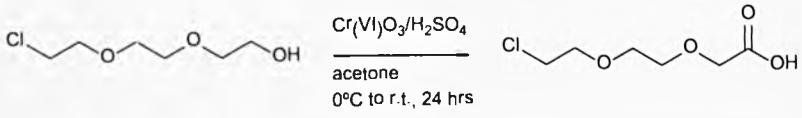


^{1}H NMR ($CDCl_3$) δ 3.72 (s, 4H, 4+5), 3.79 (t, $^{3}J_{H\text{-}H}$ 5.4 Hz, 2H, 6), 3.94 (t, $^{3}J_{H\text{-}H}$ 5.4 Hz, 2H, 7), 4.12 (s, 2H, 3), 7.74 (dd, $^{3}J_{H\text{-}H}$ 5.4 Hz, $^{4}J_{H\text{-}H}$ 3.6 Hz, 2H, 10), 7.87 (dd, $^{3}J_{H\text{-}H}$ 5.4 Hz, $^{4}J_{H\text{-}H}$ 3.6 Hz, 2H, 9)

^{13}C NMR ($CDCl_3$) δ 37.00 (7), 68.14 (6), 68.47 (3) 69.63 (5), 71.01 (4), 123.40 (10), 132.01 (9), 134.09 (11), 168.44 (8), 172.79 (2)

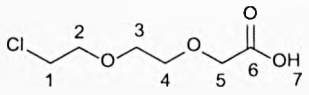
m/z (%): 294.0 (86), 316.1 (Na^+ 100)

5.4.3 Preparation of 2-(2-[2-chloroethoxy]ethoxy)ethanoic acid[52]



To a solution of chromic acid (chromium (VI) oxide - 9.28 g, 93 mmol in sulphuric acid – 100 mL of 1.5 M) was added dropwise at 0-5°C with stirring 2-(2-[2-chloroethoxy]ethoxy)ethanol (3.49 g, 21 mmol) in acetone (100 mL). After full addition the reaction mixture was allowed to return to room temperature and was stirred for 18 hrs. The reaction mixture was filtered to remove any insoluble chromium salts and concentrated on vacuum. The aqueous layer was extracted with 3x150 mL DCM, dried with magnesium sulphate and filtered. The solvent was removed on vacuum. To remove any remaining chromium products the product was redissolved in acetone and filtered over cotton wool. The acetone was removed under reduced pressure to give product 2-(2-[2-chloroethoxy]ethoxy)ethanoic acid +

methyl esters (3.4 g, 90 %; from NMRs, approximately 77 % product) as a yellow oil. No further purification was carried out.



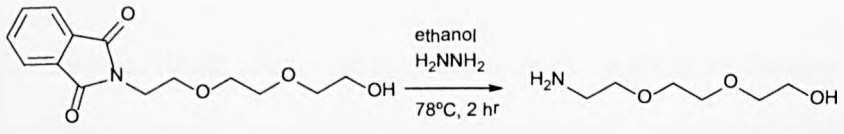
^{1}H NMR ($CDCl_3$) δ 3.58 (t, $^3J_{H\text{-}H}$ 5.7 Hz, 2H, 1), 3.65-3.68 (m, 2H, 2), 3.70-3.79 (m, 4H, 3+4), 4.15 (s, 2H, 5), 9.85 (br, 1H, 7)

^{13}C NMR ($CDCl_3$) δ 42.60 (1), 68.02 (5), 70.26/70.63 (3/4), 71.08 (2), 173.53 (6)

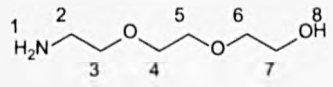
ν_{max}: 2918 (C-H, saturated), 1710 (carboxylic acid), 1109 (aliphatic ether), 663 (CH_2Cl)

m/z (%): 204.9 (Na^+, 100)

5.4.5 Preparation of 2-(2-[2-aminoethoxy]ethoxy)ethanol[43, 50, 53]



Hydrazine monohydrate (0.26 mL, 0.0053 mol, 1.2 eq) in water (0.065 mL) (to make up an 80% solution) was added to a solution of 2-(2-[2-phthalimidoethoxy]ethoxy)ethanol (1.24 g, 0.0044 mol) in ethanol (20 mL). The reaction mixture was stirred at 78°C for two hours. The white precipitate was filtered off and washed with ethanol. The filtrate was collected and the solvent removed under reduced pressure. The remaining oil was dissolved in DCM and filtered again to remove any remaining phthalohydrazide. The product was purified by column chromatography in ethyl acetate. Side products removed with 30% methanol and product eluted with 40 % methanol and 3 % ammonia, yielding 2-(2-[2-aminoethoxy]ethoxy)ethanol, yellow oil (235 mg, 35 %)



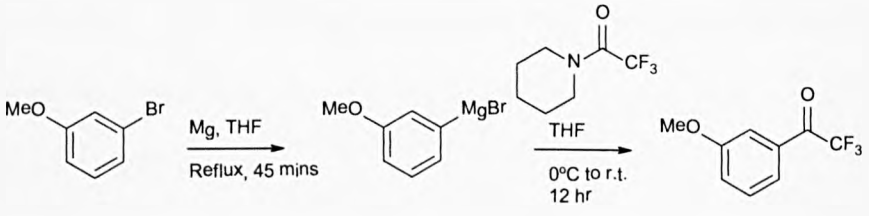
^{1}H NMR ($CDCl_3$) δ 2.90 (br t, $^3J_{H-H}$ 5.1 Hz, 2H, 2), 3.16 (br s, 4H, 1/8), 3.56 (t, $^3J_{H-H}$ 5.1 Hz, 2H, 3), 3.60 (dd, $^3J_{H-H}$ 4.7 Hz, 4.2 Hz, 2H, 6), 3.63-3.67 (m, 2H, 4/5), 3.66 (m, 2H, 4/5), 3.72 (dd, $^3J_{H-H}$ 4.7 Hz, 4.2 Hz, 2H, 7)

^{13}C NMR ($CDCl_3$) δ 41.22 (2), 61.38 (7), 70.13 (6), 70.31 (5), 72.47 (3), 72.75 (4)

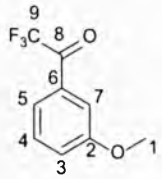
ν_{max}: 3297 (br, alcohol/amine), 2860 (C-H, saturated), 1113-1075 (aliphatic ether), 884-820 (primary amine, aliphatic)

m/z (%): 150.1 (100)

5.4.6 Preparation of 2,2,2-trifluoro-1-(3-methoxyphenyl)ethanone



In dry glassware, under nitrogen, a slurry of magnesium turnings (1.34 g, 55 mmol, 1 eq.) in dry THF (50 mL) was prepared, to which was added 3-bromoanisole (10.0 g, 53 mmol, 1 eq.) and heated slowly to reflux, at which point heat was removed and the reaction was allowed to reflux spontaneously for 45 mins. Reaction mixture was cooled to 0 °C and trifluoromethylacetylpiperidine (6.3 mL, 7.72 g, 43 mmol, 0.8 eq.) in dry THF (10 mL) was added dropwise over 30 minutes at 0 °C and then stirred at r.t. overnight. The reaction mixture was quenched with sat. NH_4Cl (10 mL), extracted into diethyl ether and washed with brine, dried over $MgSO_4$ and the solvents removed under reduced pressure. The crude product was purified by column chromatography in 2:1 Pet.ether : DCM to give title product as yellow oil (6.7 g, 77 %).



^{1}H NMR ($CDCl_3$) δ 3.66 (s, 3H, 1), 7.04 (dd, $^{3}J_{H\text{-}H}$ 8.3 Hz, $^{4}J_{H\text{-}H}$ 2.6 Hz, 1H, 5), 7.24 (t, $^{3}J_{H\text{-}H}$ 7.9 Hz, 1H, 4), 7.36 (s, 1H, 7), 7.45 (dd, $^{3}J_{H\text{-}H}$ 7.8 Hz, $^{4}J_{H\text{-}H}$ 1.3 Hz, 1H, 3)

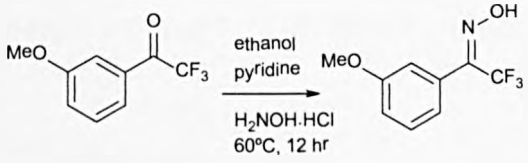
^{13}C NMR ($CDCl_3$) δ 55.43 (1), 113.97 (7), 116.55 (q, $^{1}J_{C\text{-}F}$ 292 Hz, 9), 122.19 (5), 122.66 (3), 130.08 (4), 131.04 (6), 159.98 (2), 180.30 (q, $^{2}J_{C\text{-}F}$ 38.9 Hz, 8)

^{19}F NMR ($CDCl_3$) δ - 71.68

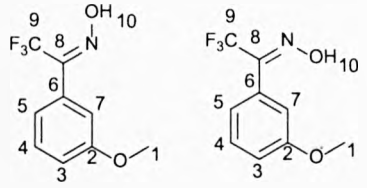
ν_{max} (cm^{-1}): 3083-2843 (C-H stretching), 1717 (ketone), 1600, 1583, 1493 (aromatic), 1251, 1138 (α-β unsaturated ether), 803 (isolated Aryl-H), 754 (3 adjacent Aryl-H), 734 (C-F)

m/z (%): 205.0 (19)

5.4.7 Preparation of 2,2,2-trifluoro-1-(3-methoxyphenyl)ethanone oxime



To 2,2,2-trifluoro-1-(3-methoxyphenyl)ethanone (6.0 g, 29 mmol) in ethanol (50 mL) and pyridine (15 mL) was added hydroxylamine hydrochloride (2.3 g, 33 mmol) and the reaction was stirred overnight at 60 °C. Solvents were removed under reduced pressure and 1M HCl (20 mL) was added and the products extracted into diethyl ether, dried over $MgSO_4$ and the solvent removed under reduced pressure. The crude product was purified by column chromatography in DCM giving the title product as yellow oil (3.7 g, 58 %).



^{1}H NMR ($CDCl_3$) δ 3.80 (s, 3H, 1), 6.99-7.40 (m, 4H, 3,4,5,7), 9.31-9.33 and 9.47-9.49 (2 br m, 8:5 ratio, 1H, 10)

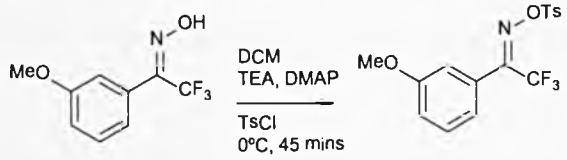
^{13}C NMR ($CDCl_3$) δ 55.36 (1), 113.86/114.22 (7), 116.14/116.29 (5), 118.23/120.47 (q, $^{1}J_{C\text{-}F}$ 282/275 Hz, 9), 120.73/120.76 (3), 129.63/129.71 (4), 147.49/147.73 (q, $^{2}J_{C\text{-}F}$ 31.3/33.3, 8), 159.40 (2)

^{19}F NMR ($CDCl_3$) δ - 62.93, -67.25

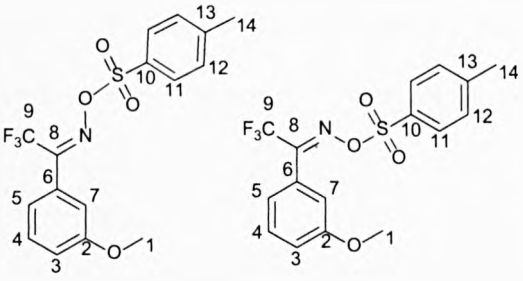
ν_{max} (cm^{-1}): 3083-2843 (C-H stretching), 1603 (oxime C=N), 1583, 1491 (aromatic), 1248, 1186 (α-β unsaturated ether), 827 (isolated Aryl-H), 766 (3 adjacent Aryl-H), 734 (C-F)

m/z (%): 220.0 (100)

Preparation of 2,2,2-trifluoro-1-(3-methoxyphenyl)ethanone oxime tosylate



2,2,2-trifluoro-1-(3-methoxyphenyl)ethanone oxime (3.38 g, 15 mmol, 1 eq.) in DCM (100 mL) was cooled to 0 °C and TEA (5.5 mL, 39 mmol, 2.5 eq.) and DMAP (100 mg, 5 mol %) were added. Tosyl chloride (3.15 g, 17 mmol, 1 eq.) was added portionwise and stirred for a further 45 mins at 0 °C. The crude product was washed with 1M HCl and water, dried over $MgSO_4$ and the solvents removed under reduced pressure, giving a cloudy yellow solution which crystallised overnight to give an off-white solid (5.24 g, 91 %). Can be further recrystallised out of hexane for very high purity.



^{1}H NMR ($CDCl_3$) δ 2.39 and 2.40 (s, 3H, 14), 3.73 and 3.74 (s, 3H, 1), 6.84-7.00 (m, 3H, 3, 5, 7), 7.24-7.34 (m, 3H, 12, 4), 7.81-7.85 (m, 2H, 11)

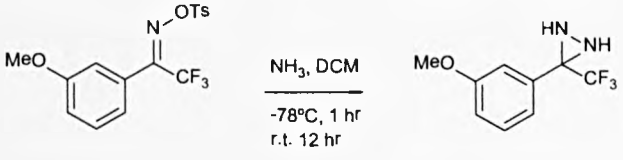
^{13}C NMR ($CDCl_3$) δ 21.66 (14), 55.32 (1), 113.88/114.27 (7), 117.10/117.39 (5), 117.22/119.52 (q, $^{1}J_{C\text{-}F}$ 285.7/277.6, 9), 120.41/121.06 (3), 129.06/129.16 (11), 129.77/129.97 (4), 129.84 (12), 131.10/131.34 (13), 146.03/146.16 (10), 153.74/153.90 (q, $^{2}J_{C\text{-}F}$ 32.0/33.8 Hz, 8), 159.46/159.49 (2)

^{19}F NMR ($CDCl_3$) δ - 62.00, -67.42

ν_{max} (cm^{-1}): 3083-2843 (C-H stretching), 1608 (oxime C=N), 1595, 1489 (aromatic), 1251, 1148 (α-β unsaturated ether), 706 (C-F)

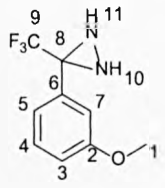
m/z (%): 375.2 (100)

5.4.8 Preparation of 3-(3-methoxyphenyl)-3-trifluoromethyl-3*H*-diaziridine



Anhydrous ammonia (roughly 20 mL, excess) was condensed into a flask at -78 °C, and 2,2,2-trifluoro-1-(3-methoxyphenyl)ethanone oxime tosylate (5.10 g, 15 mmol) in DCM (30 mL) was added. The reaction flask was maintained at -78 °C for 1 h, after which it was allowed to return to r.t. with a condenser at -78 °C maintaining the ammonia concentration. This was stirred overnight and the ammonia allowed to evaporate off. The crude product was partitioned between DCM and water, the

organic layer was washed with water, dried over $MgSO_4$ and the solvent removed under reduced pressure. This gave the title product as a yellow oil (2.81 g, 94 %).



^{1}H NMR ($CDCl_3$) δ 2.22 (d, $^3J_{H\text{-}H}$ 8.5 Hz, 1H, 10/11), 2.75 (d, $^3J_{H\text{-}H}$ 8.5 Hz, 1H, 10/11), 3.78 (s, 3H, 1), 6.93 (ddd, $^3J_{H\text{-}H}$ 8.0 Hz, $^4J_{H\text{-}H}$ 2.6 Hz, $^4J_{H\text{-}H}$ 1.0 Hz, 1H, 5) 7.11 (br s, 1H, 7), 7.16 (d, $^3J_{H\text{-}H}$ 7.5 Hz, 1H, 3), 7.29 (t, $^3J_{H\text{-}H}$ 8.0 Hz, 1H, 4)

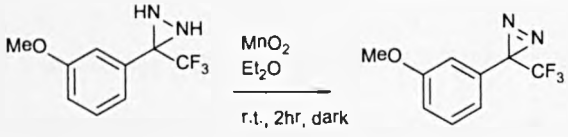
^{13}C NMR ($CDCl_3$) δ 55.29 (1), 57.96 (q, $^2J_{C\text{-}F}$ 35.9 Hz, 8), 113.59 (7), 115.77 (5), 120.25 (3), 123.48 (q, $^1J_{C\text{-}F}$ 279 Hz, 9), 129.86 (4), 132.99 (6), 159.66 (2)

^{19}F NMR ($CDCl_3$) δ - 75.94

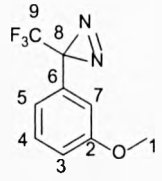
ν_{max} (cm^{-1}): 3257 (N-H stretch), 3083-2843 (C-H stretching), 1606, 1587, 1496 (aromatic), 1247, 1138 (α-β unsaturated ether), 716 (C-F)

m/z (%): 219.0 (32)

5.4.9 Preparation of 3-(3-methoxyphenyl)-3-trifluoromethyl-diazirine



MnO_2 (12.20 g, 14 mmol, 2 eq.) was added to a solution of 3-(3-methoxyphenyl)-3-trifluoromethyl-*3H*-diaziridine (1.62 g, 7.4 mmol, 1 eq.) in diethyl ether (50 mL). This was stirred vigorously for 2 h in the dark. The light sensitive product was filtered through Celite®, the solid washed with diethyl ether (100 mL). Solvent was removed under reduced pressure to yield the product as a yellow oil (1.47 g, 91%).



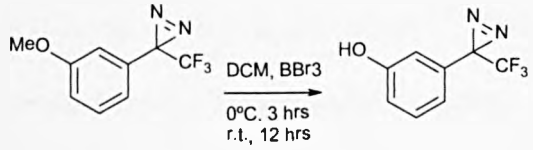
^{1}H NMR ($CDCl_3$) δ 3.73 (s, 3H, 1), 6.65 (s, 1H, 7), 6.72 (dt, $^{3}J_{H\text{-}H}$ 8.0 Hz, $^{4}J_{H\text{-}H}$ 0.8 Hz, 1H, 5), 6.88 (ddd, $^{3}J_{H\text{-}H}$ 8.0 Hz, $^{4}J_{H\text{-}H}$ 2.4 Hz, $^{4}J_{H\text{-}H}$ 0.8 Hz, 1H, 3), 7.24 (t, $^{3}J_{H\text{-}H}$ 8.0 Hz, 1H, 4)

^{13}C NMR ($CDCl_3$) δ 20.42 (q, $^{2}J_{C\text{-}F}$ 40.0 Hz, 8), 55.23 (1), 112.22 (7), 115.14 (5), 118.68 (3), 122.14 (q, $^{1}J_{C\text{-}F}$ 274 Hz, 9), 130.01 (4), 146.14 (6), 159.80 (2)

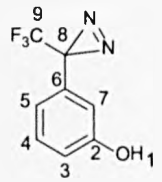
^{19}F NMR ($CDCl_3$) δ -65.55

ν_{max} (cm^{-1}): 3083-2843 (C-H stretching), 1607, 1583, 1496 (aromatic), 1262, 1147 (αβ unsaturated ether), 730 (C-F)

5.4.10 Preparation of 3-(3-trifluoromethyl-3H-diazirin-3-yl)-phenol



BBr_3 (1M in DCM, 15 mL, 15 mmol, 2 eq.) was added dropwise to a solution of 3-(3-methoxyphenyl)-3-trifluoromethyl-diazirine (1.47 g, 6.8 mmol, 1 eq.) in DCM (100 mL) at 0 °C. The reaction was maintained at 0 °C for three hours and allowed to return to r.t. with stirring overnight. The crude product was partitioned between DCM and water and the organic layer dried over $MgSO_4$. Solvents were removed under reduced pressure to give the product as a brown oil (1.02 g, 74 %).



^{1}H NMR ($CDCl_3$) δ 5.55 (br s, 1H, 1), 6.67 (s, 1H, 5), 6.73 (d, $^{3}J_{H-H}$ 8.0 Hz, 1H, 3), 6.88 (dd, $^{3}J_{H-H}$ 8.0 Hz, $^{4}J_{H-H}$ 2.6 Hz, 1H, 5), 7.24 (t, $^{3}J_{H-H}$ 8.0 Hz, 1H, 7)

^{13}C NMR ($CDCl_3$) δ 20.27 (q, $^{2}J_{C-F}$ 40.3 Hz, 8), 113.54 (7), 116.90 (5), 119.02 (3), 122.04 (q, $^{1}J_{C-F}$ 274 Hz, 9), 130.33 (4), 130.92 (6), 155.55 (2)

^{19}F NMR ($CDCl_3$) δ - 65.60

m/z (%): 202.0 (10)

5.5 Solid Phase Experimental

5.5.1 General Experimental

Wang resin (0.9 mmol/g loading, 75-100 μm) was purchased from Polymer Laboratories. Phthalimidoethoxyethoxyethanol,[50] trichloroacetimidate functionalised Wang resin,[51] and 3-(3-trifluoromethyl-3*H*-diazirin-3-yl)phenol[39] were prepared following previously published procedures. Colorimetric assays were carried out following published procedures.[44] Stated loadings were based on mass recovered product from a TFA cleavage reaction.

5.5.2 General washing procedure

(Carried out in SPE tube, drained under vacuum. Wash solvent volume used was equivalent to twice original reaction solvent)[42] solvent sequence was as follows: MeOH, Reaction solvent, Reaction solvent 1:1 H_2O, DMF, Reaction solvent 1:1 H_2O, Reaction solvent, DMF, THF, MeOH, DCM, MeOH, followed by 1 h in a vacuum oven at 25 mmHg, 50 °C for drying.

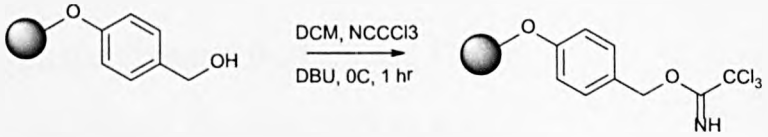
5.5.3 Cleavage procedure for Wang-immobilised compounds[42]

Wang immobilised compound (50-200 mg) was placed in a sealed SPE tube under nitrogen and dry DCM (5 mL) was added. The reaction mixture was agitated for 5 mins to allow the beads to swell. TFA (0.25 mL, 5% w.r.t. DCM) was added and agitated at room temperature for two sessions of 1h, between which the runoff is removed and fresh DCM and TFA were added. The beads were washed following the cleavage washing procedure. Runoff solution and washings were collected and solvent removed under high pressure. In specific cases, toluene (0.5 mL) added and the mixture left on high vacuum overnight.

When necessary the cleaved products were purified by a short column in ethyl acetate (for amine products the product was eluted by 57:40:3 ethyl acetate:methanol:ammonia[53]).

Products, and TFA-esters of products were recovered and analysed by NMR and IR. To hydrolyse TFA-esters, benchtop DCM was used and 0.1mL water (5% w.r.t. TFA) was added to the reaction mixture.

5.5.4 Preparation of Trichloroacetimidate-functionalised Wang Resin[51, 54]



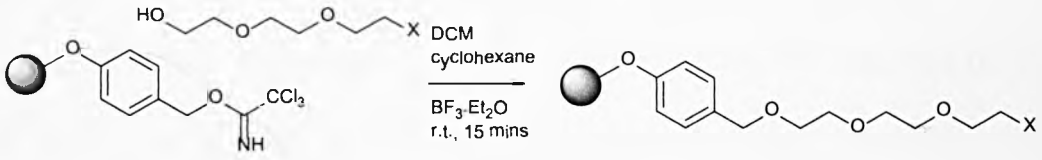
Wang resin (1.05 g, 0.9 mmol/g loading, 0.9 mmol) was placed in a sealed SPE tube with 10 mL dry DCM and the beads were allowed to swell with agitation for 5 mins. Trichloroacetonitrile (1.0 g, 7.2 mmol, 8 eq) was added and the vessel was cooled to 0 °C. DBU (0.108 mL, 0.11 g, 0.72 mmol, 0.8 eq) was added and the reaction was

agitated for 1h. The solvent was removed and the washed following the general procedure.

ν_{max}: 1663/1655 cm^{-1} (double peak), 1617/1601 cm^{-1} (double peak).

Anal. (%): C, 77.54; H, 6.35; N, 1.16

5.5.6 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethyl linked compounds[54]



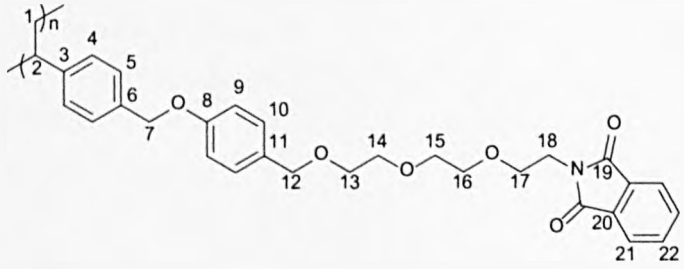
Wang-immobilised trichloroacetimidate (max. loading 0.9 mmol/g) was placed in a sealed SPE tube with dry DCM (0.1 mL/mg resin) and the beads were allowed to swell with agitation for 5 mins. End functionalised triethylene glycol (5 eq) was added in cyclohexane (0.01 mL/mg resin). Borontrifluoride diethyletherate (1 eq) was added and agitated for fifteen minutes. The solvents were removed and the beads washed following the general procedure.

5.5.6a Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylphthalmide, x= phthalimide group

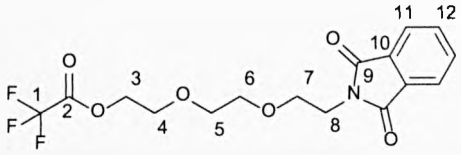
Recovered loading: 0.45 mmol/g.

ν_{max} (cm^{-1}) resin KBr: 1770 and 1720

Anal. (%): C, 86.12; H, 7.22; N, 0.59



^{13}C NMR ($CDCl_3$ Gel Phase) δ 27.04 (1), 37.40 (18), 40.55 (2), 53.55 (13), 58.00 (12), 70.24 (7), 68.07/69.22/70.70/70.77 (14-17), 114.71 (9), 123.31 (21), 125.87 (6'), 128.02 (br, 4/5), 129.46 (10), 132.24 (20), 133.99 (22), 145.35 (3), 162.67 (8), 168.36 (19)



2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylphthalimide:

^{1}H NMR ($CDCl_3$) δ 3.53-3.58 (m, 4H, 4/7), 3.65-3.68 (m, 4H, 5/6), 3.83 (t, $^{3}J_{H-H}$ 5.5 Hz, 2H, 8), 4.34 (dd, $^{3}J_{H-H}$ 5.0Hz, 4.6 Hz, 2H, 3), 7.64 (dd, $^{3}J_{H-H}$ 5.5 Hz, $^{4}J_{H-H}$ 3.1 Hz, 2H, 12), 7.78 (dd, $^{3}J_{H-H}$ 5.5 Hz, $^{4}J_{H-H}$ 3.1 Hz, 2H, 11)

^{13}C NMR ($CDCl_3$) δ 37.33 (8), 61.57 (5), 66.79 (3), 68.23 (6), 70.07 (7), 71.77 (4), 123.41 (11), 131.90 (10), 134.22 (12), 168.65 (9)

m/z (%): 280 (hydrolysed, 100), 302.0 (54), 376.0 (20)

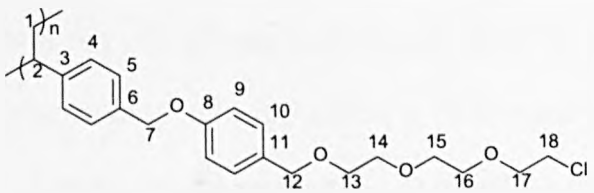
ν_{max} (cm^{-1}) film: 1780 (TFA), 1708 (5-ring unsat. imide), 1156 (ether), 720 (C-F)

5.5.6b Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylchloride, x= chloride

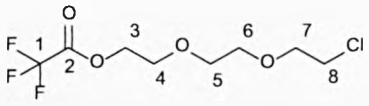
Recovered loading: 0.19 mmol/g

ν_{max} (cm^{-1}) resin KBr: 1718 , 1676 and 1638

Anal. (%): C, 84.82; H, 7.43; N, 0.045



^{13}C NMR ($CDCl_3$ Gel Phase) δ 40.56 (2), 42.85 (18), 53.56 (13), 57.93 (12), 70.13 (7), 69.26/70.81/71.49 (14-17), 114.77 (9), 125.75 (6'), 127.83 (br, 4/5), 129.46 (10), 145.20 (3), 158.72 (8)



2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethoxychloride:

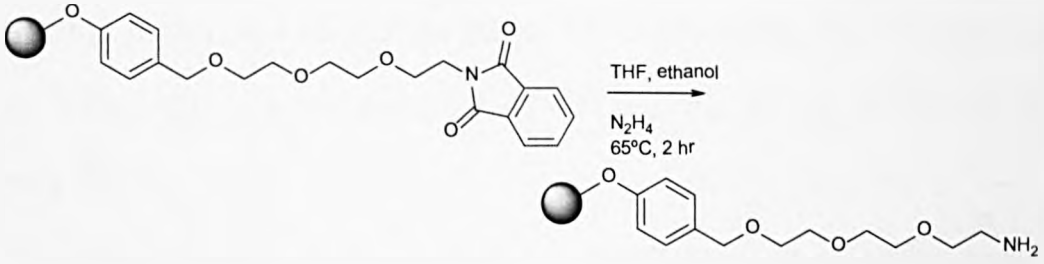
^1H NMR ($CDCl_3$) δ 3.56 (t, $^3J_{H\text{-}H}$ 5.6 Hz, 2H, 8), 3.64 (s, 4H, 5/6), 3.71 (t, $^3J_{H\text{-}H}$ 5.6 Hz, 2H, 7), 3.76 (t, $^3J_{H\text{-}H}$ 4.6 Hz, 2H, 4), 4.44 (t, $^3J_{H\text{-}H}$ 4.6 Hz, 2H, 3)

^{13}C NMR ($CDCl_3$) δ 42.71 (8), 66.85 (3), 68.30 (4), 70.64/70.74 (5/6), 71.46 (7), 125.6 (1), 218.4 (2)

m/z (%): 169.0 (hydrolysed, 100), 190.9 (44)

ν_{max} (cm^{-1}) film: 1787 (TFA), 1163 (ether), 700 (C-F), 668 (C-Cl)

5.5.7 Preparation of Wang-Immobilised 2-(2-[2-ethoxy]ethoxy)ethyl amine



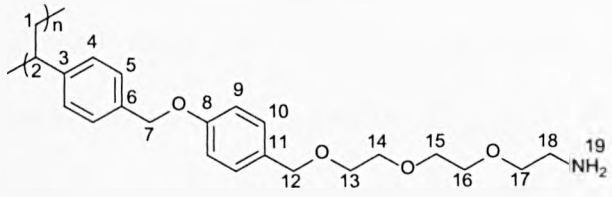
Wang immobilised 2-(2-[2-ethoxy]ethoxy)ethoxyphthalimide (406.9 mg at max 0.9 mmol/g loading, max 0.37 mmol) was placed in a sealed SPE tube with 2 mL dry THF and the beads were allowed to swell with agitation for 5 mins. Any excess THF was run off, leaving only enough to cover the beads and 2 mL ethanol was added. Hydrazine (0.01 mL, 0.0103 g, 0.32 mmol) and water (0.01 mL to make up 50% solution of hydrazine) were added and the solution was agitated for 2 h at 65 °C. The solvents were removed and the process was repeated, heating for a further 2 h. The

solvent was removed and the beads washed following the general washing procedure.

Loading by cleavage 0.4 mmol/g.

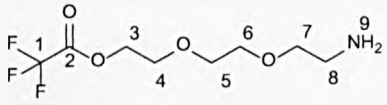
$\nu_{max}(cm^{-1})$ resin KBr: 1770 not present

Anal. (%): C, 87.43; H, 7.56; N, 0.51



^{13}C NMR ($CDCl_3$ Gel Phase) δ 25.76 (1), 40.55 (2), 41.72 (18), 53.56 (13), 57.93 (12), 70.18 (7), 69.23/70.40/70.71 (14-17), 114.79 (9), 125.82 (6'), 127.97 (br, 4/5), 129.49 (10), 145.43 (3), 158.69 (8)

2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylamine:



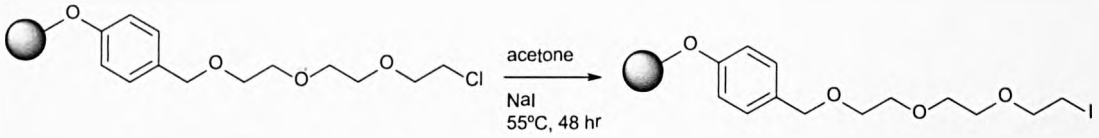
^{1}H NMR ($CDCl_3$) δ 3.22-3.27 (m, 2H, 8), 3.61-3.66 (m, 4H, 5/6), 3.73 (dd, $^{3}J_{H-H}$ 5.2 Hz, 4.8 Hz, 2H, 7), 3.76-3.80 (m, 2H, 4), 4.51 (dd, $^{3}J_{H-H}$ 6.2 Hz, 4.5 Hz, 2H, 3), 7.30 (br s, 2H, 9)

^{13}C NMR ($CDCl_3$) δ 40.23 (8), 66.05 (7), 66.45 (3), 68.38 (4), 69.87/70.49 (5/6), 130.00 (q, $^{1}J_{C-F}$ 175.6 Hz, 1), 160.70 (q, $^{2}J_{C-F}$ 39.8 Hz, 2)

m/z (%): 150.0 (hydrolysed, 100), 246 (40)

ν_{max} (cm^{-1}) film: 3452 (br, NH_2 stretch), 1784 (TFA), 1667(NH_2 bend), 1149 (ether), 700 (C-F)

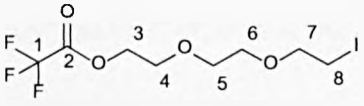
5.5.8 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethyliodide



Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylchloride (154.8 mg, max 0.9 mmol/g loading, 0.139 mmol) was placed in a sealed, foil covered SPE tube with 5 mL dry acetone (dried over anhydrous potassium carbonate, 2 h under nitrogen) and the beads were allowed to swell with agitation for 5 mins. Sodium iodide (221 mg, 1.4 mmol, 10 eq) was added and the reaction mixture was agitated at 55°C for 48 h. After allowing the reaction mixture to return to room temperature the solvent was removed and the beads washed following the general procedure.

ν_{max}(cm^{-1}) resin KBr: 1718, 1676 and 1638

2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethoxyiodide:



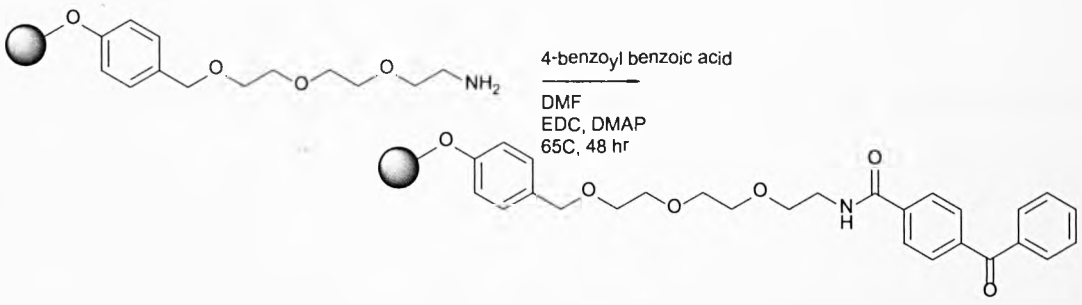
^{1}H NMR ($CDCl_3$) δ 3.24 (t, $^3J_{H\text{-}H}$ 6.7 Hz, 2H, 8), 3.61-3.67 (m, 4H, 5/6), 3.71-3.76 (m, 2H, 4), 3.78-3.82 (m, 2H, 7), 4.44 (dd, $^3J_{H\text{-}H}$ 6.2 Hz, 4.5 Hz, 2H, 3)

^{13}C NMR ($CDCl_3$) δ 2.37 (8), 66.83 (3), 68.33 (4), 70.66 (5/6), 71.99 (7)

m/z (%): 260.9 (hydrolysed, 100), 282.9 (99), 371.2 (55)

ν_{max} (cm^{-1}) film: 1787 (TFA), 1163 (ether), 720 (C-F)

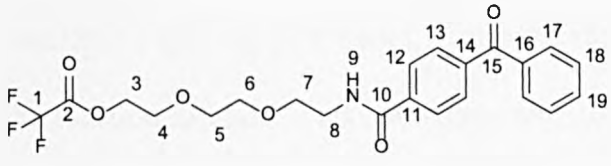
5.5.9 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylamide benzophenone[5]



Wang immobilised 2-(2-[2-ethoxy]ethoxy)ethylamine (110 mg at max 0.9 mmol/g loading, max 0.099 mmol) was placed in a sealed foil-covered SPE tube with 2 mL dry DMF and the beads were allowed to swell with agitation for 5 mins. 4-Benzoylbenzoic acid (24.9 mg, 0.11 mmol, 1 eq), EDC (42.2 mg, 0.22 mmol, 2 eq) and DMAP (5.4 mg, 0.044 mmol, 0.4 eq) were dissolved in 2 mL dry DMF and added to the reaction mixture, which was then agitated at 65 °C for 48h. After allowing the reaction mixture to return to room temperature the solvent was removed and the beads washed following the general procedure. Beads were stored in a foil-covered vial, protected from the light.

ν_{max} (cm^{-1}) resin KBr: 1720 , 1640

2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethyl-4-benzoyl benzamide:



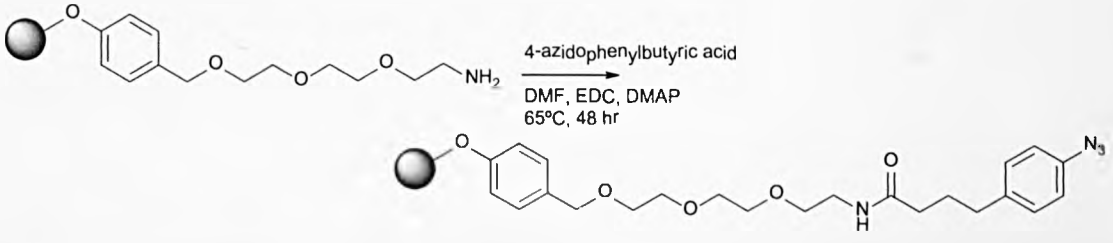
^{1}H NMR ($CDCl_3$) δ 3.63 (s, 4H, 5, 8), 3.65 (s, 4H, 5/6), 3.74 (t, $^3J_{H\text{-}H}$ 4.2 Hz, 2H, 4), 4.44 (t, $^3J_{H\text{-}H}$ 4.2 Hz, 2H, 3), 6.80 (br s, 1H, 9), 7.45 (t, $^3J_{H\text{-}H}$ 7.7 Hz, 2H, 18), 7.57 (t, $^3J_{H\text{-}H}$ 7.7 Hz, 2H, 19), 7.75 (d, $^3J_{H\text{-}H}$ 7.7 Hz, 2H, 17), 7.79 (d, $^3J_{H\text{-}H}$ 7.7 Hz, 2H, 13), 7.86 (d, $^3J_{H\text{-}H}$ 7.7 Hz, 2H, 12)

^{13}C NMR ($CDCl_3$) δ 38.68 (8), 61.59 (5), 66.73 (3), 68.29 (4), 70.67 (6/7), 114.24 (18), 128.33, 129.99, 132.56, 133.03, 137.70 (11/12/13/17/19), 131.07 (q, $^1J_{C\text{-}F}$ 168 Hz, 1), 159.26 (q, $^2J_{C\text{-}F}$ 42.6 Hz, 2), 162.60/164.40 (14/16), 197.24 (15), 197.40 (10)

m/z (%): 454.1 (100)

ν_{max} (cm^{-1}) film: 1780 (TFA), 1655 (amide/benzophenone), 1542 (amide), 1160 (ether), 700 (C-F)

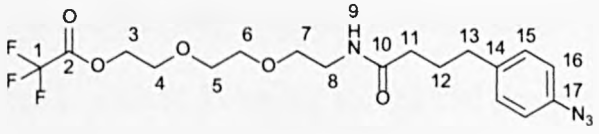
5.5.10 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylamide butyl-4-phenylazide[5]



Wang immobilised 2-(2-[2-ethoxy]ethoxy)ethylamine (105 mg at max 0.9 mmol/g loading, max 0.095 mmol) was placed in a sealed foil-covered SPE tube with 3 mL dry DMF and the beads were allowed to swell with agitation for 5 mins. 4-azidophenylbutyric acid (100 mg, 0.45 mmol, 5 eq), EDC (180 mg, 0.9 mmol, 10 eq) and DMAP (28 mg, 0.18 mmol, 4 eq) were dissolved in 2 mL dry DMF and added to the reaction mixture, which was then agitated at 65 °C for 48h. After allowing the reaction mixture to return to room temperature the solvent was removed and the beads washed following the general procedure. Beads were stored in a foil-covered vial, protected from the light.

ν_{max} (cm^{-1}) resin KBr: 2105 (azide), 1735

2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylamide butyl-4-phenylazide:



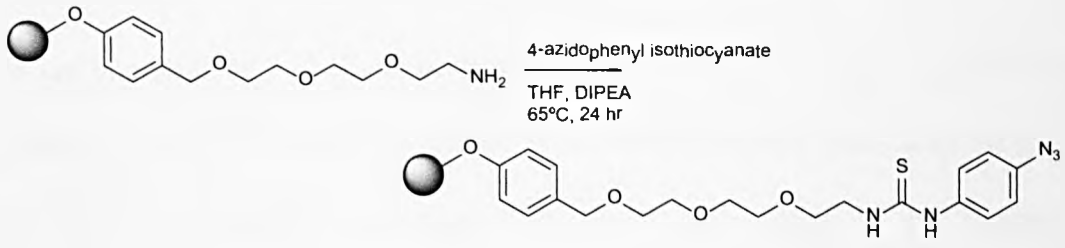
^{1}H NMR ($CDCl_3$) δ 1.92 (quint, $^{3}J_{H\text{-}H}$ 7.7 Hz, 2H, 12), 2.21 (t, $^{3}J_{H\text{-}H}$ 7.7 Hz, 2H, 11), 2.61 (t, $^{3}J_{H\text{-}H}$ 7.7 Hz, 2H, 13), 3.45 (t, $^{3}J_{H\text{-}H}$ 5.2 Hz, 2H, 8), 3.54 (t, $^{3}J_{H\text{-}H}$ 5.2 Hz, H, 7), 3.54-3.64 (m, 4H, 5/6), 3.75 (t, $^{3}J_{H\text{-}H}$ 4.4 Hz, 2H, 4), 4.48 (t, $^{3}J_{H\text{-}H}$ 4.4 Hz, 2H, 4), 6.14 (s, 1H, 9), 6.93 (d, $^{3}J_{H\text{-}H}$ 8.3 Hz, 2H, 15), 7.14 (d, $^{3}J_{H\text{-}H}$ 8.3 Hz, 2H, 16)

^{13}C NMR ($CDCl_3$) δ 27.26 (12), 34.39 (13), 35.44 (11), 39.74 (8), 66.58 (3), 68.22 (4), 69.27 (7), 69.98/70.49 (5/6), 119.05 (15), 129.76 (16), 137.71 (14), 137.93 (17), 175.83 (10)

m/z (%): 433.0 (100)

ν_{max} (cm^{-1}) film: 2110 (azide), 1787 (TFA), 1646, 1550 (amide), 1506 (Ph), 1152 (ether), 703 (C-F)

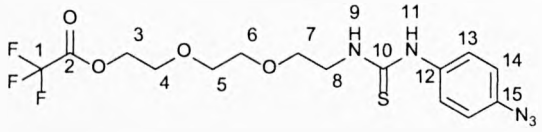
5.5.11 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylthiourea-4-phenylazide[5]



Wang immobilised 2-(2-[2-ethoxy]ethoxy)ethoxy amine (86 mg at max 0.9 mmol/g loading, max 0.077 mmol) was placed in a sealed foil-covered SPE tube with 4 mL dry THF and 1 mL ethanol and the beads were allowed to swell with agitation for 5 mins. 4-azidophenyl isothiocyanate (80 mg, 0.45 mmol, 5 eq) and diisopropylethylamine (116 mg, d = 0.74, 0.086 mL, 0.9 mmol, 10 eq) were added to the reaction mixture, which was then agitated at 65 °C for 24h. After allowing the reaction mixture to return to room temperature the solvent was removed and the beads washed following the general procedure. Beads were stored in a foil-covered vial, protected from the light.

ν_{max} (cm^{-1}) resin KBr: 2114 (azide), 1752

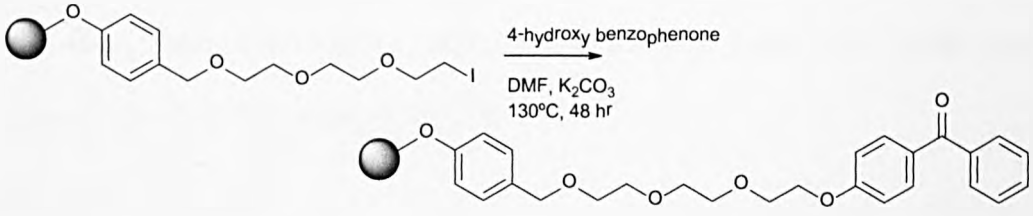
2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylthiourea-4-phenylazide:



m/z (%): 367.2 (Na^+ H_2O 100)

ν_{max} (cm^{-1}) film: 2122 (azide), 1786 (TFA), 1655 (thiourea), 1500 (Ph), 1157 (ether), 702 (C-F)

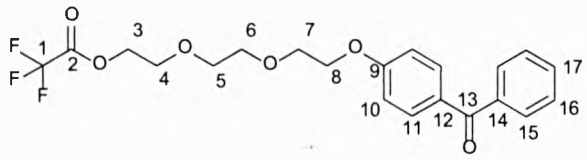
5.5.12 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethoxybenzophenone



Wang immobilised 2-(2-[2-ethoxy]ethoxy)ethoxy iodide (313 mg at max 0.9mmol/g loading, max 0.28 mmol) was placed in a round bottomed flask with 30 mL dry DMF and the beads were allowed to swell with agitation for 5 mins. 4-hydroxybenzophenone (1.8 g, 9.1 mmol, 30 eq) and potassium carbonate (1.2 g, 9.1 mmol, 30 eq) were added to the reaction mixture, which was then agitated at 130 °C for 48 h. After allowing the reaction mixture to return to room temperature the solvent was removed and the beads washed following the general procedure. Beads were stored in a foil-covered vial, protected from the light.

ν_{max} (cm^{-1}) resin KBr: 1640

2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylamide benzophenone:



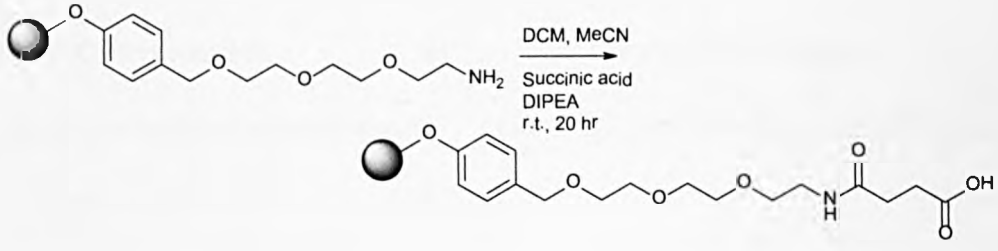
^{1}H NMR ($CDCl_3$) δ 3.55-3.70 (m, 4H, 5/6), 3.78 (t, $^{3}J_{H\text{-}H}$ 4.1 Hz, 2H, 7), 3.87 (t, $^{3}J_{H\text{-}H}$ 4.4 Hz, 2H, 4), 4.18 (t, $^{3}J_{H\text{-}H}$ 4.1 Hz, 2H, 8), 4.47 (t, $^{3}J_{H\text{-}H}$ 4.4 Hz, 2H, 3), 6.92 (s, 2H, 10), 7.45 (t, $^{3}J_{H\text{-}H}$ 7.5 Hz, 2H, 16), 7.56 (t, $^{3}J_{H\text{-}H}$ 7.2 Hz, 1H, 17), 7.72 (d, $^{3}J_{H\text{-}H}$ 7.5 Hz, 2H, 15), 7.80 (d, $^{3}J_{H\text{-}H}$ 8.6 Hz, 2H, 11)

^{13}C NMR ($CDCl_3$) δ 67.57 (7), 69.56 (8), 70.37 (4), 70.90 (5/6), 72.47 (3), 114.16 (17), 124.90/125.17/125.52/128.22/131.96/132.57/138.24 (9/10/11/12/14/15/16), 162.36 (13)

m/z (%): 331.2 (hydrolysed, 100)

ν_{max} (cm^{-1}) film: 1786 (TFA), 1652 (benzophenone), 1600, 1577, 1500 (Ph), 1257 (phenyl ether), 1170 (ether), 700 (C-F)

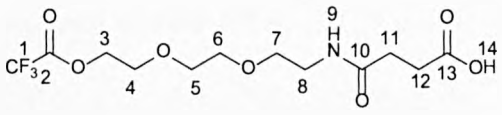
5.5.13 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylamidopropanoic acid



Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylamine (200 mg, max. loading 0.9 mmol/g, 0.18 mmol) was placed in a sealed SPE tube with DCM (5 mL) and the beads were allowed to swell with agitation for 5 mins. Acetonitrile (10 mL), succinic acid (106 mg, 1.06 mmol, 6 eq) and DIPEA (204 mg, 1.6 mmol, 8 eq) were added and the tube was agitated for 20 h at r.t. The solvents were removed and the beads were washed following the general procedure.

Recovered loading: 0.39 mmol/g.

ν_{max} (cm^{-1}) resin KBr: 1720 , 1671



2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylsuccinamic acid:

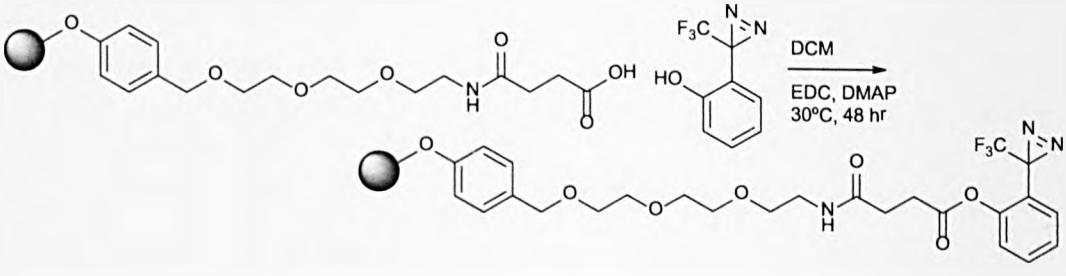
^{1}H NMR ($CDCl_3$) δ 2.54 (t, $^3J_{H-H}$ 6.4 Hz, 2H, 11), 2.66 (t, $^3J_{H-H}$ 6.4 Hz, 2H, 12), 3.39-3.42 (m, 2H, 8), 3.52 (t, $^3J_{H-H}$ 5.0 Hz, 2H, 7), 3.58-3.60/3.62-3.65 (m, 4H, 5/6), 3.77 (dd, $^3J_{H-H}$ 6.2 Hz, 4.5 Hz, 2H, 4), 4.46 (dd, $^3J_{H-H}$ 6.2 Hz, 4.5 Hz, 2H, 3), 6.93 (br s, 1H, 9), 7.14 (br s, 1H, 14)

^{13}C NMR ($CDCl_3$) δ 29.46 (12), 30.46 (11), 39.44 (8), 66.53 (3), 68.23 (4), 69.23 (7), 69.84/70.50 (5/6), 114.78 (q, $^1J_{C-F}$ 282 Hz, 1), 159.55 (q, $^2J_{C-F}$ 42.1 Hz, 2), 174.30 (10), 177.77 (13)

m/z (%): 250.1 (hydrolysed, 53), 346.1 (34), 381.2 (50)

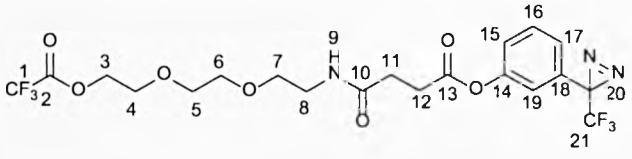
ν_{max} (cm^{-1}) film: 2721/2556 (OH), 1784 (TFA), 1719 (CO_2H), 1633, 1547 (amide), 1152 (ether), 700 (C-F)

5.5.14 Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylsuccinamic acid 3-(3-trifluoromethyl-*3H*-diazirin-3-yl)-phenyl ester



Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethylamidosuccinic acid (121 mg, max. loading 0.9 mmol/g, 0.11 mmol) was placed in a sealed SPE tube with DCM (5 mL)

and the beads were allowed to swell with agitation for 5 mins. 3-(3-trifluoromethyl-3H-diazirin-3-yl)-phenol (160 mg, 0.79 mmol, 5 eq.), EDC (210 mg, 1.4 mmol, 10 eq.) and DMAP (10 mg, 0.08 mmol, 5 mol. %) were added and the tube was agitated for 48 h at 30 °C. After allowing the reaction mixture to return to r.t. the solvent was removed and the beads washed following the general procedure. Beads were stored in a foil-covered vial, protected from the light.



2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy)ethylsuccinamic acid 3-(3-trifluoromethyl-*3H*-diazirin-3-yl)-phenyl ester:

^{1}H NMR ($CDCl_3$) δ 2.68 (t, $^{3}J_{H\text{-}H}$ 6.6 Hz, 2H, 12), 2.97 (t, $^{3}J_{H\text{-}H}$ 6.6 Hz, 2H, 11), 3.50-3.54 (m, 2H, 8), 3.57-3.62 (m, 2H, 7), 3.67-3.59 (m, 4H, 5/6), 3.73-3.77 (m, 2H, 4), 4.46-4.50 (m, 2H, 3), 6.88-6.90 (m, 1H, 19), 7.07 (d, $^{3}J_{H\text{-}H}$ 8.0 Hz, 1H, 17), 7.17 (d, $^{3}J_{H\text{-}H}$ 8.0 Hz, 1H, 15), 7.42 (t, $^{3}J_{H\text{-}H}$ 8.0 Hz, 1H, 16)

^{13}C NMR ($CDCl_3$) δ 28.80 (q, $^{2}J_{C\text{-}F}$ 42.1 Hz, 2), 29.49 (11), 30.54 (12), 39.80 (8), 66.60 (3), 68.18 (4), 68.24 (7), 69.94/70.43 (5/6), 115.29 (q, $^{1}J_{C\text{-}F}$ 282 Hz, 1), 119.82 (19), 121.92 (q, $^{1}J_{C\text{-}F}$ 275 Hz, 21), 123.11 (17), 124.07 (15), 130.06 (16), 130.73 (18), 150.69 (14), 171.33 (10/13)

^{19}F NMR ($CDCl_3$) δ -65.3, -75.1

m/z (%): 358.2 (22)

5.5.15 Colourimetric assays[44]

5.5.15a Ninhydrin Assay

Beads were washed with 2 x reaction solvent, DMF, MeOH and dried under low heat. Two stock solutions were made up: 1 g ninhydrin in 20 mL ethanol and 400 mg phenol in 1 mL ethanol. Roughly 100 mg beads were placed in a test tube. 4 drops of each stock solution were added followed by 4 drops of pyridine. This was heated at 110 °C for 2 minutes. Colouring was observed and recorded.

5.5.15b Base catalysed Fluorescein reaction

Roughly 100 mg beads were washed with THF and DMF in a test tube. Roughly 1 mL DMF was left behind. 1 drop Et_3N and 1 mg fluorescein were added to the solution and these were mixed and heated at 60 °C for 15 minutes. The beads were washed again with 2 x DMF, THF. The solution was still bright orange after heating. 1 drop HCl was added to the final THF washing. The beads returned to the expected yellow colouring. This was then washed with 2 x DMF, THF, Et_2O. Fluorescence was measured.

5.5.16 Photoimmobilisation reactions

5.5.16a Attempted photo-immobilisation of fluorescein at 254 nm for 10 minutes, analysed by fluorescence measurement

Wang resin photochemistry (**5**, **6**, **7**, **8**, 20 mg) placed in an SPE tube with fluorescein solution in tetrahydrofuran (6 mL, 0.01 mg/mL, 1 mg/mL or 2 mg/mL), agitated with irradiation from either the top or the side at 254 nm for 10 minutes.

Beads were washed following cleavage washing procedure and dried. The beads were transferred to a 96-well plate and their fluorescence read against unreacted samples.

5.5.16b Attempted photo-immobilisation of flurbiprofen, biotin or digoxin at 254 nm for 10 minutes, 30 minutes, 1 hour or 24 hours, analysed by cleavage (flurbiprofen, biotin) and antibody assay (biotin, digoxin)

Wang resin photochemistry (**5**, **6**, **7**, **8**, 80-150 mg, 5 - 10 mg/well) was placed in a 96-well filter plate (Millipore, Cat. #MSHVN4510) and small molecule solution (200 μL/well, 164 mg/mL flurbiprofen in DMSO, 1 mg/mL biotin in water, 1 mg/mL digoxin in PBS) was added. The digoxin plate was incubated in the presence of the solution for 1 hour and then was drained until dry. Plates were irradiated at 254 nm using a hand held UV lamp until the stated time had elapsed. After 24 hours, wells were entirely dry. Beads were washed into an SPE tube using methanol and then washed following the general washing procedure. After drying, ^{19}F gel phase NMR was carried out on the flurbiprofen reaction, and a sample of beads from all reactions were cleaved following the general cleavage procedure. Antibiotin-fluorescein isothiocyanate conjugate (200 μL/well, 1:100 dilutions in water) or antidigoxin-fluorescein isothiocyanate conjugate (200 μL, 1:100 dilutions in water) was added to the remaining biotin and digoxin reactions respectively and agitated for 1 hour. These beads were then washed with water, and the fluorescence read using a 96-well fluorimeter.

5.5.16c Attempted immobilisation of fluorescein, flurbiprofen or biotin at 316-400 nm for 10 minutes, 100 minutes or 17 hours, analysed by ^{19}F gel phase NMR and cleavage

Wang resin photochemstry (**5**, **6**, **7**, **8**, **10**, 35-40 mg) was placed on a small watchglass for maximum surface area and compound (200 μL, 1.6 mg/mL fluorescein in DMF, 0.016 mg/mL fluorescein in DMF, 15 mg/mL flurbiprofen in DMF, 15 mg/mL flurbiprofen in DMSO, 500 mg/mL flurbiprofen in DMSO, 11 mg/mL biotin in DMSO) was added. This watchglass was irradiated in a 16-bulb Luzchem ICH-2 photoreactor oven (16 x 8 W bulbs) for the stated time followed by 30 minutes in the dark. The beads were washed from the watchglass into a SPE tube with the same solvent and washed with DMSO, DMF, DCM. The beads were dried and ^{19}F gel phase NMR was carried out on the flurbiprofen samples. Cleavage was carried out following general cleavage procedure.

5.6 References

1. M. L. MacDonald, J. Lamerdin, S. Owens, B. H. Keon, G. K. Bilter, Z. D. Shang, Z. P. Huang, H. Yu, J. Dias, T. Minami, S. W. Michnick and J. K. Westwick, *Nat. Chem. Biol.*, 2006, **2**, 329-337.
2. N. Kanoh, S. Kumashiro, S. Simizu, Y. Kondoh, S. Hatakeyama, H. Tashiro and H. Osada, *Angew. Chem.-Int. Edit.*, 2003, **42**, 5584-5587.
3. P. J. Hergenrother, K. M. Depew and S. L. Schreiber, *J. Am. Chem. Soc.*, 2000, **122**, 7849-7850.
4. B. J. Leslie and P. J. Hergenrother, *Chem. Soc. Rev.*, 2008, **37**, 1347-1360.
5. S. J. Dilly, M. J. Bell, A. J. Clark, A. Marsh, R. M. Napier, M. J. Sergeant, A. J. Thompson and P. C. Taylor, *Chem. Commun.*, 2007, 2808-2810.
6. G. Dorman and G. D. Prestwich, *Biochemistry*, 1994, **33**, 5661-5673.
7. A. K. Schrock and G. B. Schuster, *J. Am. Chem. Soc.*, 1984, **106**, 5228-5234.
8. M. S. Platz, *Acc. Chem. Res.*, 1995, **28**, 487-492.
9. M. Platz, A. S. Admasu, S. Kwiatkowski, P. J. Crocker, N. Imai and D. S. Watt, *Bioconjug. Chem.*, 1991, **2**, 337-341.
10. J. J. Tate, J. Persinger and B. Bartholomew, *Nucleic Acids Res.*, 1998, **26**, 1421-1426.
11. S. R. Ladwa, S. J. Dilly, A. J. Clark, A. Marsh and P. C. Taylor, *ChemMedChem*, 2008, **3**, 742-744.
12. N. Kanoh, H. Takayama, K. Honda, T. Moriya, T. Teruya, S. Simizu, H. Osada and Y. Iwabuchi, *Bioconjug. Chem.*, 2010, **21**, 182-186.
13. K. Fang, M. Hashimoto, S. Jockusch, N. J. Turro and K. Nakanishi, *J. Am. Chem. Soc.*, 1998, **120**, 8543-8544.

14. T. Hiramatsu, Y. Guo and T. Hosoya, *Org. Biomol. Chem.*, 2007, **5**, 2916-2919.

15. S. S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328-1333.

16. F. Guillier, D. Orain and M. Bradley, *Chem. Rev.*, 2000, **100**, 3859-3859.

17. Y. Bing, *Acc. Chem. Res.*, 1998, **31**, 621-630.

18. S. J. Dilly, S. J. Carlisle, A. J. Clark, A. R. Shepherd, S. C. Smith, P. C. Taylor and A. Marsh, *J Polym Sci Pol Chem*, 2006, **44**, 2248-2259.

19. A. Marsh, S. J. Carlisle and S. C. Smith, *Tet. Lett.*, 2001, **42**, 493-496.

20. Y. Tanaka, M. R. Bond and J. J. Kohler, *Mol. BioSyst.*, 2008, **4**, 473-480.

21. M. Wiegand and T. K. Lindhorst, *Eur. J. Org. Chem.*, 2006, 4841-4851.

22. S. A. Fleming, *Tetrahedron*, 1995, **51**, 12479-12520.

23. C. Y. Jiao, I. D. Alves, V. Point, S. Lavielle, S. Sagan and G. Chassaing, *Bioconjug. Chem.*, 2010, **21**, 352-359.

24. Seneci and Pierfausto, *Solid Phase Synthesis and Combinatorial Technologies*, Wiley Interscience, 2000.

25. F. Z. Dorwald, *Organic Synthesis on Solid Phase*, VCH publishing, 2002.

26. T. Kodadek and K. Bachhawat-Sikder, *Mol. BioSyst.*, 2006, **2**, 25-35.

27. S. VandeVondele, J. Voros and J. A. Hubbell, *Biotechnol. Bioeng.*, 2003, **82**, 784-790.

28. E. Ostuni, L. Yan and G. M. Whitesides, *Colloid Surf. B-Biointerfaces*, 1999, **15**, 3-30.

29. M. Mrksich, C. S. Chen, Y. N. Xia, L. E. Dike, D. E. Ingber and G. M. Whitesides, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 10775-10778.

30. S. J. Dilly, M. P. Beecham, S. P. Brown, J. M. Griffin, A. J. Clark, C. D. Griffin, J. Marshall, R. M. Napier, P. C. Taylor and A. Marsh, *Langmuir*, 2006, **22**, 8144-8150.

31. T. L. Deegan, O. W. Gooding, S. Baudart and J. A. Porco, *Tet. Lett.*, 1997, **38**, 4973-4976.

32. S. Hanessian and F. Xie, *Tet. Lett.*, 1998, **39**, 733-736.

33. S. Hanessian and F. Xie, *Tet. Lett.*, 1998, **39**, 737-740.

34. A. Weissberg, A. Dahan and M. Portnoy, *J. Comb. Chem.*, 2001, **3**, 154-156.

35. K. Kokubo, K. Kitasaka and T. Oshima, *Org. Lett.*, 2006, **8**, 1597-1600.

36. H. Finkelstein, *Ber. Bunsen-Ges. Phys. Chem.*, 1910, **43**, 1528.

37. S. Botros, A. W. Lipkowski, A. E. Takemori and P. S. Portoghese, *J. Med. Chem.*, 1986, **29**, 874-876.

38. A. Furstner, G. Seidel, C. Kopiske, C. Kruger and R. Mynott, *Liebigs Ann.*, 1996, 655-662.

39. Y. Hatanaka, M. Hashimoto, H. Kurihara, H. Nakayama and Y. Kanaoka, *J. Org. Chem.*, 1994, **59**, 383-387.

40. D. Williams and I. Fleming, *Spectroscopic methods in organic chemistry*, McGraw-Hill, 2008.

41. AIST, Spectral Database for Organic Compounds SDBS, http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/direct_frame_top.cgi, Accessed 10/11/10.

42. C. Wood, Ph.D. Thesis, University Of Warwick, 2005.

43. T. Smith, Unpublished Work.

44. F. Gaggini, A. Porcheddu, G. Reginato, M. Rodriquez and M. Taddei, *J. Comb. Chem.*, 2004, **6**, 805-810.

45. G. Dorman and G. D. Prestwich, *Trends Biotechnol.*, 2000, **18**, 64-77.

46. M. R. Cline, S. M. Mandel and M. S. Platz, *Biochemistry*, 2007, **46**, 1981-1987.

47. W. T. Borden, N. P. Gritsan, C. M. Hadad, W. L. Karney, C. R. Kemnitz and M. S. Platz, *Acc. Chem. Res.*, 2000, **33**, 765-771.

48. K. A. Schnapp, R. Poe, E. Leyva, N. Soundararajan and M. S. Platz, *Bioconjug. Chem.*, 1993, **4**, 172-177.

49. P. J. A. Weber and A. G. BeckSickinger, *J. Pept. Res.*, 1997, **49**, 375-383.

50. H. Sato, E. Hayashi, N. Yamada, M. Yatagai and Y. Takahara, *Bioconjug. Chem.*, 2001, **12**, 701-710.

51. C. Quan and M. Kurth, *J. Org. Chem.*, 2004, **69**, 1470-1474.

52. I. D. Tomlinson, J. Mason, J. N. Burton, R. Blakely and S. J. Rosenthal, *Tetrahedron*, 2003, **59**, 8035-8047.

53. J. M. Zhu, Unpublished Work.

54. O. Baran, S. Sommer and H. Waldmann, *Angew. Chem., Int. Ed. Engl.*, 2004, **43**, 3195-3199.

Chapter 6 Conclusions

6.1 Phage Preparation and Biopanning Methods Conclusions

Preparation of a T7Select® phage library from nucleotide material extracted from human vascular tissue samples was successful, producing a validated phage library representing a small cohort of patients suffering ischemic heart disease. This library showed its capacity for the preparation of very small or specific cohorts, or potentially the preparation of libraries representing only one individual. This is a substantial step forward in the science of personalised medicine, allowing investigation into rare drug interactions as well as rare diseases.

Different wash conditions were assayed between rounds of biopanning, and the original wash condition, six washes with 0.5% TBS Tween® solution was found to be optimal for enrichment of selectively binding phage as estimated by phage titre assay.

The improvement of enrichment through increased number of rounds was also assayed, but after the amplification of the round three output, the phage pool was found to have disproportionated. Further assays into greater than three rounds may be of use, but assays to test for disproportionation must be used.

An assay into a more selective method of choosing late round phage outputs to sequence was carried out and showed very positive results. By picking a greater number of clones than required for sequencing in the final round and running an agarose gel on the PCR products from these clones it is possible to identify those clones which would not successfully sequence, or which would sequence to show phage vectors with no insert. It is also possible to identify any picks which contain two different clones, provided these clones are suitably different in insert length.

These picks would have given obscured sequences due to the overlap of two different sequences present in the sample, leading to a loss of information. By eliminating these phage clones from the sample sent for sequencing it is possible to maximise the amount of information gained from this stage of the biopanning process. It should be noted that picks which do not show any nucleotide band by agarose gel electrophoresis should not automatically be eliminated from the sequencing sample, as these may be dilute samples and are still possible to sequence. The competitive elution experiment, used to determine the relative binding of the different hits against the immobilised small molecule, proved to be poorly reproducible. Other methods for validating and determining strength of binding for these interactions are discussed in the future work section.

6.2 Biopanning Outputs Conclusions

6.2.1 Collection of Biopanning Outputs

Biopanning UV-immobilised photoadducts of five drug molecules using the Magic Tag® system against the prepared T7Select® human vascular tissue library returned sequences of clones which bound to these surfaces. The sequences were aligned into contiguous regions which were then assayed by querying the NCBI Reference Sequence (RefSeq) database of annotated proteins (NCBI accession numbers are listed after the protein names below). These queries returned a small collection of annotated proteins and the remainder of the sequences will be stored for later re-querying of the databases. These protein hits allow hypothesis generation regarding interactions of the small molecules under investigation.

By BLASTX algorithm analysis, the simvastatin biopan returned ribosomal protein S7 (NP_001002.1), disrupted in schizophrenia 1 (DISC1) protein

(NP_001158011.1), histone demethylase UTY (NP_872601.1), LYR motif-containing protein 4 (NP_001158312.1), signal-transducing adaptor protein 2 (NP_060190.2), myosin regulatory light chain 2 (NP_000423.2), thromboxane A2 receptor (NP_963998.2) and gap junction gamma-3 (NP_853516.1).

The rhein hydroxamic acid biopan returned ribosomal protein S7 (NP_001002.1), myosin regulatory light chain 2 (NP_000423.2), fumarylacetoacetate hydrolase domain containing protein (NP_001018114.1), signal transducing adaptor protein 2 (NP_060190.2), pyruvate dehydrogenase E1 component (NP_001166925.1) and gap junction gamma-3 (NP_853516.1).

The oxametacin biopan returned no annotated protein hits, after consideration of control wells.

The protizinic acid biopan returned histone demethylase UTY (NP_872601.1).

The rebamipide biopan returned pyruvate dehydrogenase E1 component (NP_001166925.1).

6.2.2 Simvastatin *vs.* thromboxane A_2 Receptor

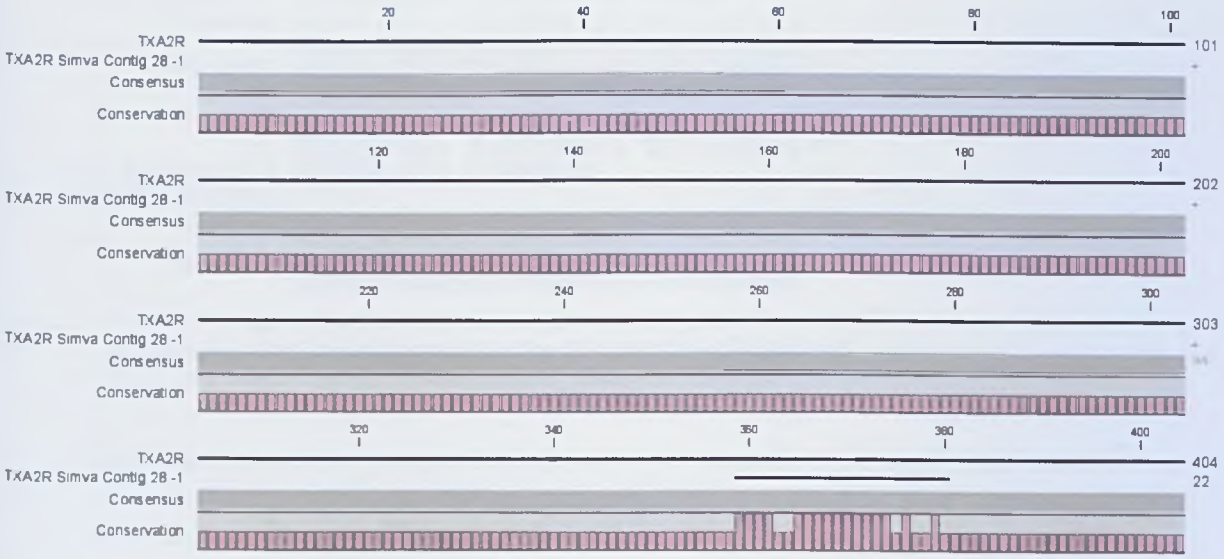
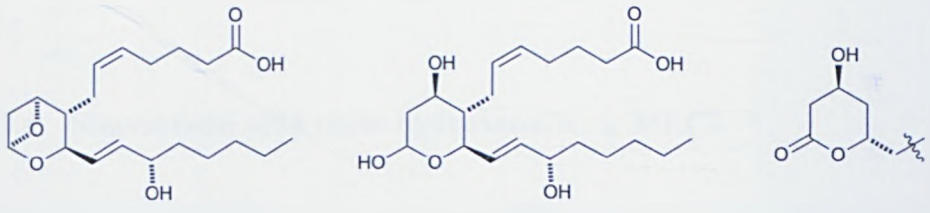


Figure 6.2.2: Alignment between thromboxane A2 receptor (NP_963998.2, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 28, transcription frame -1

While the alignment for this hit was very short (16/22 amino acids in a 404 amino acid protein), statins are known to blockade the platelet activation caused by thromboxane A_2 (TXA_2)[1, 2] and so this interaction is potentially meaningful, and definitely of interest to further study. Previous literature has suggested that this interaction is due to reduction of the biosynthesis of TXA_2, as monitored by the urinary excretion of TXB_2, its metabolite.[3] Metabolite TXB_2 has notable similarity in structure to the lactone ring of the statin.



Structures 6.2.2.1 - 3: Thromboxane A2, its metabolite thromboxane B2 and a representative statin lactone ring

6.2.3 Simvastatin and rhein hydroxamate *vs.* gap junction gamma-3

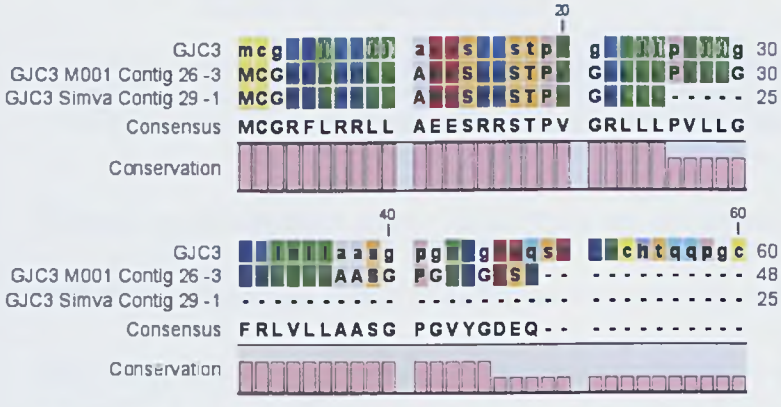


Figure 6.2.3: Alignment between gap junction gamma-3 (NP_853516.1, amino acid numbers shown above alignment) and the BLASTX hit in M001 contig 26, transcription frame -3 and simvastatin contig 29, transcription frame -1

Gap junction protein gamma-3, also known as human connexin 29, has been associated with non-syndromic hearing loss.[4] NSAIDs have also been associated with hearing loss,[5] but it is unknown if this will also be true of non-traditional

antiinflammatories such as rhein hydroxamate, where interleukin-1, rather than cyclooxygenase is the suspected target. Rhein is also known to have antioxidant activities,[6] and antioxidants have been known to have protective action against hearing loss.[7]

Certain gap junction proteins (connexins 40 and 43) have been seen to be down-regulated by statins leading to a potentially antiproliferative effect.[8] However, human connexin 29, linked to simvastatin in this screen, was not identified as down-regulated in that work.

6.2.4 Simvastatin amd rhein hydroxamate *vs.* MLC2

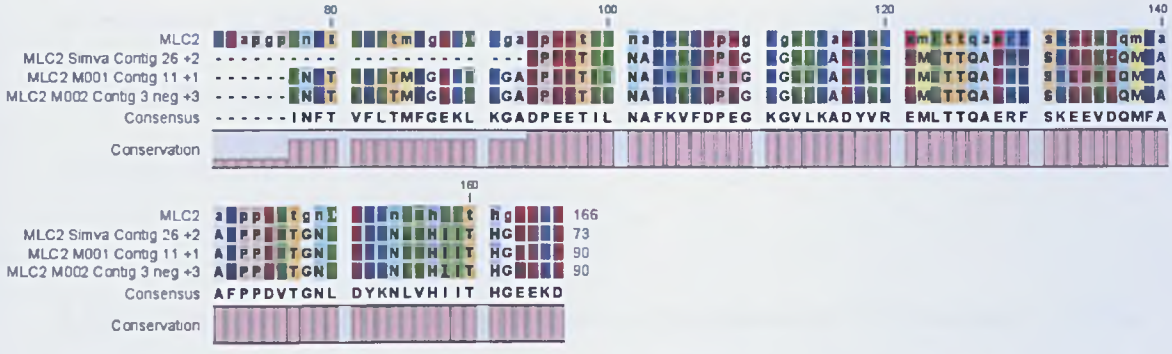


Figure 6.2.4: Alignment between myosin regulatory light chain 2 (NP_000423.2, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 26, transcription frame +2, M001 contig 11, transcription frame +1 and the control-matched alignment with M002 contig 3, transcription frame +3

Myosin regulatory light chain (MLC2) is a protein associated with cardiac myosin beta heavy chain, and contains an EF-hand region. EF-hands are defined by a helix-loop-helix structure and bind calcium ions to either act as a calcium sensor or a calcium buffer. The binding of Ca^{2+} triggers the phosphorylation of the regulatory light chain that in turn triggers contraction of the heart. Phosphorylation of myosin regulatory light chain has been observed to be attenuated by simvastatin,[9, 10] however, while MLC2 has appeared only in positives for the biopanning of simvastatin and rhein hydroxamic acid, it also appeared in the control output

sequences for oxymetacin, protizinic acid and rebamipide, and this would imply that MLC2 is a very 'sticky' peptide and so will appear to bind many different ligands, but apparently doing so without specificity for the assayed compounds.

6.2.5 Rhein hydroxamate and rebamipide *vs.* Pyruvate Dehydrogenase

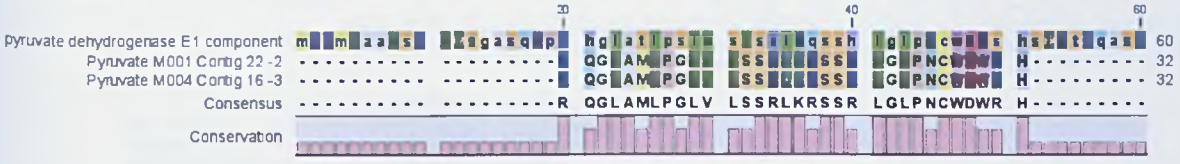


Figure 6.2.5: Alignment between pyruvate dehydrogenase E1 component (NP_001166925.1, amino acid numbers shown above alignment) and the BLASTX hit in M001 contig 22, transcription frame -2 and M004 contig 16, transcription frame -3

There are links between inflammation and impaired glucose metabolism which lead to increased lactate production through inhibition of pyruvate dehydrogenase.[11, 12] Whether this could be directly modulated by rhein hydroxymate or rebamipide has not been identified in the literature.

6.2.6 Simvastatin, rhein hydroxamate and oxametacin *vs.* Ribosomal protein S7 and Simvastatin and rhein hydroxamate *vs.* Histone Demethylase UTY

These hits have been seen in negatives and so are not considered to be of interest.

Interestingly, if the histone demethylase hit is valid and the sequence represents the DNA insert correctly, it demonstrates the tendency of the phage display library preparation method to lead to insertion or deletion of nucleotides in the sequence.[13]

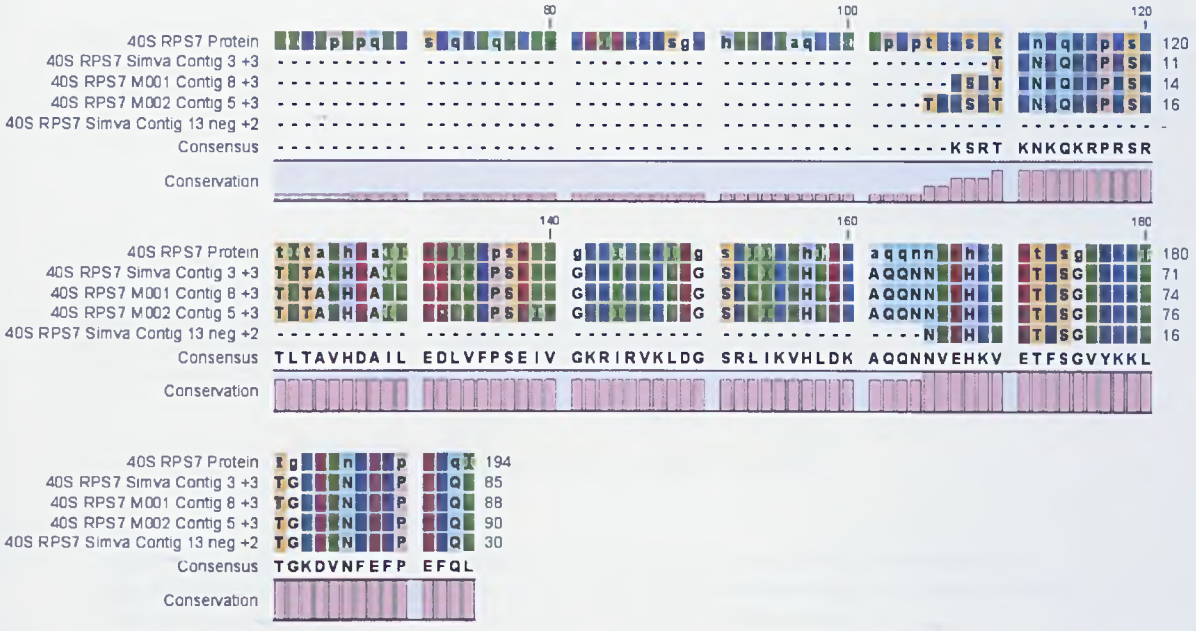


Figure 6.2.6a: Alignment between 40S ribosomal protein S7 (NP_001002.1, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 3, transcription frame +3, M001 contig 8, transcription frame +3, M002 contig 5, transcription frame +3 and the control-matched alignment with simvastatin contig 13, transcription frame +2

Figure 6.2.6b: Alignment between histone demethylase UTY (NP_872601.1, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 15, in three different transcription frames (+2, +1, +3) and control-matched alignments with M001 contig 7 in transcription frames -1 and -3.

The two contig nucleotide sequences (M001 Contig 7 Neg and Simva Contig 15) show BLASTX hits to the histone demethylase UTY (NP_872601.1) protein sequence, but across a range of translation frames. This suggests that either during preparation of the cDNA library from the extracted mRNA or during amplification of the phage clones bases have been lost or gained from the nucleotide sequence, changing the polypeptide sequence produced after translation.

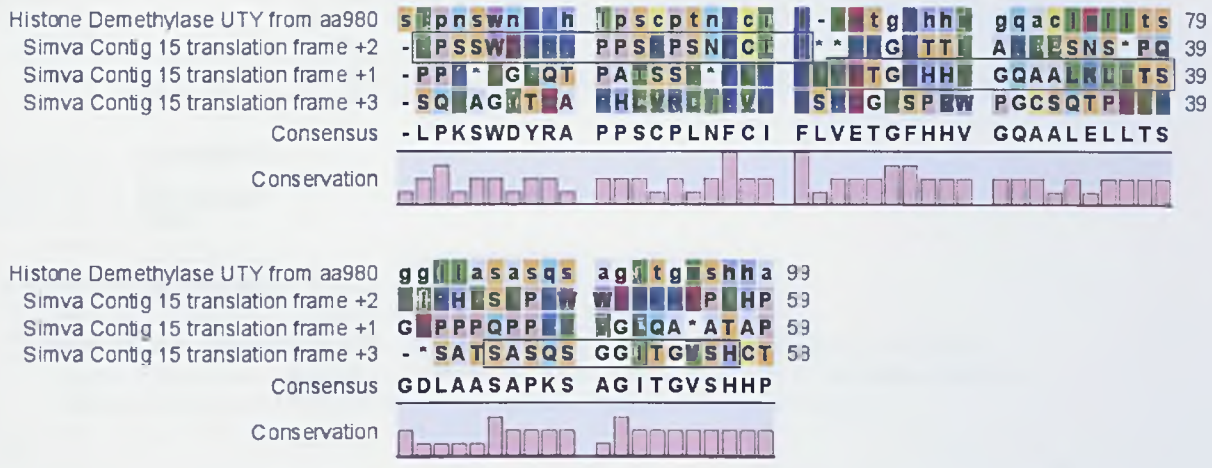


Figure 6.2.6c: Alignment between a sample of the protein sequence for histone demethylase UTY (NP_872601.1) and the polypeptide sequences produced by simvastatin contig 15 translated in three different transcription frames (+2, +1, +3). Alignments shown by black outlines.

When the polypeptide insert is produced on the phage clone, this will produce only one of these translation frames. This will result in a polypeptide sequence which represents only a small subsection of the original mRNA sequence (as shown by the outlined polypeptide sequences in each of the translation frames in 6.2.6c), surrounded by unnatural protein.

6.2.7 Simvastatin *vs.* Disrupted in Schizophrenia 1 protein, simvastatin *vs.* LYR motif containing protein, simvastatin and rhein hydroxamate *vs.* Signal Transducing Adaptor Protein and rhein hydroxamate *vs.* fumarylacetoacetate hydrolase domain

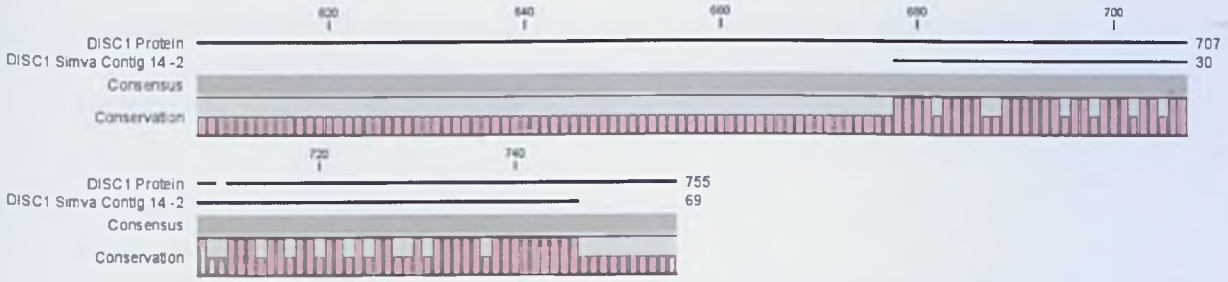


Figure 6.2.7a: Alignment between disrupted in schizophrenia 1 protein (NP_001158011.1, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 14, transcription frame -2

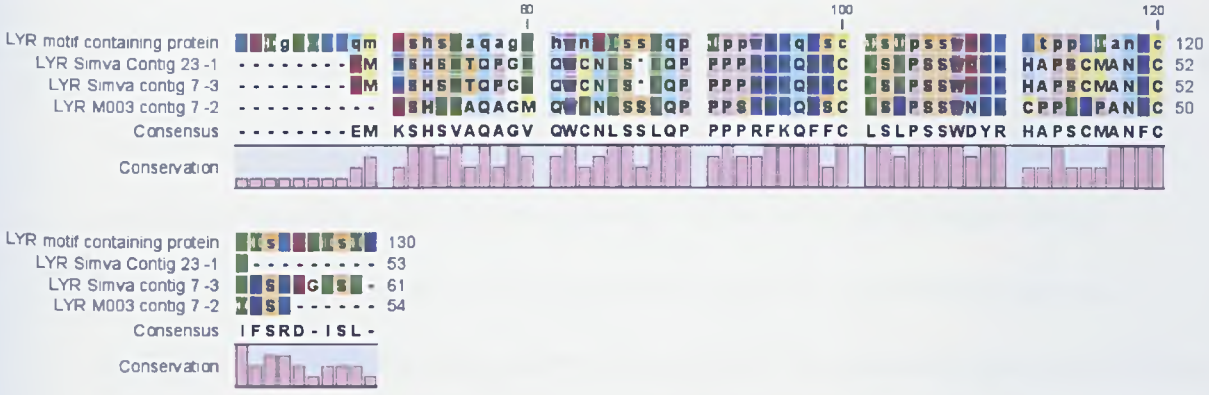


Figure 6.2.7b: Alignment between LYR motif containing protein 4 (NP_001158312.1, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 7, transcription frame -3, simvastatin contig 23, transcription frame -1 and M003 contig 7, transcription frame -2

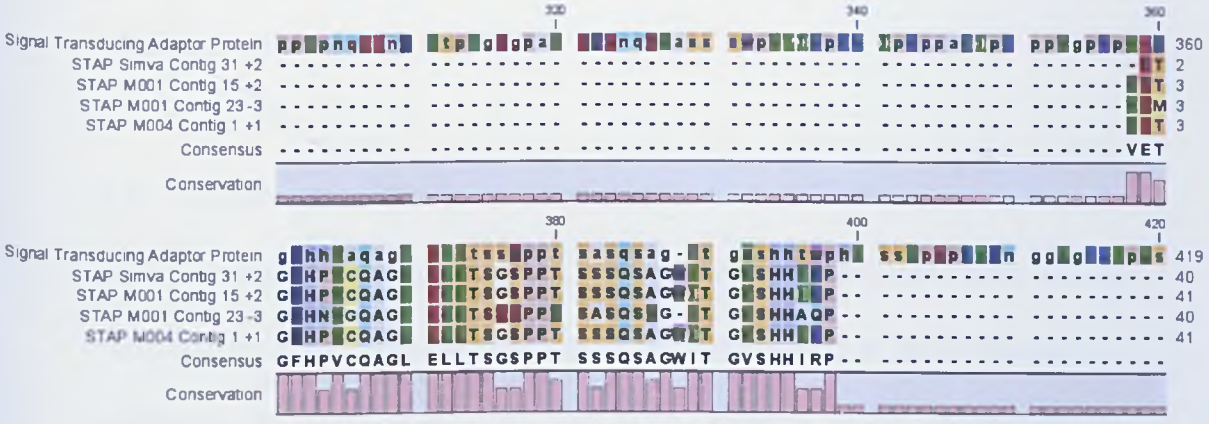


Figure 6.2.7c: Alignment between signal transducing adaptor protein 2 (NP_060190.2, amino acid numbers shown above alignment) and the BLASTX hit in simvastatin contig 31, transcription frame +2, M001 contig 15, transcription frame +2, M001 contig 23, transcription frame -3 and M004 contig 1, transcription frame +1

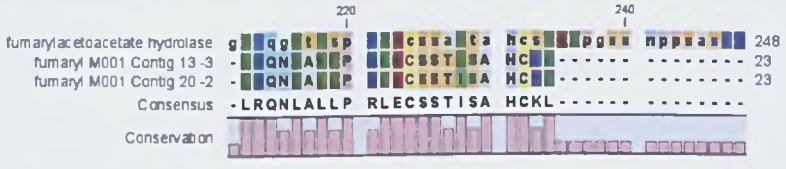


Figure 6.2.7d: Alignment between fumarylacetoacetate hydrolase domain containing protein (NP_001018114.1, amino acid numbers shown above alignment) and the BLASTX hit in M001 contig 13, transcription frame -3, M001 contig 20, transcription frame -2

Further bioinformatic studies will be needed to determine if these interactions are of meaningful interest to this project. It should be noted that the substrate of fumarylacetoacetyl hydrolase shares some structural similarity to rhein hydroxamate.

6.2.8 Negative hits

The negative wells in the biopan against rhein hydoxamate were overtaken by a single clone (M001 Contig 5 - no annotated protein hit, but a very good

MegaBLAST alignment with Homo Sapiens Chromosome 12 - 99% across 826 nucleotides). The length of the insert of this clone is unusual to be found binding control and positive wells, as longer inserts require longer to replicate during amplification and so tend to be out competed during the amplification process, leading to a diminishing appearance in later rounds. The increased appearance of this clone suggests that this clone either binds the background wells very successfully, or has some survival advantage over other clones during the amplification process which prevents its out-competition.

6.3 Cleavable Chemistry Preparation Conclusions

Five photochemical functionalities were prepared on acid-labile polystyrene resins designed for non-chemoselective immobilisation of small bioactive molecules. Preliminary investigations of the response to UV photochemical immobilisation with test ligands, suggests the Wang-resin supported surfaces are less susceptible to the original conditions used for derivatisation of the Magic Tag® plates and comparable Tentagel® beads.

6.4 Future Work

Further work will consist of validation for these interactions and analysis into their potential implications. Validation of this type of data requires several approaches, as the original data is genomic not proteomic and nothing of the strength of binding is known from the interactions observed.

Initially the clones of interest will be those which have shown BLAST hits with annotated proteins, so the first experimental follow-up must be to confirm which frame of the genetic sequence is expressed on the phage capsid protein and resulting

in interaction with the bound small molecule. This method will require over-expression of this clone in all six frames in a suitable vector with an appended marker for immunosorbant assay, followed by lysis of the cells and interaction studies. Binding of these six cell lysates to the small molecule immobilised on the appropriate chemistry, followed by enzyme-linked immunosorbant assay (ELISA) allows determination of the frame which most strongly binds to the immobilised drug molecule. A best match which is not with the expected frame may invalidate the original hit, but this information can be added to the original sequence to give a validated protein sequence for further database searching.

After confirming that the protein hit by the BLASTX algorithm was transcribed in the same frame as the strongest ELISA result, further expression of this polypeptide from the phage clone insert would allow other proteomic assays. In some cases the purification or expression of the entire annotated protein hit may be more valuable, as this will be more likely to show the accessible binding sites and natural conformations. Several of the annotated protein hits from these assays have been purified or expressed previously.[14-17]

Further assays may include;[18, 19] isothermal titration calorimetry (ITC) and equilibrium dialysis, to determine the enthalpy, stoichiometry and affinity of the interaction between the free small molecule and the polypeptide or protein; NMR spectroscopy or cocrystallisation to observe the three dimensional structure of the interaction; immunoblotting or immunoprecipitation to probe the protein with a labelled or immobilised small molecule and radioligand binding, to determine the kinetics of the binding event.

For more complete target validation the small molecule must be shown to have action against the target in the patient population, but the advantage of these

interaction studies is that they are predictive rather than retrospective, informing decisions before any expensive clinical trials are commenced.[18]

6.5 References

1. C. Borghi, G. C. Modugno and A. Pirodda, *Med. Hypotheses*, 2002, **58**, 399-402.
2. C. J. Vaughan, M. B. Murphy and B. M. Buckley, *Lancet*, 1996, **348**, 1079-1082.
3. A. Notarbartolo, G. Davi, M. Averna, C. M. Barbagallo, A. Ganci, C. Giammarresi, F. P. Laplaca and C. Patrono, *Arterioscler. Thromb. Vasc. Biol.*, 1995, **15**, 247-251.
4. H. M. Hong, J. J. Yang, C. C. Su, J. Y. Chang, T. C. Li and S. Y. Li, *Hum. Genet.*, 2010, **127**, 191-199.
5. S. G. Curhan, R. Eavey, J. Shargorodsky and G. C. Curhan, *Am. J. Med.*, 2010, **123**, 231-237.
6. G. C. Yen, P. D. Duh and D. Y. Chuang, *Food Chem.*, 2000, **70**, 437-441.
7. S. E. Heman-Ackah, S. K. Juhn, T. C. Huang and T. S. Wiedmann, *Otolaryngology-Head and Neck Surgery*, 2010, **143**, 429-434.
8. L. H. Wang, J. Z. Chen, Y. L. Sun, F. R. Zhang, J. H. Zhu, S. J. Hu and D. H. Wang, *Am. J. Hypertens.*, 2005, **18**, 1146-1153.
9. J. R. Jacobson, S. M. Dudek, K. G. Birukov, S. Q. Ye, D. N. Grigoryev, R. E. Girgis and J. G. N. Garcia, *Am. J. Respir. Cell Mol. Biol.*, 2004, **30**, 662-670.
10. K. Kani, R. Fukui, M. Hoshiga, N. Negoro, E. Kohbayashi, T. Ishihara and T. Hanafusa, *Bull. Osaka Med. Coll.*, 2007, **53**, 175-181.

11. A. Maczurek, K. Shanmugam and G. Munch, in *Maillard Reaction: Recent Advances in Food and Biomedical Sciences*, eds. E. Schleicher, V. Somoza and P. Shieberle, Blackwell Publishing, Oxford, 2008, pp. 147-151.

12. M. Jacob, I. Bjarnason and R. J. Simpson, *Clin. Sci.*, 2001, **101**, 493-498.

13. C. F. Barbas, D. R. Burton, J. K. Scott and G. J. Silverman, *Phage Display: A Laboratory Manual*, CSHL Press, 2004.

14. K. H. Ruan, V. Cervantes and J. X. Wu, *Biochemistry*, 2008, **47**, 6819-6826.

15. D. E. Timm, H. A. Mueller, P. Bhanumoorthy, J. M. Harp and G. J. Bunick, *Struct. Fold. Des.*, 1999, **7**, 1023-1033.

16. N. Nemeria, A. Volkov, A. Brown, J. Z. Yi, L. Zipper, J. R. Guest and F. Jordan, *Biochemistry*, 1998, **37**, 911-922.

17. Y. I. Tsukada and Y. Zhang, *Methods*, 2006, **40**, 318-326.

18. K. K. Kopec, D. Bozyczko-Coyne and M. Williams, *Biochem. Pharmacol.*, 2005, **69**, 1133-1139.

19. U. Rix and G. Superti-Furga, *Nat. Chem. Biol.*, 2009, **5**, 616-624.

Chapter 7 Teaching Quality Education Fund (TQEF) Project

7.1 Introduction to the Project

In the course of higher education study much information is exchanged between a tutor, supervisor or lecturer and the undergraduates or postgraduates under their supervision. Much of this information is exchanged verbally, during tutorials or group meetings or in independent research or lab work, and is not formally recorded for use by the next generation of students.

While much of this is information that can be found in text books, on the internet or through other literature searching means, the TQEF project aims to collect and present data from a variety of sources, which is relevant to specific groups of students within the university. This stored data could then be offered to the next generation of students as a 'leg up' to their studies.

In postgraduate research, some of this knowledge could potentially be harmful to IP or contain protected information, and so a data sharing website must contain locked and public-access domains, with specific information on each. The locked domain will have specific access rights to the Research group, whereas the public domain is available generally.

Due to the specific nature of postgraduate research, it's rare that two projects will require the same background knowledge, but the aim is to share the knowledge gained, and hopefully some of the information will be useful to those following me in the Marsh group.

7.2 Methods and Materials

Space on the group website was set aside for the project, with the university's Sitebuilder and Quizbuilder software allowing quick and easy website design and file uploads. This simplified web design software allowed quick updating of the pages and addition of files, chemical structures and links to presentations given at group meetings.

This software also allowed group members to be added or removed from the list of students allowed to access the group mirror of the website.

The aim was to add to the page after every meeting and learning experience. During each meeting hand written or digital notes were taken and structures noted to be later added to the website. When presentations were given by myself the slides were stored for uploading. After major learning experiences, for example the microbiology experimental carried out in preparation of the phage display library, quizzes were added to the website to cement learning and allow knowledge sharing with future students involved in similar microbiology work. Any work not relevant or specific to the IP protected work was mirrored in the publically accessible site.

7.3 Results and Discussion

The growth of the TQEF website was initially very fast, from its creation in the second year of my PhD, however when the workload increased the website was less attended, with few of the later meetings included. In the third year of my PhD the website was updated with 'how-to' pages on many pieces of laboratory equipment not generally seen in undergraduate labs as a way of introducing the uses of this equipment to new MChems and PhDs in the group, these have been the most externally used resources of the TQEF website.

For promoting the research done by a group, the addition of multimedia interviews and video presentations would have added to the site.

7.4 Conclusions and Specific Acknowledgements

While record of material discussed in meetings may be important to large groups where many students are working on similar projects, the more generic detail, such as simple introductions to unfamiliar methods and equipment are likely to be of the most use to a highly mixed multidisciplinary group. These shared materials have the potential to be of use to many students in the field, even out of the group, in the transition from undergraduate labs to research labs. An important step in the sharing of these documents has to be a moderation of their content, to ensure that the material is correct.

This project proceeded with funding from the University of Warwick under the Teaching Quality Education Fund under the guidance of Professor Donald Singer.

Chapter 8 Appendix Contents

8.5.5 Lasergene® SeqMan Analysis of M001 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.5.6 M001 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment (Electronic Only)

8.6.1 M002 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.6.2 M002 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector (Electronic Only)

8.6.3 Lasergene® SeqMan Analysis of M002 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.6.4 M002 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment (Electronic Only)

8.7.1 M003 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.7.2 M003 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector (Electronic Only)

8.7.3 Lasergene® SeqMan Analysis of M003 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.7.4 M003 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment (Electronic Only)

8.8.1 M004 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.8.2 M004 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector (Electronic Only)

8.8.3 Lasergene® SeqMan Analysis of M004 Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.8.4 M004 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment (Electronic Only)

8.9.1 M002/3/4 Negatives Biopan *vs.* Human Vascular Tissue Library Sequences (Electronic Only)

8.9.2 M002/3/4 Negatives Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector (Electronic Only)

8.10.1 Extra Chemical Experimental

NHS

National Research Ethics Service

Coventry Research Ethics Committee

Tele
Fa

Professor Donald RJ Singer

Coventry
CV2 2DX

15 October 2007

Dear Professor Singer

Full title of study: **Enhancing safety of medicines by rapid personalised testing**

REC reference number:

Thank you for your letter of 12 September 2007, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Alternative Vice-Chairman and Mr M Sakr.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised.

Ethical review of research sites

The Committee has designated this study as exempt from site-specific assessment (SSA. There is no requirement for [other] Local Research Ethics Committees to be informed or for site-specific assessment to be carried out at each site.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Application	A&B	03 August 2007
Investigator CV	Prof D Singer	03 August 2007
Protocol	Version 1	03 August 2007

Covering Letter	Prof Singer	01 August 2007
Letter from Sponsor	Mrs K Hughes	03 August 2007
Advertisement	Version 1	01 August 2007
Letter of invitation to participant	Version 1	01 August 2007
GP/Consultant Information Sheets	Version 1	01 August 2007
Participant Information Sheet: Patient Participant	Version 2	06 September 2007
Participant Information Sheet: Healthy Volunteer	Version 2	05 September 2007
Participant Consent Form: Patients	Version 1	01 August 2007
Participant Consent Form: Healthy Volunteers	Version 1	01 August 2007
Response to Request for Further Information	Clarification letter from Prof Singer	12 September 2007
Insurance Certificate	Professional Indemnity Marsh Ltd	31 July 2007
CV	Paul C Taylor	
CV	Dr A Marsh	

R&D approval

All researchers and research collaborators who will be participating in the research at NHS sites should apply for R&D approval from the relevant care organisation, if they have not yet done so. R&D approval is required, whether or not the study is exempt from SSA. You should advise researchers and local collaborators accordingly.

Guidance on applying for R&D approval is available from http://www.rdforum.nhs.uk/rdform.htm.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

Feedback on the application process

Now that you have completed the application process you are invited to give your view of the service you received from the National Research Ethics Service. If you wish to make your views known please use the feedback form available on the NRES website at:

https://www.nresform.org.uk/AppForm/Modules/Feedback/EthicalReview.aspx

We value your views and comments and will use them to inform the operational process and further improve our service.

07/H1210/126 Please quote this number on all correspondence

With the Committee's best wishes for the success of this project

Enclosures: Standard approval SL-AC2

Copy to: Mrs Kate Hughes, University of Warwick

R&D office for UHC&W NHS Trust

8.2 Human Vascular Tissue Samples and Nanodrop Results

Aortic Valve Samples - Stored

Sample ID	Date Of Extraction	Manner of Transport
MT0501	6/5/08 pm	RNA*later*
MT0502	6/5/08 pm	RNA*later*
MT0503	6/5/08 pm	RNA*later*
MT0504	6/5/08 pm	RNA*later*
MT0901	28/05/2008am	RNAlater
MT0902	28/05/2008am	RNAlater
MT0903	28/05/2008am	RNAlater
MT0904	28/05/2008am	RNAlater
MT1101	28/05/2008pm	RNAlater
MT1102	28/05/2008pm	RNAlater
MT1103	28/05/2008pm	RNAlater
MT4901	20/01/2009am	RNAlater
MT5301	27/01/2009pm	RNAlater
MT5401	28/01/2009am	RNAlater

Mammary Artery Samples - Stored

Sample ID	Date Of Extraction	Manner of Transport	Mass of Sample (mg)
MT4601	10/12/2008 midday	RNAlater	412
MT4602	10/12/2008 midday	RNAlater	499
MT4701	15/12/2008am	RNAlater	255
MT4702	15/12/2008am	RNAlater	192
MT4801	15/12/2008pm	RNAlater	237
MT4802	15/12/2008pm	RNAlater	230
MT5001	20/01/2009pm	RNAlater	397
MT5002	20/01/2009pm	RNAlater	365
MT5101	21/01/2009am	RNAlater	148
MT5102	21/01/2009am	RNAlater	365
MT5201	27/01/2009am	RNAlater	336
MT5501	03/02/2009am	RNAlater	280
MT5502	03/02/2009am	RNAlater	369
MT5601	03/02/2009pm	RNAlater	431
MT5701	04/02/2009am	RNAlater	526
MT5702	04/02/2009am	RNAlater	291
MT5901	18/02/2009am	RNAlater	Not measured
MT5902	18/02/2009am	RNAlater	Not measured

Sample ID	Date Of Extraction	Manner of Transport	Mass of Sample
MT5903	18/02/2009am	RNAlater	Not measured
MT5904	18/02/2009am	RNAlater	Not measured
MT6001	18/02/2009pm	RNAlater	Not measured
MT6002	18/02/2009pm	RNAlater	Not measured
MT6101	24/02/2009am	RNAlater	Not measured
MT6102	24/02/2009am	RNAlater	Not measured
MT6103	24/02/2009am	RNAlater	Not measured
MT6201	24/02/2009pm	RNAlater	Not measured
MT6202	24/02/2009pm	RNAlater	Not measured
MT6301	25/02/2009am	RNAlater	Not measured
MT6302	25/02/2009am	RNAlater	Not measured
MT6303	25/02/2009am	RNAlater	Not measured
MT6401	25/02/2009pm	RNAlater	Not measured
MT6501	17/03/2009am	RNAlater	Not measured
MT6502	17/03/2009am	RNAlater	Not measured
MT6503	17/03/2009am	RNAlater	Not measured
MT6504	17/03/2009am	RNAlater	Not measured
MT6601	17/03/2009pm	RNAlater	Not measured
MT6602	17/03/2009pm	RNAlater	Not measured
MT6603	17/03/2009pm	RNAlater	Not measured
MT6604	17/03/2009pm	RNAlater	Not measured
MT6701	18/03/2009pm	RNAlater	Not measured
MT6702	18/03/2009pm	RNAlater	Not measured
MT6703	18/03/2009pm	RNAlater	Not measured
MT6704	18/03/2009pm	RNAlater	Not measured
MT6801	31/03/2009am	collected by Sean James	Not measured
MT6901	07/04/2009pm	RNAlater	Not measured
MT7201	27/04/2009am	collected by Sean James	Not measured
MT7301	29/04/2009 midday	RNAlater	Not measured
MT7302	29/04/2009 midday	RNAlater	Not measured
MT7303	29/04/2009 midday	RNAlater	Not measured
MT7304	29/04/2009 midday	RNAlater	Not measured
MT7401	29/04/2009pm	RNAlater	Not measured
MT7501	03/06/2009am	RNAlater	Not measured
MT7502	03/06/2009am	RNAlater	Not measured
MT7601	03/06/2009pm	RNAlater	Not measured

Mammary Artery Processed to totalRNA

Sample ID	Date Of Extraction	Manner of Transport	Mass of Tissue (mg)	Mass totalRNA (µl)	Volume solvent (µl)	260/280	260/230
MT0101	29/4/08 am	liquid nitrogen	300	32	50	1.91	1.63
MT0301	30/4/08 pm	RNA*later*	111.3	33.7	100	1.94	1.87

Sample ID	Date Of Extraction	Manner of Transport	Mass of Tissue (mg)	Mass totalRNA (µl)	Volume solvent (µl)	260/280	260/230
MT0302	30/4/08 pm	RNA*later*	160.9	42.7	100	1.91	1.9
MT0303	30/4/08 pm	RNA*later*	104.3	34.8	100	1.93	1.82
MT0401	6/5/08 am	RNA*later*	245.2	85	100	1.91	2.16
MT0402	6/5/08 am	RNA*later*	495.1	63	100	1.98	2.02
MT0403	6/5/08 am	RNA*later*	368.2	34	100	1.96	1.95
MT0602	20/05/2008am	RNAlater	306	28.6	100	1.93	1.99
MT0603	20/05/2008am	RNAlater	224	25.9	100	1.95	1.82
MT0701	20/05/2008 pm	liquid nitrogen	998	131.7	150	1.93	1.95
MT0801	21/05/2008pm	RNAlater	61	22.4	100	1.9	1.92
MT1002	27/05/2008am	liquid nitrogen	890	129.6	150	1.87	1.85
MT1302	03/06/2008pm	RNAlater	226	13.3	100	1.78	1.59
MT1303	03/06/2008pm	RNAlater	252	22.3	100	1.78	1.59
MT1304	03/06/2008pm	RNAlater	56	22.3	100	1.88	1.8
MT1305	03/06/2008pm	RNAlater	206	24.6	100	1.87	1.66
MT1401	04/06/2008am	RNAlater	75	47.6	100	1.87	1.95
MT1502	04/06/2008pm	RNAlater	215	32.7	100	1.91	1.98
MT1503	04/06/2008pm	RNAlater	200	44.7	100	1.88	2.03
MT1504	04/06/2008pm	RNAlater	300	54.1	100	1.93	2.06
MT1505	04/06/2008pm	RNAlater	166	30.4	100	1.9	2.06
MT1601	17/06/2008pm	RNAlater	309	26	100	1.91	2.02
MT1602	17/06/2008pm	RNAlater	196	47.8	100	1.9	1.89
MT1603	17/06/2008pm	RNAlater	286	30.3	100	1.94	1.7
MT1604	17/06/2008pm	RNAlater	328	28.5	100	2.07	1.79
MT1605	17/06/2008pm	RNAlater	259	50.6	100	1.97	1.89
MT1902	18/06/2008am	RNAlater	283	3.72	100	1.9	1.78
MT1903	18/06/2008am	RNAlater	276	12.8	100	1.95	1.58
MT2002	18/06/2008pm	RNAlater	160	17.3	100	1.88	1.94
MT2003	18/06/2008pm	RNAlater	150	23.8	100	1.92	2.04
MT2101	24/06/2008am	RNAlater	275	33.5	100	2	2.15
MT2102	24/06/2008am	RNAlater	300	15	100	1.95	2.04
MT2103	24/06/2008am	RNAlater	400	29.6	100	1.91	2.11
MT2201	24/06/2008pm	RNAlater	137	21.5	100	1.88	1.85
MT2202	24/06/2008pm	RNAlater	104	22.1	100	1.92	1.75
MT2301	25/06/2008am	RNAlater	213	22.7	100	1.9	1.8
MT2501	25/06/2008pm	RNAlater	89	27.7	100	1.9	1.89
MT2502	25/06/2008pm	RNAlater	110	14.1	100	1.91	1.58
MT2601	01/07/2008pm	RNAlater	212	42.5	100	1.92	1.79
MT2602	01/07/2008pm	RNAlater	140	46.5	100	1.92	1.86
MT2603	01/07/2008pm	RNAlater	220	31.4	100	1.91	1.75
MT2604	01/07/2008pm	RNAlater	148	41.7	100	1.94	2.05
MT2605	01/07/2008pm	RNAlater	154	49.5	100	2.03	1.93
MT2701	02/07/2008pm	RNAlater	234	74.7	100	1.93	2.06
MT2702	02/07/2008pm	RNAlater	196	58.3	100	1.98	2

Sample ID	Date Of Extraction	Manner of Transport	Mass of Tissue (mg)	Mass totalRNA (µl)	Volume solvent (µl)	260/280	260/230
MT2703	02/07/2008pm	RNAlater	229	82.8	100	1.96	2.09
MT2704	02/07/2008pm	RNAlater	198	30.1	100	1.91	1.89
MT2705	02/07/2008pm	RNAlater	170	64.4	100	1.98	2.01
MT2901	29/07/2008am	liquid nitrogen	273	73.7	100	1.95	1.79
MT2902	29/07/2008am	liquid nitrogen	417	100.6	100	1.94	1.99
MT2903	29/07/2008am	liquid nitrogen	564	116.1	100	1.94	1.61
MT2904	29/07/2008am	liquid nitrogen	247	42.4	100	1.87	1.84
MT2905	29/07/2008am	liquid nitrogen	169	45.7	100	1.87	2
MT3501	10/09/2008am	RNAlater	324	74	100	2.06	1.99
MT3502	10/09/2008am	RNAlater	283	62.3	100	1.97	1.11
MT3503	10/09/2008am	RNAlater	346	51.3	100	2.05	1.92
MT3504	10/09/2008am	RNAlater	260	32.1	100	1.91	2.03
MT3601	15/09/2008am	RNAlater	211	58.6	100	1.84	2.02
MT3602	15/09/2008am	RNAlater	350	56.2	100	1.96	2.22
MT3603	15/09/2008am	RNAlater	250	24.7	100	1.91	1.88
MT3604	15/09/2008am	RNAlater	246	42.9	100	2.03	2.03
MT3701	15/09/2008pm	RNAlater	304	98.9	100	1.95	2.16
MT3702	15/09/2008pm	RNAlater	282	76.6	100	1.93	2.04
MT3703	15/09/2008pm	RNAlater	358	186.2	100	1.94	2.22
MT3704	15/09/2008pm	RNAlater	395	84.2	100	1.94	2.1
MT3801	23/09/2008am	RNAlater	345	64.9	100	1.96	2.18
MT3802	23/09/2008am	RNAlater	412	88.6	100	1.95	2.14
MT3803	23/09/2008am	RNAlater	270	86.6	100	2.01	1.98
MT3901	30/09/2008am	RNAlater	314	60.2	100	2.02	1.96
MT3902	30/09/2008am	RNAlater	306	47	100	1.87	2.02
MT4001	30/09/2008pm	RNAlater	250	98.9	100	1.9	2.17
MT4002	30/09/2008pm	RNAlater	231	86.1	100	1.81	2.06
MT4003	30/09/2008pm	RNAlater	270	75.3	100	1.8	1.93
MT4004	30/09/2008pm	RNAlater	223	78.5	100	1.79	2.08
MT4005	30/09/2008pm	RNAlater	258	66.5	100	1.93	2.07
MT4006	30/09/2008pm	RNAlater	215	50.5	100	1.9	1.95
MT4007	30/09/2008pm	RNAlater	285	98.6	100	1.91	1.86
MT4008	30/09/2008pm	RNAlater	205	42	100	1.9	1.82
MT4009	30/09/2008pm	RNAlater	250	86	100	1.87	1.99
MT4010	30/09/2008pm	RNAlater	232	83.1	100	1.89	2.17
MT4011	30/09/2008pm	RNAlater	269	87.1	100	1.85	2.06
MT4012	30/09/2008pm	RNAlater	223	50.1	100	1.91	2.08
MT4013	30/09/2008pm	RNAlater	278	66.9	100	1.91	1.82
MT4014	30/09/2008pm	RNAlater	215	80	100	1.93	2.02
MT4015	30/09/2008pm	RNAlater	287	190.2	100	1.95	2.02
MT4016	30/09/2008pm	RNAlater	205	124.7	100	1.93	2.12
MT4301	14/10/2008pm	RNAlater	131	22.3	100	1.93	1.9
MT4302	14/10/2008pm	RNAlater	157	53.4	100	1.93	2.06

Sample ID	Date Of Extraction	Manner of Transport	Mass of Tissue (mg)	Mass totalRNA (µl)	Volume solvent (µl)	260/280	260/230
MT4303	14/10/2008pm	RNAlater	187	51.1	100	1.95	1.76
MT4304	14/10/2008pm	RNAlater	165	30.4	100	1.92	2.09
MT4401	15/10/2008am	RNAlater	426	61	100	1.94	1.76
MT4402	15/10/2008am	RNAlater	484	88.1	100	1.94	1.99
MT4403	15/10/2008am	RNAlater	496	38.9	100	1.91	1.96
MT4404	15/10/2008am	RNAlater	452	71.1	100	1.96	2
MT4603	10/12/2008 midday	RNAlater	355	78	100	1.95	1.81
MT4604	10/12/2008 midday	RNAlater	404	34.6	100	2.01	1.98
MT4703	15/12/2008am	RNAlater	148	40	100	1.89	2.09
MT4704	15/12/2008am	RNAlater	261	71.7	100	1.92	1.91
MT4803	15/12/2008pm	RNAlater	224	78.1	100	1.94	1.77
MT4804	15/12/2008pm	RNAlater	212	26.4	100	1.88	1.18
MT5003	20/01/2009pm	RNAlater	457	59.8	100	1.97	1.05
MT5004	20/01/2009pm	RNAlater	369	45.9	100	1.86	1.18
MT5103	21/01/2009am	RNAlater	276	50.1	100	1.93	1.27
MT5104	21/01/2009am	RNAlater	206	48.4	100	1.89	1.16
MT5202	27/01/2009am	RNAlater	240	29.7	100	1.88	1.05
MT5503	03/02/2009am	RNAlater	352	137.6	100	1.88	1.52
MT5504	03/02/2009am	RNAlater	427	87.4	100	1.86	1.4
MT5602	03/02/2009pm	RNAlater	457	157	100	1.92	1.79
MT5703	04/02/2009am	RNAlater	327	86.5	100	1.93	2.22
MT5704	04/02/2009am	RNAlater	355	72.6	100	2.04	1.94

Saphenous Vein Processed to totalRNA

Sample ID	Date Of Extraction	Manner of Transport	Mass of Tissue (mg)	Mass totalRNA (µg)	Volume (µl)	260/280	260/230
MT0201	29/4/08 pm	RNA*later*	42.8	3.9	50	1.99	0.7
MT0601	20/05/2008am	RNAlater	50	11.6	50	1.98	1.59
MT1001	27/05/2008am	liquid nitrogen	99	23.7	100	1.94	1.8
MT1003	27/05/2008pm	liquid nitrogen	170	39.9	100	1.91	1.73
MT1201	03/06/2008am	RNAlater	130	42.2	100	1.87	1.93
MT1301	03/06/2008pm	RNAlater	127	45	100	1.68	1.78
MT1501	04/06/2008pm	RNAlater	75	16.5	100	1.9	1.23
MT1901	18/06/2008am	RNAlater	162	30.9	100	1.9	1.87
MT1904	18/06/2008am	RNAlater	173	34.4	100	1.63	1.7
MT1905	18/06/2008am	RNAlater	213	46.3	100	1.86	2
MT2001	18/06/2008pm	RNAlater	76	20.2	100	1.88	1.92

TotalRNA Samples Processed to mRNA

mRNA Sample ID	Date Processed	total RNA Sample ID	Conc. mg/ml	Volume (µl)	Mass (µg)	Collected Mass (µg)	Collected Volume (µl)	Concentration mg/ml
MT1701	10/06/08	MT0101	0.640	10	6.4			
		MT0301	0.337	10	3.37			
		MT0302	0.427	10	4.27			
		MT0303	0.348	10	3.48			
		MT0401	0.850	10	8.5			
		MT0402	0.630	10	6.3			
		MT0403	0.340	10	3.4			
		MT0602	0.286	10	2.86			
		MT0603	0.259	10	2.59			
		MT0701	0.878	10	8.78			
		MT0801	0.224	10	2.24			
		MT1002	0.864	10	8.64	60.83	120	0.507
MT1801	17/06/08	MT0101	0.640	10	6.4			
		MT0301	0.337	20	6.74			
		MT0302	0.427	20	8.54			
		MT0303	0.348	20	6.96			
		MT0401	0.850	20	17			
		MT0402	0.630	20	12.6			
		MT0403	0.340	20	6.8			
		MT0602	0.286	20	5.72			
		MT0603	0.259	20	5.18			
		MT0701	0.878	50	43.9			
		MT0801	0.224	20	4.48			
		MT1002	0.864	50	43.2	167.52	290	0.578
MT2401	25/06/08	MT1502	0.327	100	32.7			
		MT1503	0.447	100	44.7			
		MT1504	0.541	100	54.1			
		MT1505	0.304	100	30.4			
		MT1601	0.260	100	26			
		MT1602	0.478	100	47.8			
		MT1603	0.303	100	30.3			
		MT1604	0.285	100	28.5			
		MT1605	0.506	100	50.6			
		MT1902	0.037	100	3.72			
		MT1903	0.128	100	12.8			
		MT2002	0.173	100	17.3			
		MT2003	0.238	100	23.8			
		MT2101	0.335	100	33.5			
		MT2102	0.150	100	15			
		MT2103	0.296	100	29.6	480.82	1600	Concentrated to 150 µl

mRNA Sample ID	Date Processed	total RNA Sample ID	Conc. mg/ml	Volume (µl)	Mass µg	Collected Mass µg	Collected Volume µl	Concentration mg/ml
MT2801	02/06/08	MT1302	0.133	100	13.3			
		MT1303	0.223	100	22.3			
		MT1304	0.223	100	22.3			
		MT1305	0.246	100	24.6			
		MT1502	0.327	100	32.7			
		MT1503	0.447	100	44.7			
		MT1504	0.541	100	54.1			
		MT1505	0.304	100	30.4			
		MT1601	0.260	100	26			
		MT1602	0.478	100	47.8			
		MT1603	0.303	100	30.3			
		MT1604	0.285	100	28.5			
		MT1605	0.506	100	50.6			
		MT1902	0.037	100	3.72			
		MT1903	0.128	100	12.8			
		MT2002	0.173	100	17.3			
		MT2003	0.238	100	23.8			
		MT2101	0.335	100	33.5			
		MT2102	0.150	100	15			
		MT2103	0.296	100	29.6			
		MT2201	0.215	100	21.5			
		MT2202	0.221	100	22.1			
		MT2301	0.227	100	22.7			
		MT2501	0.277	100	27.7	657.32	2400	Concentrated to 210 µl

mRNA Sample ID	Date Processed	Nano-drop ng/µl	260/280	260/230	Volume Magnetic Beads	Mass mRNA µg	Volume (µl)	260/280	260/230
MT1701	10/06/08	0.4775	1.84	1.86	120 µl	0.636	20	1.65	2.98
MT1801	17/06/08	0.6186	1.92	1.92	340 µl	2.482	20	1.74	2.03
MT2401	25/06/08	0.6133	1.97	2.07	150 µl	0.018	20	0.88	1.82
MT2801	02/06/08	0.2643	1.97	2.03	210 µl	0.988	20	1.68	2.33

8.3.1 Library Validation Sequences

Clone ID	Sequence
Lib1	No Signal
Lib2	TCGGCACACAAAGCAGTGGTAGTCATAATAATTTCANANNNACAAAGGGAAAAAGGCAATAAATTAAA ATACACTTCAAGGTAAATTAGAAACACTGTTATTTATTTCTTTATTCATTTATTTATTTACTTTAGAGACAT GGTCTCACTCTGTTGCCACAATCATAGCTCACTGCAGCCTCAAGCTTGCGGCCGCACTCGAGTAACTA GTTAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGNNTANNNTAGTTACTCGAGTGCGGCCGCAAGC TTGAGGCTGCAGTGAGCTATGATTGTGGCAACAGAGTGAGACCATGTCTCTAAAGTAAATAAATAAAT GAATAAAGAAATAAATAACAGTGTTTCTAATTTACCTTGAAGTGTATTTTAATTTATTGCC
Lib3	No Signal
Lib4	CCCGGGANCTCGTCGANAAGCTTGCGGCCGCACTCGAGTAACTAGTTAACCCCTTGGGGCCTCTAAA CGGGTCTTGAGGGGTTAANTAGTTACTCGAGTGCGGCNGCAAGCTTGTCGACGAGCTCNNGGGATAT CCCTGCAGGAGAATTCGGATCCNCGAGCATCACACCTGACTGGAATACGACAGC
Lib5	No Signal
Lib6	TANTTAACCCCTTGGGGCCTCTAAACGGGTCTTGAGG
Lib7	AGGGAATCCTCCCTAACTCATTTTATGAGGCCAGCNTCNTTCTGATACCAAAGCTGGGCAGAGACNCN ACCAANAAAGAGAATTTTAGCCCAATATCCTTGNNNAACATTGATGCAAAAATCCTCAATAAAATACTG GCTAAACANATCCAGCAGCTCATCAAAAAGCTTGCGGCCGCNNT
Lib8	TTTTGGCTTCCCCTGGGCCACNTCATANAAGAATTGTCTTGGGCCACCCATAAAATACACTAACACTAA TAATAGCTGATGAGCTAAAA
Lib9	AANACACAAACTCCACATTTCTTTTTANNANNGCCCNNNCTCTATATCCTGAGNCATGTCATAAACATG GNATTGAAANAATTAAATAANGNGNANACATTGANNGATGTGTTNAAAAATAATGNACCACANCTTCTA CAANCCTNANNGAGGNTATCGGNGCACATTCCNGGGGGGNNNNGNTNAAAATGNNNNGNTCTCTGT CNGNANNACTGNCCTTTTCNNATAANGNAAAACCAAATGAANAGTNGCTGCNCCATAAANCTTGCGNC CCCACTCNANNAACTATTNAACCCCTTGGGGCCTCTAACCGGNNCTTGAGG
Lib10	GCCCATNNNATCATTTTTCTCNNNNTGGCCTCCTGGCCTGNGACGGGNNNNGCNGCTGNNNNNTCN GACNTGNNCNNNANACNTNTTCC
Lib11	No Signal
Lib12	CTCCNGGCCAGGAATGTCCAAAATTTTGGCTTCCCCTGGGCCNCNTCATANAAGAATTGTCTTGGGCC ACCCATAAAATACACTAACACTAATAATAGCTGATGAGCTAAAAA
Lib13	No Signal
Lib14	CTCCANGCCNGGAATGTCCAAAATTTTGGCTTCCCCTGGGCCACATCATAGAAGAATTGTCTTGGGCC ACCCATAAAATACACTAACACTAATAATAGCTGATGAGCTAAAAA
Lib15	No Signal
Lib16	CCCGGGANCTCGTCGANAAGCTTGCGGCCGCACTCNANNAACTAGTTAACCCCTTGGGGCCTCTAAA CGGGTCTTGAGGGGTTANCTAGTTACTCGAGTGCGGCCGCAAGCTTGTCGACGAGCTCCCGGGATAT CCCTGCAGGAGAATTCGGATCCCCGAGCATCACACCTGACTGGAATACGACAGCTCCA

8.3.2 Library Sequences Aligned against T7Select® Vector

Clone ID	Sequence
Lib4	CCCGGGANCTCGTCGANAAGCTTGCGGCCGCACTCGAGTAACTAGTTAACCCCTTGGGGCC TCTAAACGGGTCTTGAGGGGTTAANTAGTTACTCGAGTGCGGCNGCAAGCTTGTCGACGAG CTCNNGGGATATCCCTGCAGGAGAATTCGGATCCNCGAGCATCACACCTGACTGGAATACG ACAGC
Lib16	CCCGGGANCTCGTCGANAAGCTTGCGGCCGCACTCNANNAACTAGTTAACCCCTTGGGGCC TCTAAACGGGTCTTGAGGGGTTANCTAGTTACTCGAGTGCGGCCGCAAGCTTGTCGACGAG CTCCCGGGATATCCCTGCAGGAGAATTCGGATCCCCGAGCATCACACCTGACTGGAATACG ACAGCTCCA

```
BLASTN 2.2.24+
Reference: Stephen F. Altschul, Thomas L. Madden, Alejandro
A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and
David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new
generation of protein database search programs", Nucleic
Acids Res. 25:3389-3402.

RID: EXS0YNBM11R

Query= Lib 2
Length=404
                                                          Score     E
Sequences producing significant alignments:               (Bits)  Value
lcl|51059                                                  60.8    2e-14

ALIGNMENTS
>lcl|51059
Length=116

 Score = 60.8 bits (66),  Expect = 2e-14
 Identities = 33/33 (100%), Gaps = 0/33 (0%)
 Strand=Plus/Plus
Query  181  CAAGCTTGCGGCCGCACTCGAGTAACTAGTTAA  213
            |||||||||||||||||||||||||||||||||
Sbjct  84   CAAGCTTGCGGCCGCACTCGAGTAACTAGTTAA  116


 Score = 51.8 bits (56),  Expect = 1e-11
 Identities = 28/28 (100%), Gaps = 0/28 (0%)
 Strand=Plus/Minus
Query  251  TAGTTACTCGAGTGCGGCCGCAAGCTTG  278
            ||||||||||||||||||||||||||||
Sbjct  111  TAGTTACTCGAGTGCGGCCGCAAGCTTG  84

Query= Lib 4
Length=188
                                                          Score     E
Sequences producing significant alignments:               (Bits)  Value
lcl|51059                                                  176     2e-49

ALIGNMENTS
>lcl|51059
Length=116

 Score =  176 bits (194),  Expect = 2e-49
 Identities = 102/107 (96%), Gaps = 0/107 (0%)
 Strand=Plus/Minus
Query  82   TTAANTAGTTACTCGAGTGCGGCNGCAAGCTTGTCGACGAGCTCNNGGGATATCCCTGCA  141
            |||| |||||||||||||||| |||||||||||||||||||  ||||||||||||||
Sbjct  116  TTAACTAGTTACTCGAGTGCGGCCGCAAGCTTGTCGACGAGCTCCCGGGATATCCCTGCA  57

Query  142  GGAGAATTCGGATCCNCGAGCATCACACCTGACTGGAATACGACAGC  188
            |||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  56   GGAGAATTCGGATCCCCGAGCATCACACCTGACTGGAATACGACAGC  10
```

```
 Score = 82.4 bits (90),  Expect = 3e-21
 Identities = 47/49 (96%), Gaps = 0/49 (0%)
 Strand=Plus/Plus
Query  1    CCCGGGANCTCGTCGANAAGCTTGCGGCCGCACTCGAGTAACTAGTTAA  49
            ||||||| |||||||| ||||||||||||||||||||||||||||||||||
Sbjct  68   CCCGGGAGCTCGTCGACAAGCTTGCGGCCGCACTCGAGTAACTAGTTAA  116


Query= Lib 7
Length=181
                                                                 Score     E
Sequences producing significant alignments:                      (Bits)  Value
lcl|51059                                                        31.9    4e-06

ALIGNMENTS
>lcl|51059
Length=116

 Score = 31.9 bits (34),  Expect = 4e-06
 Identities = 22/25 (88%), Gaps = 0/25 (0%)
 Strand=Plus/Plus
Query  154  AGCTCATCAAAAAGCTTGCGGCCGC  178
            ||||| || | ||||||||||||||
Sbjct  74   AGCTCGTCGACAAGCTTGCGGCCGC  98


Query= Lib 16
Length=192
                                                                 Score     E
Sequences producing significant alignments:                      (Bits)  Value
lcl|51059                                                         196    2e-55

ALIGNMENTS
>lcl|51059
Length=116

 Score =  196 bits (216),  Expect = 2e-55
 Identities = 109/110 (99%), Gaps = 0/110 (0%)
 Strand=Plus/Minus
Query  82   TTANCTAGTTACTCGAGTGCGGCCGCAAGCTTGTCGACGAGCTCCCGGGATATCCCTGCA  141
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  116  TTAACTAGTTACTCGAGTGCGGCCGCAAGCTTGTCGACGAGCTCCCGGGATATCCCTGCA  57

Query  142  GGAGAATTCGGATCCCCGAGCATCACACCTGACTGGAATACGACAGCTCC  191
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  56   GGAGAATTCGGATCCCCGAGCATCACACCTGACTGGAATACGACAGCTCC  7

 Score = 71.6 bits (78),  Expect = 5e-18
 Identities = 44/49 (90%), Gaps = 0/49 (0%)
 Strand=Plus/Plus
Query  1    CCCGGGANCTCGTCGANAAGCTTGCGGCCGCACTCNANNAACTAGTTAA  49
            ||||||| |||||||| |||||||||||||||||  |  ||||||||||
Sbjct  68   CCCGGGAGCTCGTCGACAAGCTTGCGGCCGCACTCGAGTAACTAGTTAA  116
```

8.3.3 Lasergene® SeqMan Analysis of Library Sequences

>Lib Contig 1
NNACAAAGGGAAAAAGGCAATAAATTAAAATACACTTCAAGGTAAATTAGAAACACTGT
TATTTATTTCTTTATTCATTTATTTATTTACTTTAGAGACATGGTCTCACTCTGTTGCCAC
AATCATAGCTCACTGCAGCCTCAAGCTTGCGGCCGCACTCGAGTAACTAGTTAACCCC
TTGGGGCCTCTAAACGGGTCTTGAGGNNTANNNTAGTTACTCGAGTGCGGCCGCAAG
CTTGAGGCTGCAGTGAGCTATGATTGTGGCAACAGAGTGAGACCATGTCTCTAAAGTA
AATAAATAAATGAATAAAGAAATAAATAACAGTGTTTCTAATTTACCTTGAAGTGTATTTT
AATTTATTGCCTTTTTCCCTTTGTCCTCTGAAATTATTATGACTACCACTGCTTTGTGTG
NNG

>Lib Contig 2
AGGGAATCCTCCCTAACTCATTTTATGAGGCCAGCNTCNTTCTGATACCAAAGCTGGG
CAGAGACNCNACCAANAAAGAGAATTTTAGCCCAATATCCTTGNNNAACATTGATGCA
AAAATCCTCAATAAAATACTGGCTAAACANATCCAGCA

>Lib Contig 3
GTCCAAAATTTTGGCTTCCCCTGGGCCACNTCATANAAGAATTGTCTTGGGCCACCCA
TAAAATACACTAACACTAATAATAGCTGATGAGCTAAAAAAAAAAAAAAAAATCNNAAAA
AAAAATCCCCAAAAAAATC

>Lib Contig 4
CTCCANGCCNGGAATGTCCAAAATTTTGGCTTCCCCTGGGCCACATCATAGAAGAATT
GTCTTGGGCCACCCATAAAATACACTAACACTAATAATAGCTGATGAGCTAAAAAAAAA
AAAAAAAATCNCAAAAAAAAAATCCCCAAAAAAATCTCATAN

Lib2	(54>194)	Lib Contig 3
Lib7	(41>202)	Lib Contig 4
Lib8	(84>180)	Lib Contig 4
Lib12	(80>508)	Lib Contig 1
Lib14	(90>247)	Lib Contig 2

8.3.4 Library Contigs MegaBLAST Alignment

104 Pages, MegaBLAST algorithm text output of library contigs aligned against RefSeq RNA database.

Electronic copy only.

8.4.1 Simvastatin Biopan *vs.* Human Vascular Tissue Library Sequences

9 Pages, Sequences output from simvastatin biopan *vs.* human vascular tissue library.

Electronic copy only.

8.4.2 Simvastatin Biopan *vs.* Human Vascular Tissue Library Aligned against T7Select® Vector

16 Pages, BLAST alignment of two sequences (bl2seq algorithm). Alignment of sequences output from simvastatin biopan *vs.* human vascular tissue library against T7Select® vector sequence.

Electronic copy only.

8.4.3 Lasergene® SeqMan Analysis of Simvastatin Biopan *vs.* Human Vascular Tissue Library Sequences

5 Pages, Lasergene® SeqMan output of contiguous regions, after alignment of all sequences output from simvastatin biopan *vs.* human vascular tissue library. Electronic copy only.

8.4.4 Simvastatin Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment

73 Pages, BLASTX algorithm text output of simvastatin contigs (Appendix 8.4.3) aligned against RefSeq protein database.

Electronic copy only.

8.4.5 Simvastatin Biopan *vs.* Human Vascular Tissue Library:

Competitive Elution Data

25 Clones Selected for Competitive Elution Studies

Clone ID	Well Chemistry	Contig	Top BLAST Hit (BLASTX, MEGABLAST, BLASTN)	Top Accession No.
SimvaCP43	C	1	Homo sapiens ribosomal protein L21 (RPL21), mRNA	NM_000982.3
SimvaCP11	C	1	Homo sapiens ribosomal protein L21 (RPL21), mRNA	NM_000982.3
SimvaEP172	E	2	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 3B (HTR3B), mRNA	NM_006028.3
SimvaCP65	C	7	LYR motif-containing protein 4 isoform 2	NP_001158312.1
SimvaBP35	B	7	LYR motif-containing protein 4 isoform 2	NP_001158312.1
SimvaAP4	A	7	LYR motif-containing protein 4 isoform 2	NP_001158312.1
SimvaCP88	C	7	LYR motif-containing protein 4 isoform 2	NP_001158312.1
SimvaCP170	C	14	disrupted in schizophrenia 1 protein isoform c	NP_001158011.1
SimvaDP9	D	15	Homo sapiens NAD(P)H dehydrogenase, quinone 1 (NQO1), transcript variant 1, mRNA	NM_000903.2
SimvaBP12	B	16	-	
SimvaBP13	B	17	Homo sapiens arylformamidase (AFMID), transcript variant 2, mRNA	NM_001145526.1
SimvaBP14	B	18	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 3B (HTR3B), mRNA	NM_006028.3
SimvaDP20	D	19	Homo sapiens actin, alpha 1, skeletal muscle	NM_001100.3
SimvaBP24	B	21	Homo sapiens metallophosphoesterase domain containing 2	NM_001145399.1
SimvaBP36	B	22	Homo sapiens cathepsin C (CTSC), transcript variant 3, mRNA	NM_001114173.1
SimvaAP39	A	23	LYR motif-containing protein 4 isoform 2	NP_001158312.1
SimvaBP57	B	25	-	
SimvaBP58	B	26	myosin regulatory light chain 2, ventricular/cardiac muscle isoform	NP_000423.2
SimvaDP62	D	27	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 3B (HTR3B), mRNA	NM_006028.3
SimvaBP68	B	28	thromboxane A2 receptor isoform beta	NP_963998.2
SimvaCP77	C	31	signal-transducing adaptor protein 2 isoform 1	NP_060190.2
SimvaCP87	C	-	Homo sapiens nucleolar protein 9 (NOL9), mRNA	NM_024654.4
SimvaEP128	E	-	Homo sapiens chromatin modifying protein 1B (CHMP1B), mRNA	NM_020412.4
SimvaEP138	E	-	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 3B (HTR3B), mRNA	NM_006028.3
SimvaEP151	E	-	-	

First Experiment (Water)

Clone ID	Pos (plaque count)	Neg (plaque count)	ratio
SimvaBP12	1616	2976	0.54
SimvaBP13	2176	3504	0.62
SimvaBP14	1304	2296	0.57
SimvaBP24	10000	10000	1.00
SimvaBP35	10000	10000	1.00
SimvaBP36	1808	2536	0.71
SimvaBP57	10000	10000	1.00
SimvaBP58	10000	10000	1.00
SimvaBP68	10000	2936	3.41
SimvaAP4	10000	10000	1.00
SimvaAP39	10000	2560	3.91
SimvaDP9	3488	2040	1.71
SimvaDP20	132	302	0.44
SimvaDP62	2320	10000	0.23
SimvaCP11	3328	2320	1.43
SimvaCP43	10000	10000	1.00
SimvaCP65	2984	2400	1.24
SimvaCP77	2464	10000	0.25
SimvaCP87	1280	246	5.20
SimvaCP88	1488	2256	0.66
SimvaCP170	1896	1160	1.63
SimvaEP128	20000	10000	2.00
SimvaEP138	10000	20000	0.50
SimvaEP151	2640	3784	0.70
SimvaEP172	10000	20000	0.50

Second Experiment (Water)

Clone ID	Pos (plaque count)	Neg (plaque count)	ratio
SimvaBP12	79	57	1.39
SimvaBP13	20	25	0.80
SimvaBP14	7	5	1.40
SimvaBP24	17.8	39	0.46
SimvaBP35	436	5	87.20
SimvaBP36	5	890	0.01
SimvaBP57	452	11	41.09
SimvaBP58	1359	891	1.53
SimvaBP68	76	12	6.33
SimvaAP4	366	1012	0.36
SimvaAP39	193	257	0.75
SimvaDP9	105	119	0.88
SimvaDP20	42	9	4.67
SimvaDP62	10	84	0.12
SimvaCP11	4	250	0.02
SimvaCP43	483	1020	0.47
SimvaCP65	329	129	2.55
SimvaCP77	83	975	0.09
SimvaCP87	56	8	7.00
SimvaCP88	78	6	13.00
SimvaCP170	92	25	3.68
SimvaEP128	380	642	0.59
SimvaEP138	523	617	0.85
SimvaEP151	204	666	0.31
SimvaEP172	88	1113	0.08

Third Experiment (Water)

Clone ID	Pos (plaque count)	Neg (plaque count)	ratio
SimvaBP12	166	287	0.58
SimvaBP13	515	496	1.04
SimvaBP14	540	679	0.80
SimvaBP24	497	584	0.85
SimvaBP35	898	1298	0.69
SimvaBP36	717	376	1.91
SimvaBP57	294	76	3.87
SimvaBP58	251	456	0.55
SimvaBP68	-	4	0.00
SimvaAP4	-	-	0.00
SimvaAP39	370	2136	0.17
SimvaDP9	-	4	0.00
SimvaDP20	5	611	0.01
SimvaDP62	307	764	0.40
SimvaCP11	4	24	0.17
SimvaCP43	6	9	0.67
SimvaCP65	-	402	0.00
SimvaCP77	52	-	0.00
SimvaCP87	810	-	0.00
SimvaCP88	-	-	0.00
SimvaCP170	-	9	0.00
SimvaEP128	120	1176	0.10
SimvaEP138	149	573	0.26
SimvaEP151	644	499	1.29
SimvaEP172	504	968	0.52

Fourth Experiment (TBS)

Clone ID	Pos (plaque count)	Neg (plaque count)	Ratio
SimvaBP12	1286	780	1.65
SimvaBP13	1258	702	1.79
SimvaBP14	998	882	1.13
SimvaBP24	687	711	0.97
SimvaBP35	129	155	0.83
SimvaBP36	1132	1208	0.94
SimvaBP57	828	1180	0.70
SimvaBP58	1292	1212	1.07
SimvaBP68	1088	1236	0.88
SimvaAP4	1484	756	1.96
SimvaAP39	750	936	0.80
SimvaDP9	101	117	0.86
SimvaDP20	114	1252	0.09
SimvaDP62	1344	792	1.70
SimvaCP11	129	113	1.14
SimvaCP43	960	1488	0.65
SimvaCP65	1480	1028	1.44
SimvaCP77	131	107	1.22
SimvaCP87	123	120	1.03
SimvaCP88	85	122	0.70
SimvaCP170	1492	1080	1.38
SimvaEP128	1040	1212	0.86
SimvaEP138	960	1052	0.91
SimvaEP151	942	1104	0.85
SimvaEP172	760	70	10.86

Fifth Experiment (Water)

Clone ID	Pos (plaque count)	Neg (plaque count)	Ratio
SimvaBP12	1268	33	38.42
SimvaBP13	49	912	0.05
SimvaBP14	25	617	0.04
SimvaBP24	780	790	0.99
SimvaBP35	920	740	1.24
SimvaBP36	980	950	0.13
SimvaBP57	124	85	11.41
SimvaBP58	970	840	0.13
SimvaBP68	113	107	10.50
SimvaAP4	1124	97	1.03
SimvaAP39	100	69	3.78
SimvaDP9	261	67	1.06
SimvaDP20	71	97	0.73
SimvaDP62	71	96	0.88
SimvaCP11	84	88	0.78
SimvaCP43	69	82	1.30
SimvaCP65	107	81	0.99
SimvaCP77	80	440	0.22
SimvaCP87	97	101	0.79
SimvaCP88	80	80	1.04
SimvaCP170	83	89	1.13
SimvaEP128	101	107	0.70
SimvaEP138	75	83	0.94
SimvaEP151	78	83	0.94
SimvaEP172	78	90	0.87

Sixth Experiment (TBS)

Clone ID	Pos (plaque count)	Neg (plaque count)	Ratio
SimvaBP12	1316	556	2.37
SimvaBP13	1788	798	2.24
SimvaBP14	1816	2316	0.78
SimvaBP24	3864	2012	1.92
SimvaBP35	1320	15800	0.08
SimvaBP36	790	850	0.93
SimvaBP57	1130	52480	0.02
SimvaBP58	1824	1480	1.23
SimvaBP68	2684	3600	0.75
SimvaAP4	1656	1420	1.17
SimvaAP39	2800	2008	1.39
SimvaDP9	798	846	0.94
SimvaDP20	3856	1024	3.77
SimvaDP62	358	123	2.91
SimvaCP11	4312	333	12.95
SimvaCP43	1952	213	9.16
SimvaCP65	1824	1012	1.80
SimvaCP77	1664	184	9.04
SimvaCP87	1412	804	1.76
SimvaCP88	3440	1428	2.41
SimvaCP170	828	90	9.20
SimvaEP128	351	952	0.37
SimvaEP138	1328	77	17.25
SimvaEP151	774	321	2.41
SimvaEP172	1264	564	2.24

Comparison secondary screen using water, duplicates prepared simultaneously

SimvaBP35

Well Chemistry	Pos (plaque count)	Neg (plaque count)	ratio
B	199	374	0.53
B	1099	331	3.32
A	776	280	2.77
A	524	328	1.60
C	2344	788	2.97
C	10000	2112	4.73
B Neg	334	1976	0.17
B Neg	436	136	3.21
A Neg	194	254	0.76
A Neg	188	10000	0.02
C Neg	121	712	0.17
C Neg	243	1212	0.20

SimvaAP4

Well Chemistry	Pos (plaque count)	Neg (plaque count)	ratio
B	253	200	1.27
B	306	249	1.23
A	851	960	0.89
A	664	560	1.19
C	1022	1016	1.01
C	2156	1040	2.07
B Neg	194	10000	0.02
B Neg	135	2616	0.05
A Neg	90	206	0.44
A Neg	3116	205	15.20
C Neg	176	235	0.75
C Neg	419	217	1.93

SimvaCP88

Well Chemistry	Pos (plaque count)	Neg (plaque count)	ratio
B	1776	110	16.15
B	331	134	2.47
A	231	492	0.47
A	350	395	0.89
C	558	2976	0.19
C	1256	1232	1.02
B Neg	145	136	1.07
B Neg	131	101	1.30
A Neg	260	224	1.16
A Neg	122	131	0.93
C Neg	136	161	0.84
C Neg	544	100	5.44

SimvaBP35	Ratio 1	Ratio 2
B	0.53	3.32
A	2.77	1.60
C	2.97	4.73
B Neg	0.17	3.21
A Neg	0.76	0.02
C Neg	0.17	0.20

SimvaAP4	Ratio 1	Ratio 2
B	1.27	1.23
A	0.89	1.19
C	1.01	2.07
B Neg	0.02	0.05
A Neg	0.44	15.20
C Neg	0.75	1.93

SimvaCP88	Ratio 1	Ratio 2
B	16.15	2.47
A	0.47	0.89
C	0.19	1.02
B Neg	1.07	1.30
A Neg	1.16	0.93
C Neg	0.84	5.44

	SimvaBP35		SimvaAP4		SimvaCP88	
Chemistry	Ratio 1	Ratio 2	Ratio 1	Ratio 2	Ratio 1	Ratio 2
B	0.53	3.32	1.27	1.23	16.15	2.47
A	2.77	1.60	0.89	1.19	0.47	0.89
C	2.97	4.73	1.01	2.07	0.19	1.02
B Neg	0.17	3.21	0.02	0.05	1.07	1.30
A Neg	0.76	0.02	0.44	15.20	1.16	0.93
C Neg	0.17	0.20	0.75	1.93	0.84	5.44

8.5.1 M001 Biopan *vs.* Human Vascular Tissue Library: Phage Titres

Before Biopanning

Dilution Factor	Number of Plaques	Volume µl	pfu/ml $x10^{10}$
10^{-5}	lysis	90	N/A
10^{-6}	1672	90	1.86
10^{-7}	123	90	1.37
10^{-8}	15	100	1.67

Round 1 titres:

Dilution Factor 10^{-2}	Number of Plaques in 100 µl				
	PCB	PCA	PCD	PCC	PCE
Wash A	952	1036	1240	1150	944
Wash B	271	572	1032	713	597
Wash C	1532	N/A	1428	N/A	N/A
Wash D	2144	N/A	N/A	1222	2404

Repeats

Dilution Factor 10^{-3}	Number of Plaques in 100 µl				
	PCB	PCA	PCD	PCC	PCE
Wash C	N/A	171	N/A	105	182
Wash D	N/A	383	114	N/A	N/A

	pfu/ml $x10^{5}$				
	PCB	PCA	PCD	PCC	PCE
Wash A	9.52	10.4	12.4	11.5	9.44
Wash B	2.71	5.72	10.3	7.13	5.97
Wash C	15.3	17.1	14.3	10.5	18.2
Wash D	21.4	38.3	11.4	12.2	24.0

Round 2 titres:

Dilution Factor 10^{-2}	Number of Plaques in 100 µl				
	PCB	PCA	PCD	PCC	PCE
Wash A	592	500	1035	500	445
Wash B	291	631	700	470	500
Wash C	508	777	700	618	500
Wash D	163	145	6	215	81

	pfu/ml $x10^5$				
	PCB	PCA	PCD	PCC	PCE
Wash A	5.92	5.00	10.4	5.00	4.45
Wash B	2.91	6.31	7.00	4.70	5.00
Wash C	5.08	7.77	7.00	6.18	5.00
Wash D	1.63	1.45	0.06	2.15	0.81

Round 3 titres:

Dilution Factor 10^{-2}	Number of Plaques in 100 µl				
	PCB	PCA	PCD	PCC	PCE
Wash A	551	1657	413	1037	N/A
Wash B	434	573	389	84	N/A
Wash C	477	354	304	211	179
Wash D	332	232	460	1412	126

Dilution Factor 10^{-3}	Number of Plaques in 100 µl				
	PCB	PCA	PCD	PCC	PCE
Wash A	64	134	68	117	617
Wash B	55	48	52	15	400
Wash C	N/A	N/A	N/A	N/A	N/A
Wash D	29	20	42	195	42

	pfu/ml $x10^5$				
	PCB	PCA	PCD	PCC	PCE
Wash A	5.96	15.0	5.47	11.0	61.7
Wash B	4.92	5.27	4.55	1.17	40.0

Wash C	4.77	3.54	3.04	2.11	1.79
Wash D	3.11	2.16	4.40	16.8	2.73

ROUND 1	pfu/ml x10^5					Washing
	PCB	PCA	PCD	PCC	PCE	AVERAGE
Wash A	9.52	10.4	12.4	11.5	9.44	10.6
Wash B	2.71	5.72	10.3	7.13	5.97	6.37
Wash C	15.3	17.1	14.3	10.5	18.2	15.1
Wash D	21.4	38.3	11.4	12.2	24.0	21.5

ROUND 2	pfu/ml x10^5					Washing
	PCB	PCA	PCD	PCC	PCE	AVERAGE
Wash A	5.92	5.00	10.4	5.00	4.45	6.14
Wash B	2.91	6.31	7.00	4.70	5.00	5.18
Wash C	5.08	7.77	7.00	6.18	5.00	6.21
Wash D	1.63	1.45	0.06	2.15	0.81	1.22

ROUND 3	pfu/ml x10^5					Washing
	PCB	PCA	PCD	PCC	PCE	AVERAGE
Wash A	5.96	15.0	5.47	11.0	61.7	19.8
Wash B	4.92	5.27	4.55	1.17	40.0	11.2
Wash C	4.77	3.54	3.04	2.11	1.79	3.05
Wash D	3.11	2.16	4.40	16.8	2.73	5.84

MT Average	6.94	9.83	7.52	7.54	14.9

8.5.2 M001 Biopan *vs.* Human Vascular Tissue Library: Insert Length Data

Prepanning Library (M001) Insert Lengths

Lane	Well identity	Size of Insert
1	A1	450
2	B1	450
3	C1	200
4	D1	550
5	E1	No Band
6	F1	350
7	G1	700
8	H1	450
9	A2	<200
10	B2	<200
11	C2	400
12	D2	400
13	E2	250
14	F2	300
15	G2	320
16	H2	380
17	A3	420
18	B3	200
19	C3	400
20	D3	300
21	E3	<200
22	F3	450
23	G3	200
24	TC1	350
25	TC2	(350) + 400 + 550
26	negative	

Round 1 Wash A Insert Lengths

Chemistry B Wash A	Chemistry A Wash A	Chemistry D Wash A	Chemistry C Wash A	Chemistry E Wash A
200	200	200	200	200
200	220	200	200	200
200	220	200	200	300
200	280	200	200	300
250	280	250	200	300
250	300	250	200	320
300	300	250	200	320
300	320	300	210	350
320	350	300	300	350
320	350	300	300	380
320	380	300	320	380
350	390	320	320	380
350	400	320	320	380
350	400	320	350	400
350	420	350	350	420
350	450	350	350	450
420	550	370	350	520
500	550	380	350	600
500	600	380	350	650
550	750	400	380	750
550	780	400	380	750
550	950	410	500	1000
600		450	500	
750		550	500	
		580	550	
		600	550	
		600	600	
		620	700	
		900	700	
			700	
<200	<200	<200	<200	<200
2	5	0	2	7
unknown	unknown	unknown	unknown	unknown
7	4	2	3	2

Round 1 Wash B Insert Lengths

Chemistry B Wash B	Chemistry A Wash B	Chemistry A Wash B PCR Repeat	Chemistry D Wash B	Chemistry D Wash B PCR Repeat	Chemistry C Wash B	Chemistry C Wash B PCR Repeat	Chemistry E Wash B
200	200	200	200	200	200	200	200
200	200	210	200	220	200	200	230
200	220	220	200	250	200	200	300
200	250	250	200	250	200	220	300
200	300	300	200	250	220	220	300
220	350	310	280	300	250	220	320
250	350	320	300	320	300	250	320
280	350	320	300	320	320	280	320
280	350	330	300	330	350	280	350
300	350	350	300	350	350	320	380
320	380	370	320	350	380	320	400
320	450	380	350	350	400	320	420
320	450	380	350	350	420	320	420
350	450	380	350	350	600	320	450
350	450	380	350	350		350	550
350	600	450	400	350		350	700
350		450	450	370		380	950
380		450	700	380		380	
450		450		380		380	
500		500		380		400	
550		600		550		400	
550		700		600		450	
600		700		600		450	
700				750		500	
				750		900	
<200	<200	<200	<200	<200	<200	<200	<200
4	4	4	4	2	1	3	2
unknown	unknown	unknown	unknown	unknown	unknown	unknown	unknown
6	12	7	10	6	16	5	10

Round 1 Wash C Insert Lengths

Chemistry B Wash C	Chemistry A Wash C	Chemistry D Wash C	Chemistry C Wash C	Chemistry E Wash C
200	200	200	200	200
320	200	220	220	200
320	200	220	220	200
350	210	220	220	280
350	300	220	250	300
350	320	220	300	350
350	330	250	300	380
380	350	250	300	380
380	350	300	300	380
420	380	320	300	400
750	400	330	320	400
	420	350	350	450
	550	350	350	450
	600	350	350	450
	650	380	350	450
	750	380	380	500
	1600	380	380	600
		380	400	
		400	450	
		400	500	
		400		
		400		
		450		
		500		
		500		
		620		
		700		
		700		
<200	<200	<200	<200	<200
1	7	6	3	4
unknown	unknown	unknown	unknown	unknown
16	5	4	11	10

Round 2 Wash A Insert Lengths

Chemistry B Wash A	Chemistry A Wash A	Chemistry D Wash A	Chemistry C Wash A	Chemistry E Wash A
200	200	200	310	200
200	210	310	310	210
200	210	320	330	230
210	300	330	330	250
230	320	330	350	290
250	320	350	350	290
300	330	350	370	310
300	350	350	380	310
350	350	370	380	310
350	350	370	380	310
350	350	380	400	310
380	350	380	420	310
380	350	380	420	330
380	350	380	550	330
380	380	380	550	330
380	380	450	580	330
380	400	750	600	330
380	430	800	650	330
400	500		700	330
650	520		1000	350
650	700		1800	350
650				350
700				370
750				370
780				370
				380
				400
				500
				550
				700
				700
<200	<200	<200	<200	<200
5	8	9	7	3
unknown	unknown	unknown	unknown	unknown
3	2	3	4	1

Round 2 Wash B Insert Lengths

Chemistry B Wash B	Chemistry A Wash B	Chemistry D Wash B	Chemistry C Wash B	Chemistry E Wash B
210	220	250	210	200
210	250	290	230	210
210	300	290	270	230
230	300	300	270	270
300	300	320	290	270
320	300	330	310	310
350	300	330	310	310
350	310	350	310	350
350	330	370	320	350
370	350	370	330	350
370	350	370	330	370
370	350	370	330	370
370	350	370	330	370
370	350	370	350	400
380	350	380	350	450
380	350	500	350	500
390	370	500	370	600
390	450	550	400	600
450	700	580	450	650
500	700	650	500	650
500	750	650	500	1800
500	1800	750	600	
650	1800	750	600	
		750	650	
			1700	
<200	<200	<200	<200	<200
6	5	10	8	9
unknown	unknown	unknown	unknown	unknown
1	3	1	1	0

Round 2 Wash C Insert Lengths

Chemistry B Wash C	Chemistry A Wash C	Chemistry D Wash C	Chemistry C Wash C	Chemistry E Wash C
210	210	200	210	230
270	230	210	230	270
300	270	300	290	290
310	290	300	300	290
330	310	300	300	310
330	330	350	310	310
330	330	350	350	310
330	330	350	350	310
330	330	350	350	310
350	330	350	370	310
370	370	350	370	330
370	370	350	370	330
370	380	350	370	350
370	390	380	370	350
370	400	380	380	350
370	450	380	380	370
400	450	380	400	370
450	500	380	500	380
500	700	450	700	450
600		600	700	450
600		600	850	550
650		650	1600	600
750		750		650
1700				750
<200	<200	<200	<200	<200
8	12	7	6	7
unknown	unknown	unknown	unknown	unknown
1	1	1	3	0

Round 3 Wash A Insert Lengths

Chemistry B Wash A	Chemistry A Wash A	Chemistry D Wash A	Chemistry C Wash A	Chemistry E Wash A
200	200	200	210	250
210	200	290	300	290
210	220	310	310	290
220	300	330	320	300
300	300	350	350	310
300	310	370	350	350
320	330	370	350	370
320	350	380	350	380
330	350	380	380	380
350	350	380	400	380
350	370	430	400	380
350	370	450	450	380
350	380	650	500	500
370	380	800	500	510
400	380	800	500	550
600	380	1000	500	700
700	380	1600	510	750
700	400	1700	520	1700
750	420	1800	520	
1600	420		520	
1700	450		700	
1700	550			
	700			
	750			
<200	<200	<200	<200	<200
8	5	8	6	12
unknown	unknown	unknown	unknown	unknown
2	4	5	4	3

Round 3 Wash B Insert Lengths

Chemistry B Wash B	Chemistry A Wash B	Chemistry D Wash B	Chemistry C Wash B	Chemistry E Wash B
300	300	310	210	210
310	310	310	290	290
310	320	310	290	300
330	320	330	290	310
330	320	330	310	310
330	350	330	310	320
330	370	330	310	330
350	430	370	310	330
370	450	370	310	330
370	600	370	320	330
450	650	410	330	330
600	650	430	350	330
650	650	800	350	330
900	700	800	370	330
1500	1500	1500	430	330
1500	1600	1600	430	330
	1600	1600	470	350
	1600	1800	600	350
	1600		650	370
	1700		1600	390
	1700		1600	430
	1800			510
	1800			650
				650
				650
				900
				1700
				1700
				1700
<200	<200	<200	<200	<200
13	7	10	9	6
unknown	unknown	unknown	unknown	unknown
2	1	3	2	0

Round 3 Wash C Insert Lengths

Chemistry B Wash C	Chemistry A Wash C	Chemistry D Wash C	Chemistry C Wash C	Chemistry E Wash C
290	210	200	320	210
310	320	300	330	300
310	330	300	330	330
330	330	330	330	330
330	330	350	350	330
330	330	380	350	330
330	330	380	350	330
330	350	480	350	330
330	350	500	350	330
370	350	700	350	330
370	370	1500	370	350
370	370	1500	370	370
390	370	1600	370	370
410	380	1600	370	370
600	400	1800	370	370
650	550	1800	370	370
650	600	1800	380	370
700	650		380	390
1500	650		390	430
1600	650		450	750
1700	750		550	1700
	1900		650	1700
	2000		800	
			1600	
			1700	
<200	<200	<200	<200	<200
7	7	8	5	10
unknown	unknown	unknown	unknown	unknown
2	2	7	0	3

Round 3 Control Wells Insert Lengths

Chemistry B	Chemistry A	Chemistry D	Chemistry C	Chemistry E
370	330	300	320	310
370	330	350	350	310
500	330	350	350	850
500	390	350	350	1600
550	390	350	380	1600
550	390	350	380	1600
1600	390	350	380	1700
1800	390	350	500	1700
3700	1600	350	1600	1700
	1700	350	1600	1900
	1800	1600	1600	1900
<200	<200	<200	<200	<200
4	1	1	1	0
unknown	unknown	unknown	unknown	unknown
1	0	0	0	1

Round 3 After Amplification

Insert Size	Count
<200	3
300	1
350	22
400	2
600	2
650	2
800	1

M001 Clones Sent For Sequencing

Sequencing Plate 1

	A	B	C	D	E	F	G	H
1	R3 Chem B WashC	R3 Chem A WashB	R3 Chem B WashA	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
2	R3 Chem B WashC	R3 Chem A WashB	R3 Chem B WashA	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
3	R3 Chem E WashB	R3 Chem A WashB	R3 Chem B WashA	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
4	R3 Chem E WashB	R3 Chem A WashB	R2 Chem C WashC	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
5	R3 Chem C WashB	R3 Chem A WashB	R2 Chem B WashC	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
6	R3 Chem D WashB	R3 Chem A WashB	R2 Chem E WashB	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
7	R3 Chem D WashB	R3 Chem E WashA	R2 Chem A WashB	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
8	R3 Chem D WashB	R3 Chem D WashA	R2 Chem A WashB	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
9	R3 Chem D WashB	R3 Chem D WashA	R2 Chem C WashA	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
10	R3 Chem A WashB	R3 Chem D WashA	R2 Chem C WashA	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
11	R3 Chem A WashB	R3 Chem B WashB	R1 Chem A WashC	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control
12	R3 Chem A WashB	R3 Chem B WashB	R1 Chem E WashA	R3 Chem E Control	R3 Chem C Control	R3 Chem D Control	R3 Chem A Control	R3 Chem B Control

Sequencing Plate 2

	A	B	C	D	E	F	G	H
1	R3 Chem B WashC	R1 Chem E WashA	R2 Chem A WashA	R3 Chem C WashA	R1 Chem E WashA	R2 Chem B WashA	R3 Chem A WashB	Prepanning Library
2	R3 Chem A WashC	R1 Chem B WashB	R2 Chem E WashB	R3 Chem C WashA	R1 Chem A WashB	R2 Chem D WashB	R3 Chem D WashB	Prepanning Library
3	R3 Chem A WashC	R1 Chem E WashB	R2 Chem D WashC	R3 Chem C WashA	R1 Chem B WashC	R2 Chem E WashC	R3 Chem E WashB	Prepanning Library
4	R3 Chem D WashC	R1 Chem B WashA	R2 Chem C WashC	R3 Chem C WashA	R1 Chem A WashA	R2 Chem B WashC	R3 Chem B WashC	Prepanning Library
5	R3 Chem D WashC	R1 Chem B WashA	R2 Chem C WashA	R3 Chem E WashB	R1 Chem E WashA	R2 Chem B WashA	R3 Chem E WashA	Prepanning Library
6	R3 Chem D WashC	R1 Chem A WashB	R2 Chem A WashB	R3 Chem B WashA	R1 Chem A WashC	R2 Chem D WashC	R3 Chem E WashC	Prepanning Library
7	R3 Chem C WashC	R1 Chem D WashB	R2 Chem A WashB	R3 Chem E WashA	R1 Chem A WashB	R2 Chem B WashC	R3 Chem A WashC	R2 Chem C WashB
8	R3 Chem C WashC	R1 Chem E WashA	R2 Chem D WashC	R3 Chem C WashC	R1 Chem E WashB	R2 Chem B WashB	R3 Chem D WashA	R2 Chem C WashB
9	R3 Chem D WashC	R1 Chem D WashB	R2 Chem D WashC	R3 Chem C WashC	R1 Chem E WashB	R2 Chem B WashC	R3 Chem D WashA	R3 Chem C WashA
10	R3 Chem D WashC	R1 Chem D WashB	R2 Chem D WashC	R3 Chem C WashC	R1 Chem C WashA	R2 Chem B WashA	R3 Chem D WashA	R3 Chem B WashA
11	R3 Chem D WashC	R1 Chem D WashB	R2 Chem D WashA	R3 Chem C WashC	R1 Chem B WashA	R2 Chem C WashC	R3 Chem B WashC	R3 Chem D WashA
12	R3 Chem D WashC	R1 Chem E WashA	R2 Chem E WashA	R3 Chem C WashC	R1 Chem B WashB	R2 Chem D WashB	R3 Chem B WashC	R3 Chem D WashA

Sequencing Plate 1 Estimated Insert Lengths

	A	B	C	D	E	F	G	H
1	1500	1500	1700	850	1600	350	380/400*	<200
2	1700	1600	1700	1600	350	350	380/400*	350/380*
3	1700	1600	1600	1700	380	<200	320/350*	500/550*
4	1700	1800	1600	1600	350	350	1600	<200
5	1600	1700	1700	300/320*	<200	350	380/400*	500/550*
6	1600	1600	1800	1900	500	1600	380/400*	<200
7	1800	1700	1800	300/320*	380	350	1800	350/380*
8	1600	1600	1800	1600	1600	300	1700	<200
9	1500	1800	1800	1900	1600	350	380/400*	1800
10	1800	1700	1000	1700	350	350	300/320*	350/380*
11	1700	380	1600	1700	380	350	<200	1600
12	1600	No Band	1000	No Band	320	350	320/350*	No Band

Sequencing Plate 2 Estimated Insert Lengths

	A	B	C	D	E	F	G	H
1	1600	350	350	350	320	750	650	450
2	2000	<200	350	350	700	750	300/320*	450
3	1900	350	350	350	750	750	650	450
4	1800	350	350	350	750	750	650	450
5	1800	350	350	350	650	650	750	200
6	1600	350	350	350	650	650	200/220*	200
7	1700	350	350	350	700	650	750	220/250*
8	1600	350	350	350	700	650	800	450/500*
9	1600	350	350	350	<200	<200	<200	500/520*
10	1500	350	350	350	<200	<200	<200	300
11	1500	350	350	350	<200	<200	<200	800
12	1800	350	350	350	<200	<200	<200	380

* Observed as doubled band due to percentage of agarose in gel

8.5.3 M001 Biopan *vs.* Human Vascular Tissue Library Sequences

32 Pages, Sequences output from rhein hydroxymate (M001) biopan *vs.* human vascular tissue library.

Electronic copy only.

8.5.4 M001 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector

22 Pages, BLAST alignment of two sequences (bl2seq algorithm). Alignment of sequences output from rhein hydroxamate (M001) biopan *vs.* human vascular tissue library against T7Select® vector sequence.

Electronic copy only.

8.5.5 Lasergene® SeqMan Analysis of M001 Biopan *vs.* Human Vascular Tissue Library Sequences

7 Pages, Lasergene® SeqMan output of contiguous regions, after alignment of all sequences output from rhein hydroxamate (M001) biopan *vs.* human vascular tissue library.

Electronic copy only.

8.5.6 M001 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment

75 Pages, BLASTX algorithm text output of rhein hydroxamate (M001) contigs (Appendix 8.4.3) aligned against RefSeq protein database.

Electronic copy only.

8.6.1 M002 Biopan *vs.* Human Vascular Tissue Library Sequences

7 Pages, Sequences output from oxametacin (M002) biopan *vs.* human vascular tissue library.

Electronic copy only.

8.6.2 M002 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector

15 Pages, BLAST alignment of two sequences (bl2seq algorithm). Alignment of sequences output from oxametacin (M002) biopan *vs.* human vascular tissue library against T7Select® vector sequence.

Electronic copy only.

8.6.3 Lasergene® SeqMan Analysis of M002 Biopan *vs.* Human Vascular Tissue Library Sequences (against M00 negatives)

5 Pages, Lasergene® SeqMan output of contiguous regions, after alignment of all sequences output from oxametacin (M002) biopan *vs.* human vascular tissue library. Electronic copy only.

8.6.4 M002 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment

44 Pages, BLASTX algorithm text output of oxametaxin (M002) contigs (Appendix 8.4.3) aligned against RefSeq protein database.

Electronic copy only.

8.7.1 M003 Biopan *vs.* Human Vascular Tissue Library Sequences

10 Pages, Sequences output from protizinic acid (M003) biopan *vs.* human vascular tissue library.

Electronic copy only.

8.7.2 M003 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector

18 Pages, BLAST alignment of two sequences (bl2seq algorithm). Alignment of sequences output from protizinic acid (M003) biopan *vs.* human vascular tissue library against T7Select® vector sequence.

Electronic copy only.

8.7.3 Lasergene® SeqMan Analysis of M003 Biopan *vs.* Human Vascular Tissue Library Sequences

7 Pages, Lasergene® SeqMan output of contiguous regions, after alignment of all sequences output from protizinic acid (M003) biopan *vs.* human vascular tissue library.

Electronic copy only.

8.7.4 M003 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment

82 Pages, BLASTX algorithm text output of protizinic acid (M003) contigs (Appendix 8.4.3) aligned against RefSeq protein database.

Electronic copy only.

8.8.1 M004 Biopan *vs.* Human Vascular Tissue Library Sequences

9 Pages, Sequences output from rebamipide (M004) biopan *vs.* human vascular tissue library.

Electronic copy only.

8.8.2 M004 Biopan *vs.* Human Vascular Tissue Library Sequences Aligned against T7Select® Vector

21 Pages, BLAST alignment of two sequences (bl2seq algorithm). Alignment of sequences output from rebamipide (M004) biopan *vs.* human vascular tissue library against T7Select® vector sequence.

Electronic copy only.

8.8.3 Lasergene® SeqMan Analysis of M004 Biopan *vs.* Human Vascular Tissue Library Sequences

6 Pages, Lasergene® SeqMan output of contiguous regions, after alignment of all sequences output from rebamipide (M004) biopan *vs.* human vascular tissue library. Electronic copy only.

8.8.4 M004 Biopan *vs.* Human Vascular Tissue Library Contigs BLASTX Alignment

65 Pages, BLASTX algorithm text output of rebampide (M004) contigs (Appendix 8.4.3) aligned against RefSeq protein database.

Electronic copy only.

8.9.1 M002/3/4 Negatives Biopan *vs.* Human Vascular Tissue Library Sequences

7 Pages, Sequences output from antiinflammatories biopan (control wells) *vs.* human vascular tissue library.

Electronic copy only.

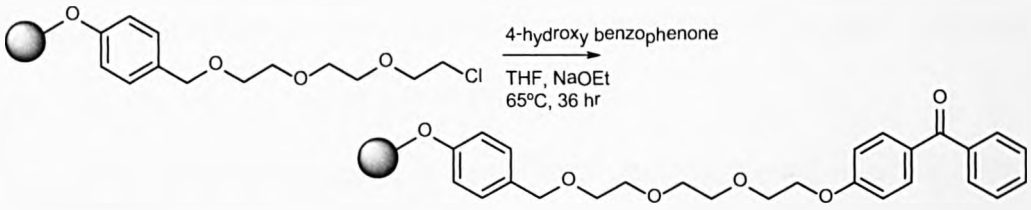
8.9.2 M002/3/4 Negatives Biopan *vs.* Human Vascular Tissue Library Sequences Aligned Against T7Select® Vector

9 Pages, BLAST alignment of two sequences (bl2seq algorithm). Alignment of sequences output from antiinflammatories biopan (control wells) *vs.* human vascular tissue library against T7Select® vector sequence.

Electronic copy only.

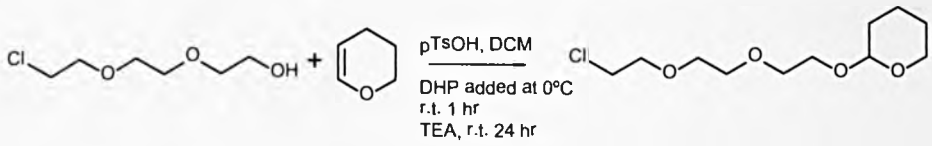
8.10.1 Extra Chemistry Experimental

8.10.1a Unsuccesful Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethyl-4-benzoyl benzamide

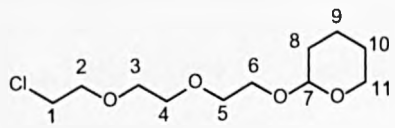


Wang immobilised 2-(2-[2-ethoxy]ethoxy)ethoxy chloride (192 mg at max 0.9mmol/g loading, max 0.18 mmol) was placed in a sealed foil-covered SPE tube with 10 mL dry THF and the beads were allowed to swell with agitation for 5 mins. 4-hydroxybenzophenone (376 mg, 1.8 mmol, 10 eq) and sodium ethoxide (123 mg, 1.8 mmol, 10 eq) were dissolved in 5 mL ethanol and added to the reaction mixture, which was agitated at 65 °C for 36 h. Reaction mixture was allowed to return to room temperature, the solvent was removed and the beads washed following the general procedure. Beads were stored in a foil-covered vial, protected from the light. This reaction is proceeding at very low yield as observed by NMR.

8.10.1b Preparation of 2-[2-(2-[2-chloroethoxy]ethoxy)ethoxy]tetrahydropyran[1]



2-(2-[2-(2-chloroethoxy)ethoxy]ethoxy)tetrahydropyran was prepared as a clear oil (78.3 %) following literature procedures.



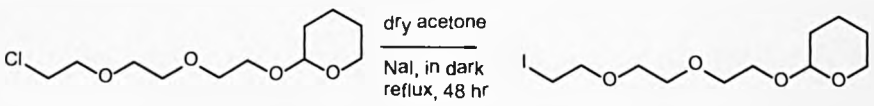
^{1}H NMR ($CDCl_3$) δ 1.54 (m, 3H, 10_{eq} + 9), 1.60 (m, 1H, 8_{eq}), 1.70 (m, 1H, 8_{ax}), 1.82 (m, 1H, 10_{ax}), 3.49 (m, 1H, 11_{eq}), 3.59 (m, 1H, 6_{eq}), 3.61 (t, J^3_{H-H}5 Hz, 2H, 1), 3.67 (m, 6H, 3 + 4 + 5), 3.75 (t, J^3_{H-H} 5 Hz, 2H, 2), 3.85 (m, 2H, 6_{ax} + 11_{ax}), 4.62 (dd, J^3_{H-H} 4 Hz, 2 Hz, 1H, 7)

^{13}C NMR ($CDCl_3$) δ 19.30 (10), 25.41 (9), 30.55 (8), 42.69 (1), 62.17 (11), 66.62 (6), 70.60 (3 + 4 + 5), 71.36 (2), 98.90 (7)

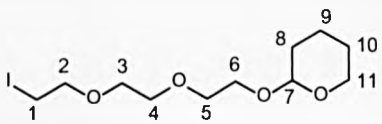
ν_{max}: 2940, 2868 (C-H saturated), 1120, 1074 (aliphatic ethers), 665 (CH_2Cl)

m/z (%): 275.1 (100)

8.10.1c Preparation of 2-[2-(2-[2-iodoethoxy]ethoxy)ethoxy]tetrahydropyran[2]



2-(2-[2-(2-iodoethoxy)ethoxy]ethoxy)tetrahydropyran was prepared as a yellow oil (51.7 %) following literature procedures.



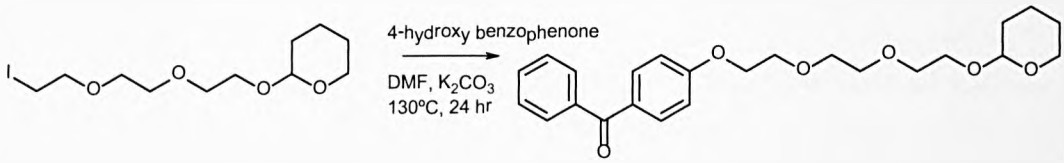
^{1}H NMR ($CDCl_3$) δ 1.47 (m, 3H, 10_{eq} + 9), 1.52 (m, 1H, 8_{eq}), 1.66 (m, 1H, 8_{ax}), 1.77 (m, 1H, 10_{ax}), 3.20 (t, $^3J_{H-H}$ 6.9 Hz, 2H, 1), 3.44 (m, 1H, 11_{eq}), 3.55 (m, 1H, 6_{eq}), 3.61 (m, 6H, 3 + 4 + 5), 3.71 (t, $^3J_{H-H}$ 6.9 Hz, 2H, 2), 3.80 (m, 2H, 6_{ax} + 11_{ax}), 4.57 (dd, $^3J_{H-H}$ 4.5 Hz, 3.7 Hz, 1H, 7)

^{13}C NMR ($CDCl_3$) δ 3.18 (1), 19.70 (10), 25.63 (9), 30.77 (8), 62.43 (11), 66.85 (6), 70.45 (5), 70.81 (3 + 4), 72.19 (2), 99.14 (7)

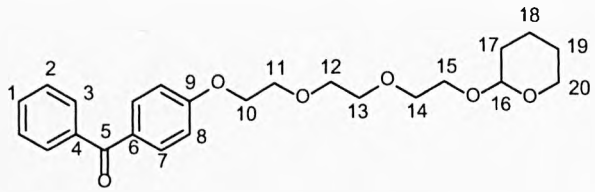
m/z (%): 367 (Na^+ 100)

ν_{max}: 2937, 2866 (C-H saturated), 1120, 1074 (aliphatic ethers), no peak at 600-800 (CH_2Cl)

8.10.1d Preparation of 2-[2-(2-[2-tetrahydropyranoxyethoxy]ethoxy)ethoxy]benzophenone[4]



2-[2-(2-[2-tetrahydropyranoxyethoxy]ethoxy)ethoxy]benzophenone was prepared as a pale yellow oil (39%) following literature procedures.



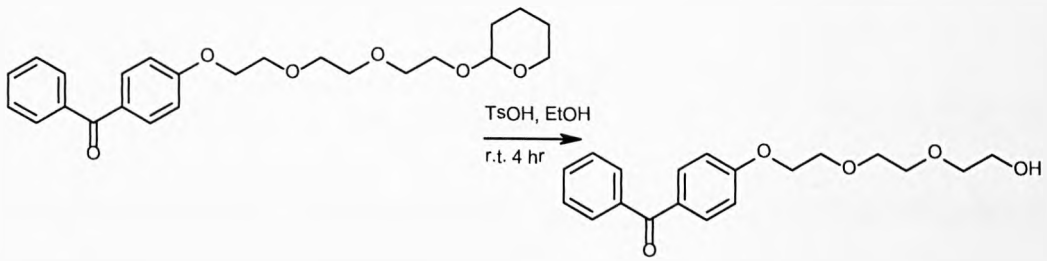
^{1}H NMR ($CDCl_3$) δ 1.52 (m, 3H, 18 + 19_{eq}), 1.61 (m, 1H, 17_{eq}), 1.73 (m, 1H, 17_{ax}), 1.84 (m, 1H, 19_{ax}), 3.51 (m, 1H, 20_{eq}), 3.63 (m, 1H, 15_{eq}), 3.73 (m, 6H, 12, 13, 14), 3.91 (m, 2H, 15_{ax}, 20_{ax}), 3.93 (t, $J^3_{H\text{-}H}$4.6 Hz, 2H, 11), 4.24 (t, $J^3_{H\text{-}H}$4.6 Hz, 2H, 10), 4.64 (t, $J^3_{H\text{-}H}$3.3 Hz, 1H, 16), 7.00 (d, $J^3_{H\text{-}H}$7.9 Hz, 2H, 8), 7.49 (t, $J^3_{H\text{-}H}$7.9 Hz, 2H, 2), 7.58 (t, $J^3_{H\text{-}H}$7.9 Hz, 1H, 1), 7.77 (d, $J^3_{H\text{-}H}$7.9 Hz, 2H, 3), 7.84 (d, $J^3_{H\text{-}H}$7.9 Hz, 2H, 7)

^{13}C NMR ($CDCl_3$) δ 195.53 (5), 162.49 (9), 138.30 (4), 132.53 (7), 131.89 (1), 130.28 (6), 129.74 (3), 128.20 (2), 114.15 (8), 98.98 (16), 70.96/70.70/70.63 (12-14), 69.57 (11), 67.68 (10), 66.67 (15), 62.27 (20), 30.59 (17), 25.44 (18), 19.56 (19)

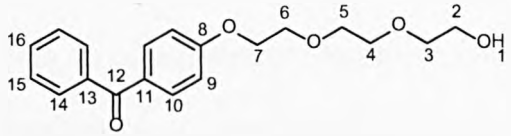
m/z (%): 437.2 (100)

ν_{max}: 2939, 2869 (C-H stretching), 1650 (carbonyl, diaryl), 1597, 1577, 1504 (aromatic ring),1121, 1077 (aliphatic ethers)

8.10.1c Deprotection of 2-[2-(2-[2-tetrahydropyranoxyethoxy]ethoxy)ethoxy]benzophenone[4]



Triethylene glycol benzophenone was prepared as a yellow oil (83.9 %) following literature procedures.

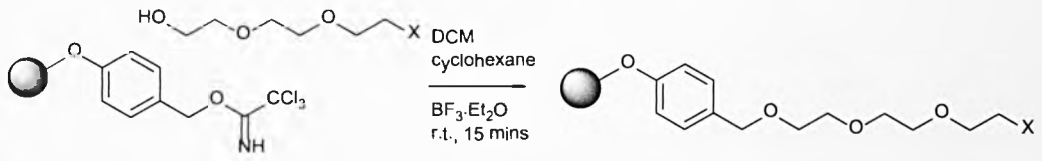


^{1}H NMR ($CDCl_3$) δ 3.37 (br s, 1H, 1), 3.61 (m, 2H, 2), 3.70 (m, 2H, 6), 3.73 (m, 4H, 4/5), 3.88 (m, 2H, 3), 4.21 (m, 2H, 7), 6.98 (d, $^{3}J_{H\text{-}H}$ 8.7 Hz, 2H, 9), 7.46 (t, $^{3}J_{H\text{-}H}$ 7.6 Hz, 2H, 15), 7.55 (t, $^{3}J_{H\text{-}H}$ 7.6 Hz, 1H, 16), 7.74 (d, $^{3}J_{H\text{-}H}$ 7.6 Hz, 2H, 14), 7.81 (d, $^{3}J_{H\text{-}H}$ 8.7 Hz, 2H, 10)

^{13}C NMR ($CDCl_3$) δ 62.26 (2), 66.67 (6), 67.68 (7), 69.56 (3), 70.63/70.96 (4/5), 114.17 (9), 128.20 (15), 129.76 (14), 130.30 (11), 131.91 (16), 132.52 (10), 138.30 (13), 162.49 (8), 195.37 (12)

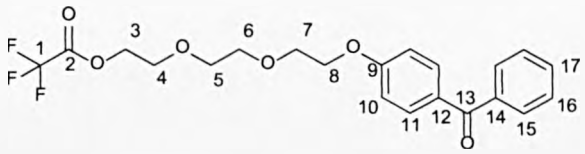
m/z (%): 331.2 (100)

8.10.1f Preparation of Wang-immobilised 2-(2-[2-ethoxy]ethoxy)ethyl-4-benzoyl benzamide, x= 4-benzoyl benzamide



Wang-immobilised trichloroacetimidate (max. loading 0.9 mmol/g) was placed in a sealed SPE tube with dry DCM (0.1 mL/mg resin) and the beads were allowed to swell with agitation for 5 mins. End functionalised triethylene glycol (5 eq) was added in cyclohexane (0.01 mL/mg resin). Borontrifluoride diethyletherate (1 eq) was added and agitated for fifteen minutes. The solvents were removed and the beads washed following the general procedure.

2-(2-[2-(2,2,2-Trifluoro-acetoxy)ethoxy]ethoxy) ethyl-4-benzoyl benzamide:



As final product, see main text.