



University of Bradford eThesis

This thesis is hosted in [Bradford Scholars](#) – The University of Bradford Open Access repository. Visit the repository for full metadata or to contact the repository team



© University of Bradford. This work is licenced for reuse under a [Creative Commons Licence](#).

**SYNTHESIS OF BESPOKE MATRICES TO INVESTIGATE A
NOVEL ANTI-TUMOUR MOLECULAR TARGET USING AFFINITY
CHROMATOGRAPHY**

H. R. EVANS

PhD

UNIVERSITY OF BRADFORD

2010

**SYNTHESIS OF BESPOKE MATRICES TO INVESTIGATE A
NOVEL ANTI-TUMOUR MOLECULAR TARGET USING AFFINITY
CHROMATOGRAPHY**

The design, synthesis and evaluation of biotinylated
biarylheterocycles used as novel affinity probes in the identification
of anti-tumour molecular targets

Hayley Ruth EVANS

Submitted for the degree
of Doctor of Philosophy

Institute of Cancer Therapeutics and School of Pharmacy

School of Life Sciences

University of Bradford

2010

ABSTRACT

Three novel, synthetic biarylheterocycles bearing imidazole terminal groups had previously been discovered with high cytotoxicity (IC_{50} 16–640 nM) against a number of human tumour cell lines. Notably, this biological activity was independent of duplex DNA binding affinity. The compounds were tested in the NCI 60-cell line panel and COMPARE analysis suggests they have a novel mechanism of action, targeting the product of a 'gene-like sequence' of unidentified function.

The identity of likely protein targets was explored using a chemical proteomic strategy. Bespoke affinity matrices for chromatography were prepared in which test compounds were attached to a solid support through a biotin tag. A synthetic route to hit compounds containing a biotin moiety in place of one of the imidazole sidechains was developed. Chemosensitivity studies confirmed that the biotinylated compounds retained their activity showing $IC_{50} = 6.25 \mu\text{M}$ in a susceptible cell line, compared with $> 100 \mu\text{M}$ for an insensitive cell line.

The biotinylated ligands were complexed to a streptavidin-activated affinity column and exposed to cell lysates from the susceptible cell lines. Bound proteins were eluted from the column and separated using SDS-PAGE. Proteins were characterised by MALDI MS and MS/MS and identified using Mascot database searches. Heterogeneous nuclear ribonuclear protein A2/B1 was found to selectively bind to the affinity probes.

TO MY FAMILY

ACKNOWLEDGMENTS

I would firstly like to thank Yorkshire Cancer Research for funding this project, also BMSS, School of Life Sciences and the Frank Hudson Memorial Fund for travel bursaries.

My heartfelt thanks go to all of my supervisors, Dr. Richard Wheelhouse, Dr. Chris Sutton and Dr. Roger Phillips for all their help and support throughout the last three years. I would also like to thank Dr. Nazira Karodia, Dr. Dawen Rong, Dr. Elrashid Elobaid Garelnabi, Tariq Mahmood and Dr. Nina Sahabo for their daily support and discussions in the lab.

For all their help and guidance in the proteomics laboratory I would like to thank Hannah Bateson, Kelly McMahon, Saira Saleem and Nitin Rustogi.

Finally I would like to say a big thank to all my family and friends for their endless support and encouragement throughout this project. Also a special thanks to my friends in D company Yorkshire (N & W) Army Cadet Force for the fantastic weekends away, showing me that there is much more to life than my work.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Introduction.....	1
1.2 The drug discovery process.....	3
1.2.1 The National Cancer Institute screening programme.....	4
1.2.1.1 COMPARE programme.....	8
1.2.1.2 Self-organised maps.....	10
1.2.2 Phortress: an example of compound-driven drug discovery....	15
1.2.3 Target orientated drug discovery.....	20
1.2.3.1 Target identification.....	20
1.2.3.2 Target validation.....	22
1.2.3.3 Hit compound identification.....	23
1.2.3.4 Lead compound optimisation.....	24
1.3 Biarylheterocyclic hit compounds.....	27
1.4 Aplidine.....	33
1.5 Target deconvolution.....	36
1.6 Proteomics.....	37
1.6.1 Mass spectrometry.....	39
1.6.1.1 Matrix-assisted laser desorption ionisation.....	40
1.6.1.2 Time of flight mass analysers.....	42
1.6.1.3 Tandem mass spectrometry.....	46
1.6.1.4 Instrumentation.....	46
1.6.2 Protein identification.....	48

1.6.3 Chemical proteomics.....	52
1.7 Conclusion.....	55
1.8 Aims.....	56

CHAPTER 2: DESIGN OF THE AFFINITY PROBES

2.1 Introduction.....	57
2.2 Chromatography medium.....	58
2.2.1 Biotin/streptavidin bond.....	59
2.3 Bespoke matrices for affinity chromatography.....	62
2.4 Attachment linker.....	66
2.5 Proteomic Strategy.....	68

CHAPTER 3: SYNTHESIS

3.1 Introduction.....	70
3.2 Established synthesis.....	74
3.2.1 Stage 1 – Formation of the biarylheterocyclic ring system.....	74
3.2.1.1 The Suzuki coupling reaction.....	74
3.2.1.1.1 Variants on the Suzuki coupling reaction.....	81
3.2.1.1.2 Suzuki reactions carried out in this project.....	82
3.2.2 Stage 2 – Hydrolysis of the methyl esters.....	84
3.2.2 Stage 3 – Formation of amide bonds.....	84
3.3 Synthesis of non-symmetrical compounds.....	87

3.3.1	Synthesis of the attachment linker.....	87
3.3.2	Route A – Breaking symmetry during amide formation.....	88
3.3.3	Route B – Breaking symmetry during ester hydrolysis.....	91
3.3.4	Route C and Route D – Breaking symmetry during formation of the biarylheterocyclic ring.....	93
3.3.4.1	Sequential Suzuki coupling.....	94
3.3.4.2	Route C.....	100
3.3.4.3	Route D.....	101
3.3.4.4	Acylation with imidazole side chains.....	102
3.3.5	Conclusions.....	104
3.4	Synthesis on non-symmetrical biotinylated compounds	106
3.4.1	Conclusion.....	108
3.5	Attachment of a linker to the pyrimidine ring.....	108
3.5.1	Formation of the biarylheterocyclic ring system.....	109
3.5.2	Attachment of the linker.....	110
3.5.3	Conclusion.....	116
3.6	Conclusion.....	118

CHAPTER 4: BIOLOGICAL EVALUATION AND PROTEOMICS

4.1	Introduction.....	119
4.2	<i>In vitro</i> chemosensitivity.....	120
4.3	Proteomics.....	122
4.3.1	Method 1.....	124
4.3.2	Method 2.....	129

4.3.3 Method 3.....	130
4.3.4 Method 4.....	138
4.3.2 Heterogeneous nuclear ribonuclear protein A2/B1.....	153
4.3.5 Conclusions.....	155

CHAPTER 5: CONCLUSIONS AND FURTHER WORK

5.1 Conclusions.....	156
5.2 Target validation.....	158
5.3 Synthesis of alternative affinity probes.....	159

EXPERIMENTAL

Experimental details for Chapter 3: Synthesis.....	162
Experimental details for Chapter 4: Biological evaluation and proteomics....	200

REFERENCES	210
-------------------	------------

APPENDIX

Appendix 1: NMR data for compound 92	252
Appendix 2: Mass spectrometry data for peptides identified as hnRNP A2/B1.....	257
Appendix 3: Publications.....	263

ABBREVIATIONS

Δ	Reflux
ΔT_m	Change in melting temperature
$^{\circ}\text{C}$	Degrees Celcius
2-D	Two Dimensional
2DE	Two Dimensional Electrophoresis
3-D	Three Dimensional
A	Absorbance
A	Alanine
ABPP	Activity-Based Probe Profiling
Ac	Acetyl
ACTH (clip)	Adrenocorticotrop Hormone (Corticotrophin-Like-Intermediate Lobe Peptide)
ADME	Absorption, Distribution, Metabolism and Elimination
Ambic	Ammonium bicarbonate
APS	Ammonium persulfate
Ar	Aromatic
ATP	Adenosine Triphosphate
ATR	Attenuated Total Reflection
BLAST	<i>Basic Local Alignment Search Tool</i>
BOC	<i>Tert</i> -butyloxycarbonyl
br	Broad
BRCA1	Breast Cancer 1

Bu	Butyl
C	Cysteine
C7BzO	3-(4-Heptyl)phenyl-3-hydroxypropyl)dimethylammonio)propanesulfonate
CCCP	Compound Centric Chemical Proteomics
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHCA	α -Cyano-4-hydroxycinnamic acid
CID	Collision Induced Dissociation
ClogP	Log of <i>n</i> -Octanol/Water Partition Coefficient
CNS	Central Nervous System
COSY	Correlated Spectroscopy
CT	Calf Thymus
CYP	Cytochrome P450
d	Doublet
D	Aspartic Acid
DCM	Dichloromethane
DDG	Drug Development Group
DDQ	Dichlorodicyanoquinone
DEPT	Distortionless Enhancement by Polarisation Transfer
DMF	Dimethyl Formamide
DMSO	Dimethyl Sulfoxide
DMW	Database Molecular Weight
DNA	Deoxyribonucleic acid
DTP	Developmental Therapeutics Programme
DTT	Dithiothreitol

E	Glutamic Acid
ECACC	European Collection of Cell Cultures
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal Growth Factor Receptor
eq.	Equivalent
ER+	Oestrogen Receptor positive
ESI	Electrospray Ionisation
Et	Ethyl
EtOH	Ethanol
exMOA	Expanded mechanism of action
ExPASy	Expert Protein Analysis System
F	Phenylalanine
FBS	Fetal Bovine Serum
FDA	Federal Drug Administration
FT-IR	Fourier Infrared
FMW	Fragment Molecular Weight
G	Glycine
GC	Gas Chromatography
GI	Gastrointestinal
GI ₅₀	Growth Inhibition; the concentration of drug required to inhibit growth in 50% of cells
GSH	Glutathione
h	Hours
HBSS	Hank's Buffered Salt Solution
HER-2	Human epidermal growth factor receptor

HMBC	Heteronuclear Multiple Bond Correlation
hnRNP	Heterogeneous nuclear ribonuclear protein
HOBt	<i>N</i> -Hydroxybenzotriazole
HPLC	High pressure liquid chromatography
HSP90	Heat Shock Protein 90
HSQC	<i>Heteronuclear Single Quantum Coherence</i>
HTS	High Throughput Screen
I	Isoleucine
IC ₅₀	the concentration of drug required to kill 50% of cells
IPA	Isopropyl alcohol / propan-2-ol
ISD	In-Source Decay
<i>J</i>	Coupling constant
JNK	c-Jun N-terminal kinase
K	Lysine
K _a	Acid dissociation constant
L	Leucine
LC	Liquid Chromatography
LC ₅₀	Lethal Concentration; the concentration of drug required to kill 50% of cells.
Lit	Literature
<i>L-Pro</i>	L-Proline
Ltd	Limited
m	Multiplet
M	Molar
M	Methionine

m.p.	Melting Point
m/z	Mass:Charge ratio
MALDI	Matrix-Assisted Laser Desorption Ionisation
MAP	Mitogen Activated Protein
MAPK	Mitogen Activated Protein Kinase
MD	Maryland
MDM2	Murine Double Minute
Me	Methyl
MeCN	Acetonitrile
MeOH	Methanol
min	Minutes
MIND	Mining Information for New Direction
MKP-1	MAP Kinase Phosphate-1
MMC	Mitomycin C
mmol	Millimole
MOA	Mechanism of Action
MOWSE	Molecular Weight Search Engine
mRNA	Messenger Ribonucleic Acid
MS	Mass Spectroscopy
MS/MS	Tandem Mass Spectrometry
MTP	Massive Target Plate
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Mwt	Molecular Weight
N	Asparagine
NAP	Nucleic Acid Purification

NCI	National Cancer Institute
NMR	Nuclear Magnetic Resonance
NSC	Non-Small Cell
OAT	Ornithine aminotransferase
ODC	Ornithine Decarboxylase
p	Probability
P	Proline
PBS	Phosphate Buffered Saline
PCC	Pearsons Correlation Coefficient
PCIS	Precursor Ion Selector
PCR	Polymerase Chain Reaction
Ph	Phenyl
PhMe	Toluene
PIC	Protease Inhibitor Cocktail
PKC	Protein Kinase C
PLMS	Post-LIFT Metastable Suppressor
PMF	Peptide Mass Fingerprint
ppm	Parts per million
PSD	Post Source Decay
PyBOP	Benzotriazol-1-yl-oxytrispyrrolidino phosphonium hexafluorophosphate
q	Quartet
Q	Glutamine
quint	Quintet
R	Arginine

RAND	Rapid Access to NCI Discovery Resources
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
RPMI	Roswell Park Memorial Institute
RT	Room Temperature
s	Singlet
S	Serine
SAR	Structure-Activity Relationship
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
siRNA	Small Interfering Ribonucleic Acid
SOM	Self-Organised Map
T	Threonine
TEMED	Tetramethylethylenediamine
TGI	Total growth inhibition; the concentration of drug required to inhibit 100% of cell growth
THF	Tetrahydrofuran
TLC	Thin layer chromatography
T_m	Melting Temperature
TOF	Time of Flight
Tris	Trishydroxymethylaminomethane
UV	Ultraviolet
V	Valine
VEGF	Vascular endothelial growth factor
Xphos	2-Dicyclohexylphosphino-2',4',6'-triisopropylbiphenyl
Y	Tyrosine

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Cancer is a leading cause of death worldwide and in the UK the lifetime risk of developing cancer is more than one in three. In 2008, cancer was the cause of 7.6 million deaths and it is estimated that globally there were 12.4 million newly-diagnosed cancer cases. Considering the continued growth and age of the world's population, cancer mortality is predicted to keep rising, with an estimated 26.4 million incidences and 17.0 million cancer deaths by 2030.¹ In conjunction with reducing incidence and early detection, the design and development of new and superior treatments for cancers is essential.

Cancer treatment can take numerous approaches including surgery, radiotherapy, chemotherapy, immunotherapy and endocrine therapies. Surgery is the oldest and most frequently used treatment method and one of the few curative therapies. However, this type of treatment is only suitable for localised, solid tumours. Similarly radiation therapy can be used in place of surgery, as a less invasive approach, for localised cancers. However, both methods fail if cancer metastasises and spreads to secondary sites within the body.

Cancer is characterised by increased rates of cellular proliferation and many current chemotherapeutic strategies target this. Conventional chemotherapeutic treatments have been successful since they were first implemented in the 1940s, after toxicological studies of mustard nerve gas, and continue to extend many lives today. Nowadays chemotherapy refers to the use of cytotoxic drugs

to treat a cellular component, such as DNA, RNA or proteins, required for proliferation resulting in cell death.²

Despite advances in the treatment of some cancers, the majority of cancers do not respond well to current chemotherapy, therefore treatments can be described as palliative rather than curative. As conventional chemotherapeutic agents are not specific to neoplastic cells, the treatments can lead to toxic side effects, such as bone marrow toxicity, neurotoxicity and carcinogenicity. Furthermore, this usually limits the dose of drug that can be administered. Also chemotherapeutic treatments may prove to be effective for defined periods of time; however, prolonged use can lead to acquired drug resistance.

The problems with drug resistance and toxic side effects present significant limitations for the treatment of cancers with classic chemotherapeutic agents. Hence, a priority of current research is developing new and improved drugs which are specific to cancers and selective for antitumour targets. These recent therapies are often referred to as 'molecular cancer therapeutics' and attempt to target the precise molecular mechanisms responsible for cancer initiation and progression.³

The first successful, small, synthetic molecular-targeted, cancer therapeutic was Glivec (Imatinib) and in many ways represents a prototype for drugs targeting oncogenic signal transduction pathways.⁴ In addition to small molecules, monoclonal antibodies (e.g. Avastin and Herceptin) can be used as oncogenic therapies.

1.2 THE DRUG DISCOVERY PROCESS

Potential new anticancer agents originate from a number of sources, ranging from random screening of natural products and synthetic chemicals to rational drug discovery.⁵ Synthetic organic chemistry is considered the underpinning field⁶ and has historically driven the drug discovery process. However as the biological understanding of cancer progresses, the successful discovery and development of small-molecule anticancer drugs is becoming multidisciplinary and highly dependent on the creative interaction between the different areas.⁷

Classical drug discovery (Figure 1.1) involves continual cycles of synthetic chemistry and biological evaluation, allowing refinement and optimisation of a target compound.⁷ Once optimised, the compound can enter preclinical *in vivo* evaluation generating efficacy and toxicological data. Before a compound can enter clinical trials, formulation is required.

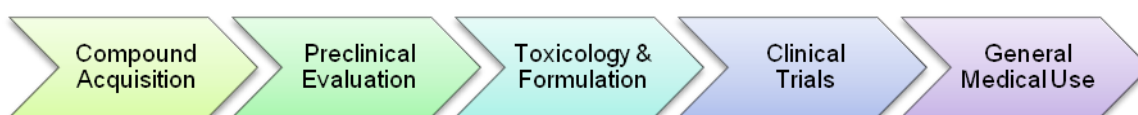


Figure 1.1. The Drug Discovery Process

Traditionally drug discovery was based on random screening of natural products, substrate analogues (antimetabolites) and small synthetic compounds *in vitro* and *in vivo* in the search for agents.⁸ Nowadays there are two main ways in which drugs are discovered, random drug screening and target

orientated drug discovery. A major programme runs at the NCI which screens many compounds arising from industry and academia allowing drugs to be discovered *via* an empirical drug discovery approach.

1.2.1 THE NATIONAL CANCER INSTITUTE SCREENING PROGRAMME

The National Cancer Institute (NCI), Bethesda, MD, USA, has been involved with drug discovery and development since its establishment in 1955. This was when it started a screening service for the large number of compounds arising from academia and industry⁹ – The Developmental Therapeutics Programme (DTP). The DTP allows selection of possible drug compounds as clinical candidates.¹⁰

Initial screening models used by the NCI were based on three transplantable rodent tumour models *in vivo* (sarcoma 180, carcinoma 755 and leukemia L1210). In 1975, the scheme was changed and the NCI tested all potential anticancer agents against a P388 or L1210 murine leukaemia pre-screen before testing on a panel of tumours.¹⁰ The programme produced drugs that were effective in the treatment of many cancers, mostly childhood leukaemia. However, there was limited success in the treatment of the common adult solid tumours.¹¹ Also, several compounds that failed to show activity in pre-screens showed activity in the second panel of tumour models and in clinical trials. In addition, the classes of anticancer drugs which were active against the rodent tumours were mainly cytotoxic alkylating agents and other DNA damaging agents and novel structures had not been discovered for over 20 years.¹²

Consequently the NCI changed its approach to large scale screening. In April 1990 a novel programme began, part of the Rapid Access to NCI Discovery Resources (R·A·N·D) programme, which involves *in vitro* screening of potential anti-cancer drugs against numerous human tumour cell lines.¹³ This massive-scale drug screening initiative is known as the 'NCI-60 DTP Human Tumour Cell Line Screen'. These primary pre-screens are followed by *in vivo* evaluation by hollow fibre assays¹⁴ and human tumour xenografts,¹² if compounds prove to be potent and selective against a particular disease category.

The screen has allowed vast number of compounds to be tested each year for potential anticancer activity. During the 1980s and 1990s the NCI were screening 10,000 compounds per year, however, screening throughput has now been reduced to approximately 3,000 compounds per year due to reductions in workforce.¹⁵

The primary screening programme involves *in vitro* testing of potential anticancer agents against 60 well characterised human tumour cell lines, grouped into the nine major histological classes of malignancies (leukaemia, melanoma, lung, colon, prostate, breast, renal, ovarian and CNS cancer).¹⁶ It was intended that the screen would exploit unique features of each tumour type and therefore replace the need for the primary *in vivo* screening previously employed. However, the patterns of drug sensitivity and resistance generated with the NCI 60 screen were found to reflect the mechanisms of drug action with standard anticancer agents, therefore providing an unexpected dimension to the

screening model.¹⁷ This has produced a number of discoveries which have greatly contributed to the field of targeted anticancer therapy.¹⁵

Data derived from the *in vitro* screening assay can be displayed and analysed in different ways. Dose response curves can be generated expressing percentage cell survival at varying concentrations of drug, relative to a drug-free control. From these curves LC₅₀ values (the concentration of drug required to kill 50 % of the original cell population), GI₅₀ values (the concentration of drug required to inhibit growth in 50 % of the cells) and TGI (the concentration of drug required to inhibit 100 % of cell growth) can be calculated¹⁵ (Figure 1.2).

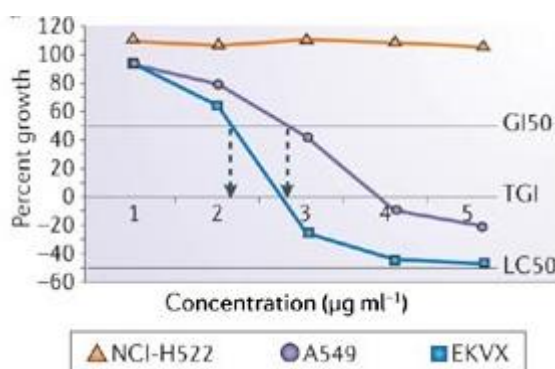


Figure 1.2. Dose-response curves showing the activity of a compound on three different NSC lung cancer cell lines. Linear interpolation gives the GI₅₀, TGI and LC₅₀. Adapted from Shoemaker 2006.¹⁵

The data collected can also be expressed in the form of a 'mean graph'. This is derived from a typical horizontal bar graph of the logarithms of the GI₅₀ (or TGI, or LC₅₀) values. In a mean graph each bar is plotted along a centred vertical line which is representative of the mean value of the log GI₅₀ for all cell lines. For

each cell line, the difference between the individual measured value and the mean, 'delta', is plotted¹⁸ (Figure 1.3).

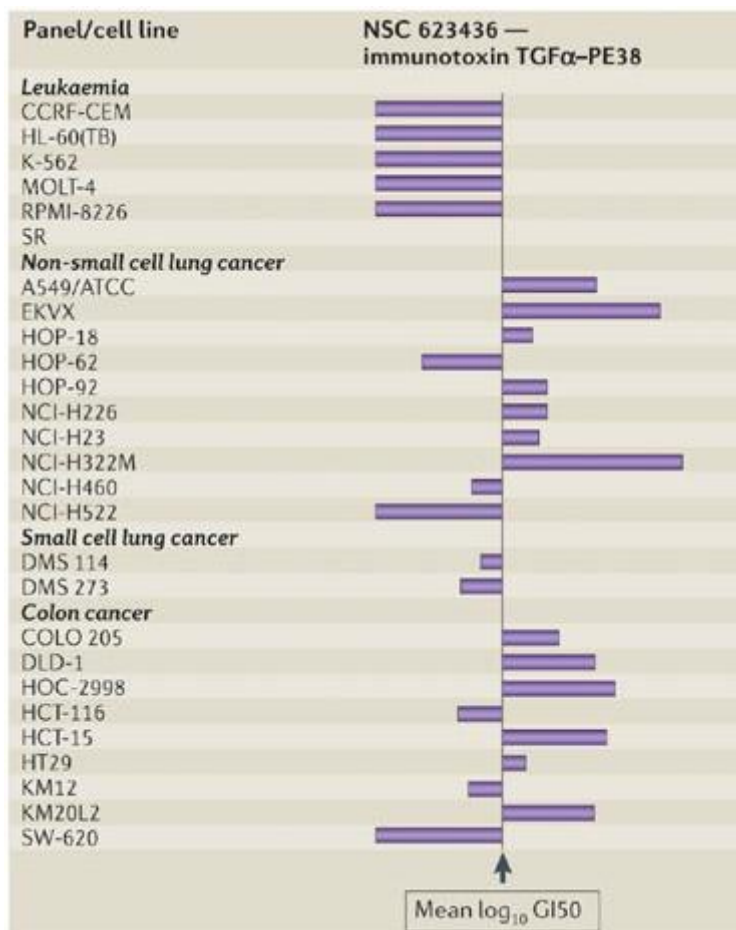


Figure 1.3. An example of a DTP mean graph showing GI₅₀ data for three types of malignancy treated with the immunotoxin transforming growth factor α . Deltas projecting to the right have a GI₅₀ greater than the mean. Adapted from Shoemaker 2006.¹⁵

The mean graphs have deltas projecting to either the right or left of the mean depending on the sensitivity of a particular cell line to an experimental drug compared with the mean.¹⁹ The mean graphs are a compact method of representing the test data and afford a profile for each compound which emphasises the differential effects and provides a unique data set – a

'Chemosensitivity fingerprint'. Comparison of the mean graphs generated provides an insight into the mechanism of action (MOA) of a drug. An automated method of comparison with large screening databases is also provided by the NCI – the 'COMPARE' programme.

1.2.1.1 COMPARE Programme

COMPARE is a computerised, pattern-recognition algorithm which evaluates data generated by the NCI 60 screen.¹⁸ The programme calculates the linear correlation coefficient between the data of an investigational seed compound over the 60 cell lines against all the profiles of previously tested compounds within a certain database.²⁰ Similarities between compounds are expressed as a quantitative figure – the Pearson correlation coefficient (PCC). Each COMPARE analysis affords a rank-ordered list of compounds. Compounds with comparable mean graphs, work by the same or similar mechanisms of action¹⁶,¹⁹ Therefore, compounds with low PCCs suggest novel MOA. The NCI databases available include:

- The synthetic database, which is subdivided into the standard agent database and the 'diversity set'. The standard agents database contains 175 FDA-approved anticancer drugs.²⁰ 122 of these compounds have well-established mechanisms of action²¹ and are present within the Anti-cancer Agent Mechanism Database. The diversity set contains 1990 compounds available from the NCI.²⁰

- The molecular characterisation database, which is subdivided into the molecular targets database and the NCI 60 microarray panel. The latter contains DNA chip measurements for all of the cell lines.^{22, 23} All of the 60 cell lines have been characterised for targets and potential targets, including protein levels, RNA measurements, mutation status and enzyme activity levels. The molecular targets database contains approximately 200 target levels and activity measurements.
- The natural product extracts database, which contains data for approximately 15,000 natural products.²⁰
- All compounds previously tested at the NCI stored in a publicly available database.

COMPARE is an information-intensive approach to drug discovery in cancer which can indicate novel chemical structural classes of compounds that exhibit a biochemical mechanism of action, in addition to compounds with novel mechanisms of action. The natural products database also allows crude extracts to be analysed for the identity of compounds possessing anticancer activity.¹⁸

Compounds are taken forward a stage in development for *in vivo* testing if they present a preferential activity against a certain type of cancer. Also if they exhibit a unique mean-graph fingerprint, indicating a novel mechanism of action, and have correlation with specific molecular target expression.

1.2.1.2 Self-Organised Maps

The extensive screening programme implemented by the NCI has generated great volumes of data relating to established anticancer drugs, diverse chemical compounds with unknown biological activity and molecular targets. Systematic investigations of the patterns of mean graph results, create a possibility of discovering new anticancer drugs and improving molecular taxonomy.²⁴ Computer algorithms have been developed to analyse these data patterns of $\log(GI_{50})$ measurements against the various tumour cell lines received from the 60-cell line screens.

The basis of this analysis relays the pairwise statistical correlations between the growth inhibition measurements for different drugs tested against the cell lines and establishes the similarities in mechanism of action, modes of resistance and molecular structure for all of the compounds tested.²⁵ A method of determining the biological significance in cellular response, between the different cell lines, can lead to identification of tumour-selective agents, new molecular targets and new lead compounds.²⁴

The Data mining programme – 3D MIND (mining information for novel discoveries) – implemented by the NCI, provides tools for analysis of data which complements COMPARE results using clustering methods.¹⁵ The self-organised map (SOM) is a useful method for exploration of data *via* two-dimensional response maps.²⁴ The SOM utilises iterative minimisation techniques to group multidimensional data into response vectors.²⁵ SOMs have been used in

various analyses of large datasets that exhibit large amounts of random noise²⁶ and missing data, making it suitable for mining the data from the NCI screen.

Although the raw data generated by the screening programme determine the potency of a compound for growth inhibition, the most useful data lie in the biological significance of the cellular response patterns. Therefore, raw data are normalised using Z-score conditioning to enhance the biological response signal.²⁴ Hence, improving the quality of data by providing a common mean reference and scale which leads to the enhancement.

The SOM method can be divided into two elements, clustering in high dimensional space and projections into a lower dimensional display space, thus providing a means of visual translation.²⁴ Visualisations of the clustering results, are 2D hexagonal projections containing 1066 clusters²⁵ (Figure 1.4). Compounds represented within each cluster have a unique pattern of growth inhibition.

Projection of the cellular profiles in the molecular targets database and microarray subpanels onto the complete DTP SOM allows the clusters to be divided into regions corresponding to the apparent cellular activity. Regions have been defined on the map that group individual map nodes with the most similar response profiles and nine functional clusters have been defined. Namely, mitotic (M), nucleic acid metabolism (S), membrane function (N), metabolic stress and cell survival (Q) and phosphatases/kinases and oxidative stress (P). The four remaining clusters are uncharacterised and have been named R, F, J and V.²⁷

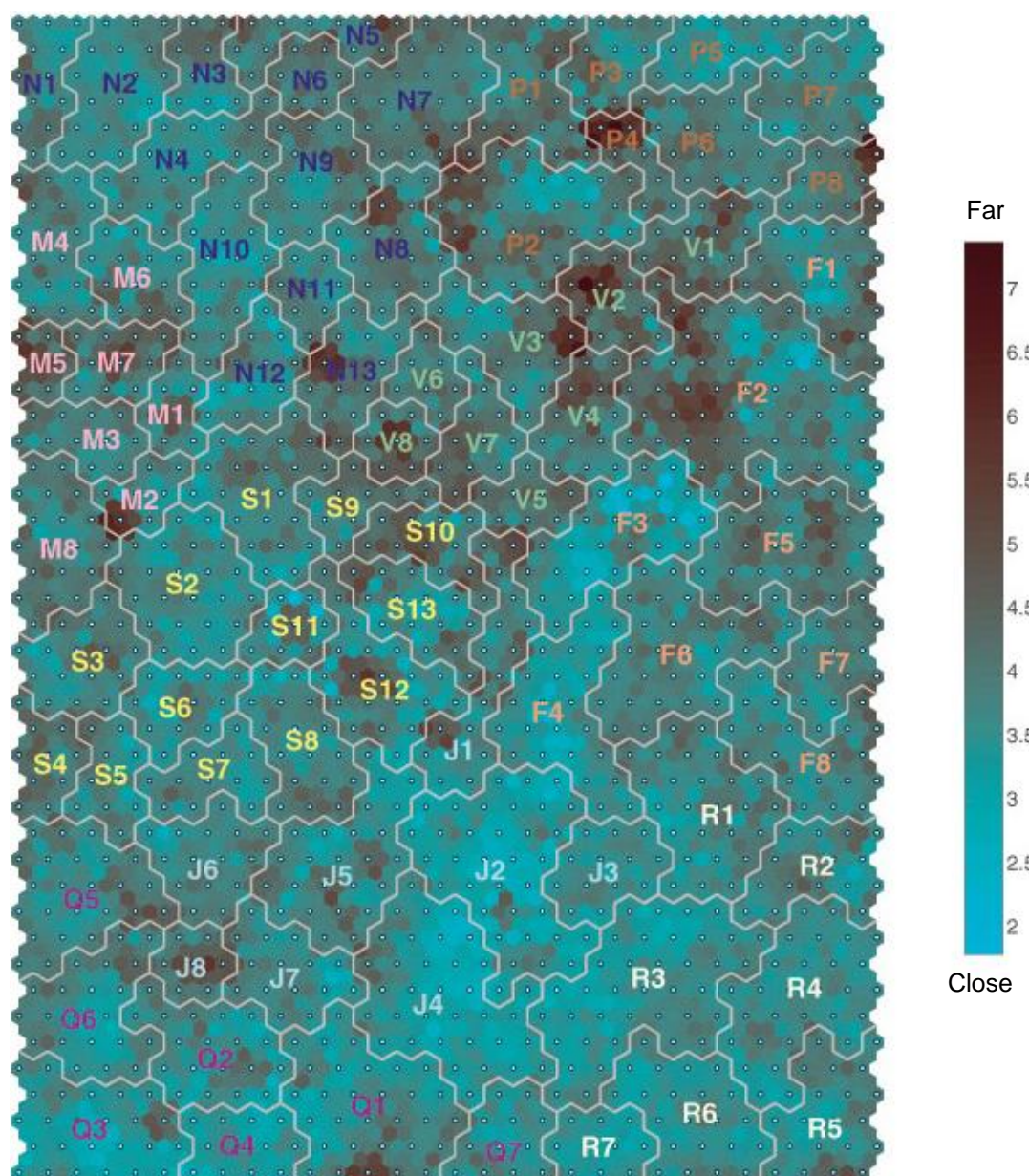


Figure 1.4. Complete DTP SOM map for approx. 20,000 compounds (Feb 2003 version). Colour is used to represent the distance between clusters.²⁸

In addition to the complete DTP SOM, the NCI 3D MIND programme also provides a bit vector map. This SOM relays data from a group of chemical descriptors for approximately 20,000 small compounds which are also used in the complete DTP map. The vectors represent the chemical properties of a

given molecule, such as hydrogen bond donors, hydrogen bond acceptors, and aromatic rings.

The data from the NCI 60 screen for each cluster can be viewed from the SOM cluster graphs as a data vector (Figure 1.5). This is a graphical representation of the variation in GI₅₀ data across the cell lines.

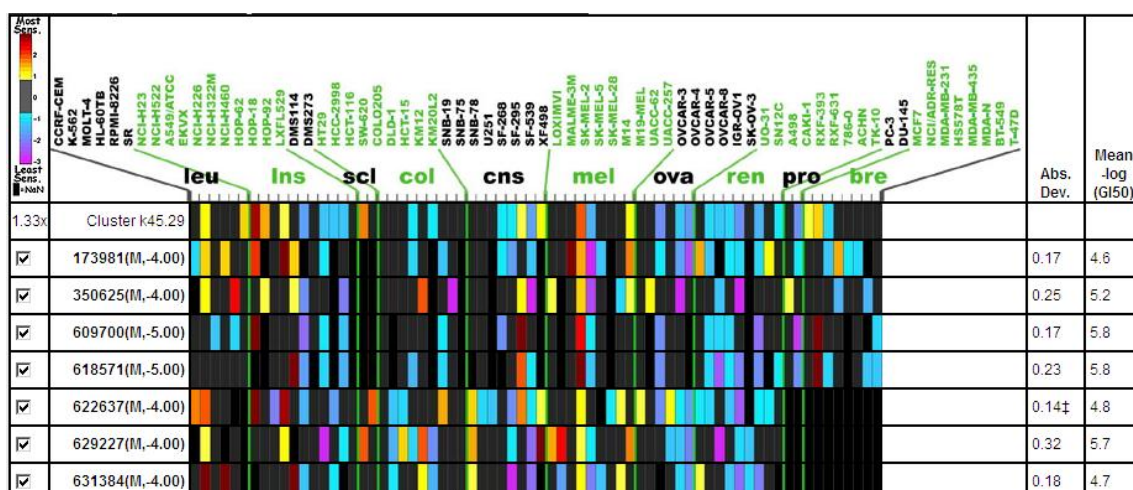


Figure 1.5. A data vector for a cluster of the complete DTP map. The x-axis consists of the different cell lines and the y-axis the compounds present within the vector.²⁸

The NCI 3D MIND programme also provides a mechanism of action map which is a small subset of the complete DTP map (Figure 1.6). The data expressed within the map are a set of 362 compounds identified as having one of six classes of molecular action: antimetabolic agents, RNA/DNA antimetabolites, DNA antimetabolites, topoisomerase I inhibitors, topoisomerase II inhibitors, and alkylating agents.

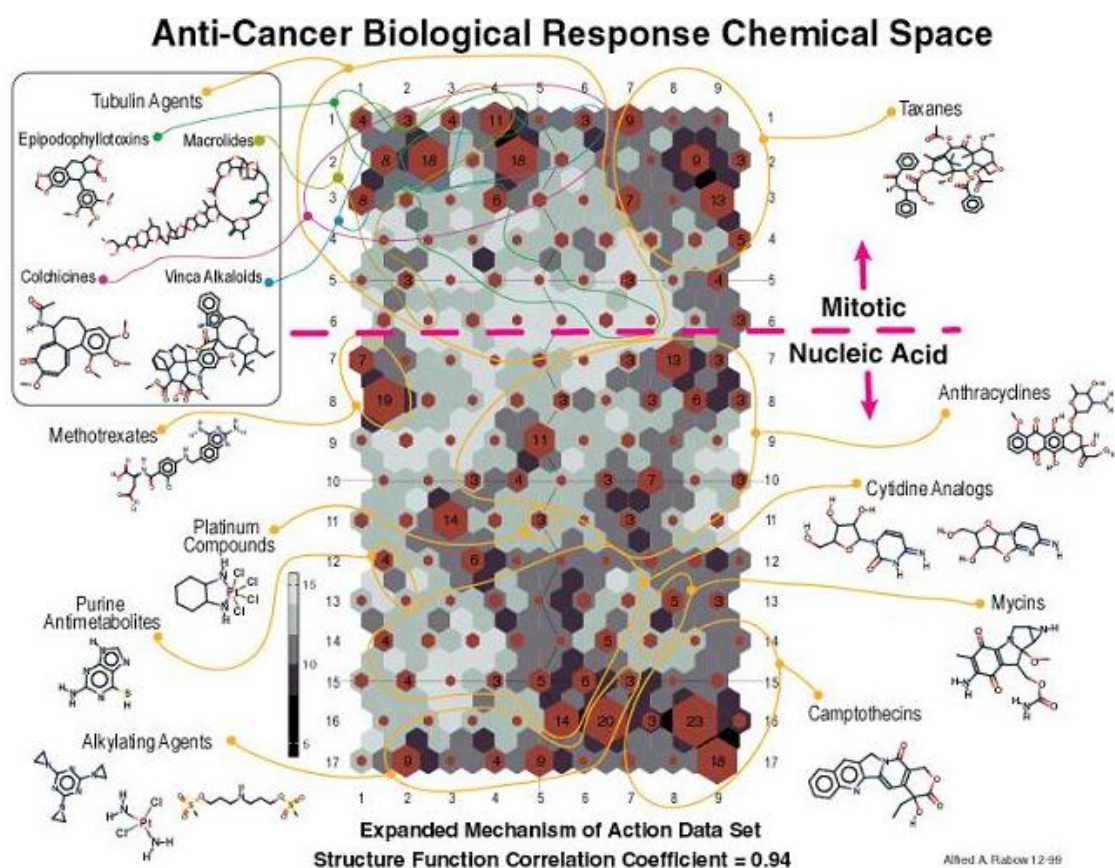


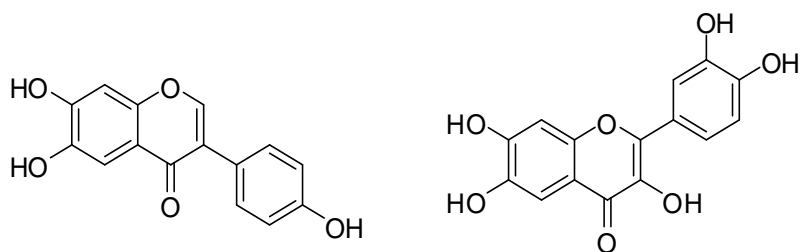
Figure 1.6. A SOM cluster map showing the expanded mechanism of action data set.²⁸

The tools provided by the NCI DTP programme have been responsible for the discovery and development of a number of anticancer agents including Phortress, Taxol and Halichondrin B.

1.2.2 PHORTRESS: AN EXAMPLE OF COMPOUND-DRIVEN DRUG DISCOVERY

Phortress is a novel antitumour agent currently in Phase I clinical trials in the UK. It has a unique mechanism of action relating to the activation of cytochrome P450 1A1 (CYP1A1).

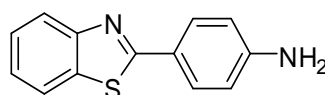
The development of Phortress has taken many years and began with the synthesis of polyhydroxylated-2-phenyl benzothiazoles related to genistein and quercetin.²⁹



Genistein

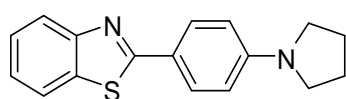
Quercetin

Biological evaluation of benzothiazole CJM 126 revealed selectivity towards MCF-7 ER+ breast carcinoma cells ($GI_{50} < 1$ mM). It was also noted that growth inhibition was associated with an unusual biphasic dose response.²⁹

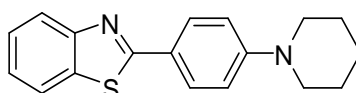


CJM 126

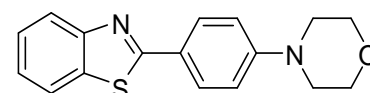
Structure-activity relationship studies found CJM 126 has a 10-fold increased activity compared with the benzoxazole and a 1000-fold increased activity compared with the benzimidazole analogue. The production of tertiary amine analogues pyrrolidin-1-yl, piperidin-1-yl or morpholin-4-yl groups showed increased growth inhibition ($GI_{50} = 0.1\mu\text{M} - 1\mu\text{M}$).³⁰



Pyrrolidin-1-yl analogue



Piperidin-1-yl analogue



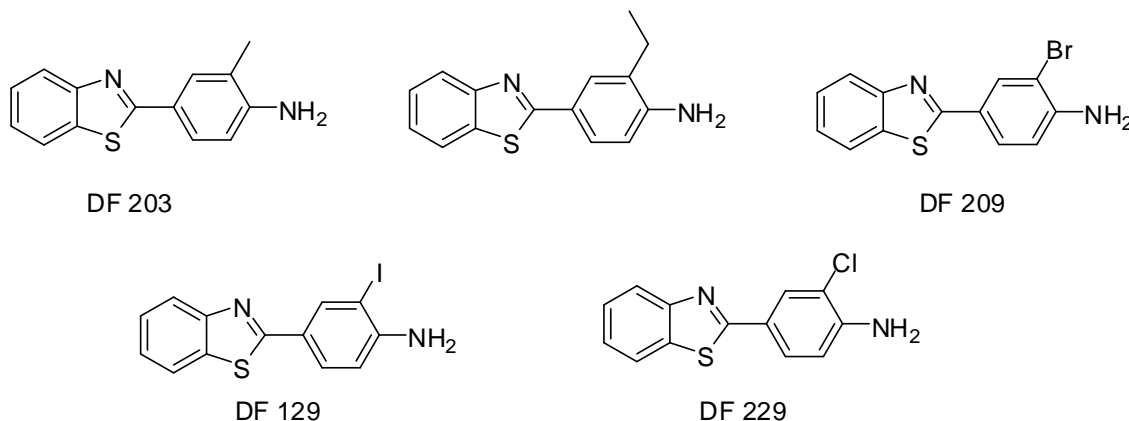
Morpholin-1-yl analogue

Alkoxy or hydroxy groups substituted into the benzothiazole nucleus led to reduced biological responses. Also, replacement of the 2-(4-aminophenyl) group of CJM 126 with a 2-(pyridin-4-yl) or 2-(2-aminopyridin-5-yl) group showed activity with $GI_{50} < 1\mu\text{M}$.

Introduction of a bromo group to yield 2-(2-amino-3-bromopyridin-5-yl) benzothiazole enhanced activity by more than 100-fold compared with the unbrominated analogue. Replacement of the amino group by a nitro group in 3' substituted 2-(4-aminophenyl)benzothiazole analogues reduced cellular activity ($GI_{50} < 1\mu\text{M}$).

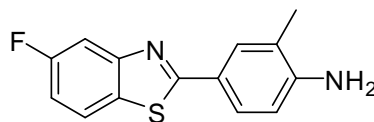
Introduction of a substituent to the 2' position of the aminophenyl group of CJM 126 showed an increase in the GI_{50} values. However, potent activity was demonstrated with 3'-methyl (DF 203), 3'-ethyl, 3'-bromo (DF 209), 3'-iodo (DF

129) and 3'-chloro (DF 229) substituted analogues ($GI_{50} < 0.1$ nM). Although 3'-cyano and 3'-hydroxy substituents showed a decrease in cellular response.



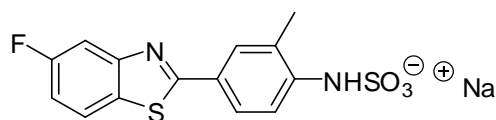
In 1994 the NCI DTP received the parent benzothiazole compound (DF 203) and its analogues.³¹ The NCI 60 tumour cell line screen demonstrated potent activity against breast, ovarian, renal, colon, melanoma and NSC lung cancer cell lines ($GI_{50} < 10$ nm).²⁹ The biphasic dose response previously seen was confirmed and the COMPARE programme demonstrated a unique fingerprint. The molecular target SMART screen confirmed the target to be of unknown function. Collectively these data suggested a novel MOA.

Fluorination of a compound prevents undesirable hydroxylation of bioactive substrates. Therefore, a series of mono- and di-fluoro analogues of the 3'-substituted 2-(4-aminophenyl)benzothiazole compounds was synthesised. All analogues retained their potency and selectivity. Moreover, only the 5- and 7-substituted compounds showed a conventional dose response, i.e. the biphasic response was eradicated.³⁰



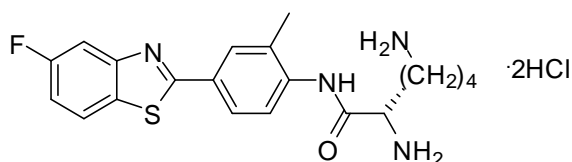
5F 203

In 1996 analogues of 5F 203 were presented to the Drug Development Group (DDG). However, the compounds were not accepted due to high lipophilicity and poor water solubility. Initially 2-(4-aminophenyl)benzothiazole sulfamate salts were synthesised but they were found to be sparingly soluble in aqueous medium and degraded rapidly to the free base under acidic conditions. Also they showed a decrease in *in vitro* activity.



5F 203 Sulfamic acid salt

Alanyl- and lysyl- amide hydrochloride salts showed higher degrees of water solubility and chemical stability. The prodrugs also rapidly reverted to their parent compounds *in vitro*.



5F 203 Lysyl- salt (Phortress)

Screening of the amino acid prodrugs in the NCI 60-cell line panel showed selectivity was retained with the same potency as the parent amine (Figure

1.7).³⁰ In 1999 5F 203 lysylamide dihydrochloride salt was submitted to the NCI DDG and was approved for development as a clinical candidate in 2002. The DDG provided synthesis, formulation, pharmacokinetic and toxicity resources.

In 2004, Pharminox acquired exclusive rights to develop and commercialise 5F 203 lysylamide dihydrochloride salt under the name of Phortress.³¹

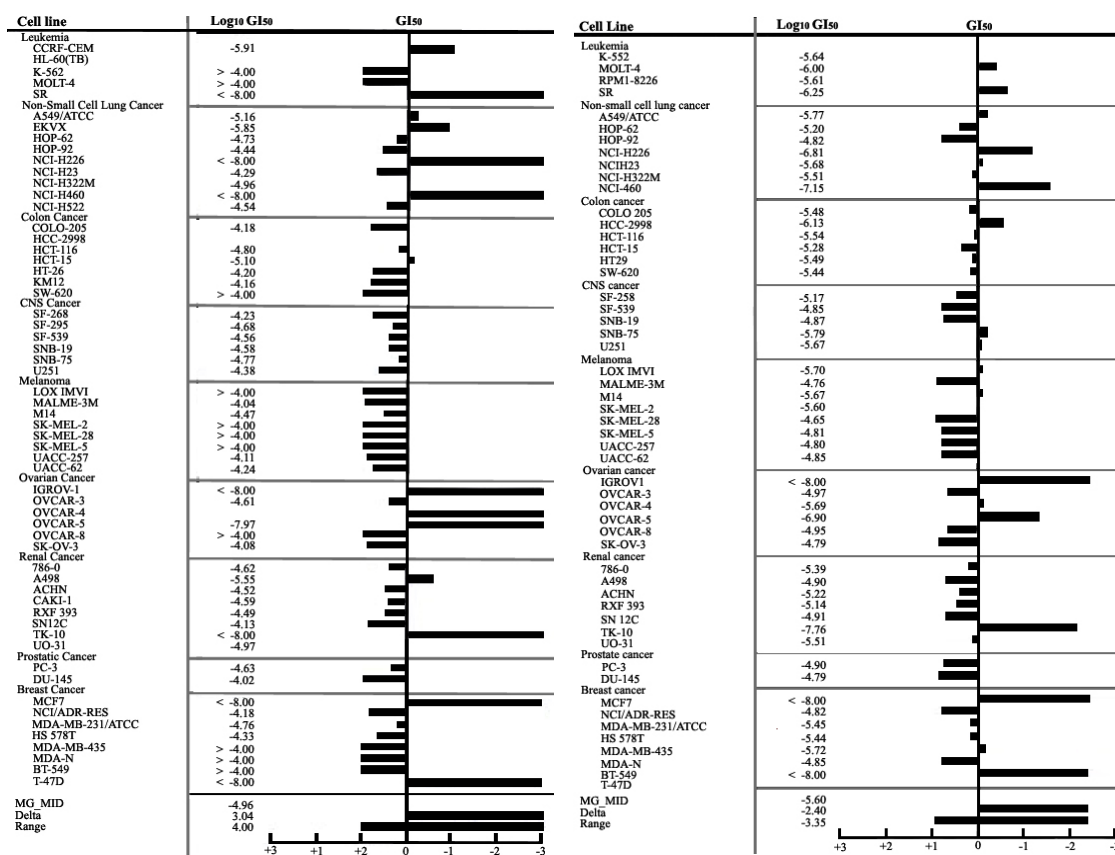


Figure 1.7. Comparison of the activities of 5F 203 (left) and its lysylamide dihydrochloride salt (Phortress, right).³⁰

The information generated by the NCI COMPARE programme has been effective in identifying Phortress has a novel mechanism of action, which has successfully advanced the drug to the clinic.

1.2.3 TARGET-ORIENTATED DRUG DISCOVERY

The discovery of anticancer drugs today is taking a more rational approach, leading to more selective agents that act on targets specific to cancer (Figure 1.8).



Figure 1.8. Molecular-Targeted Drug Discovery.

1.2.3.1 Target Identification

Understanding the molecular mechanisms underpinning cancer development and progression unveils many potential new drug targets. Human cancers are known to evolve through a multistage process which extends over several years.³² The carcinogenesis process is driven by the accumulation of mutations and epigenetic alterations in gene expression with important biochemical roles.^{32, 33} Hanahan and Weinberg outlined six essential alterations between normal and cancer cells that dictate malignant growth termed 'The Hallmarks of Cancer' (Figure 1.9).³⁴

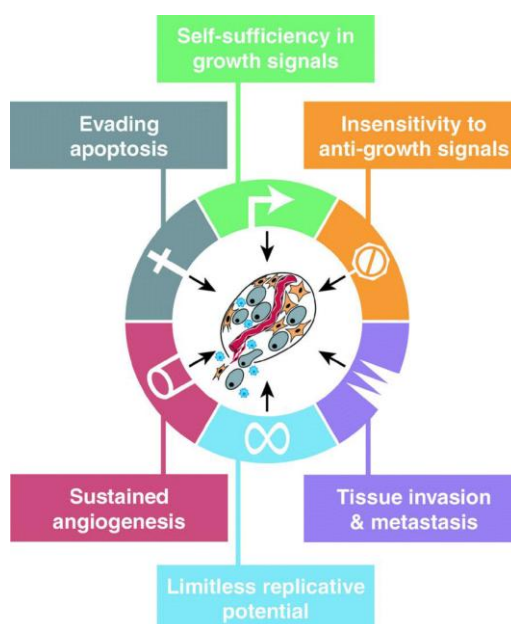


Figure 1.9. Acquired Capabilities of Cancer.³⁴

The selection of a target is based on numerous factors including its role in the initiation and progression of cancer and its 'druggability'.³ Alterations in cancer cells can include mutated genes, such as *p53*³⁵ or *B-raf*³⁶, or amplification and over-expression of specific cancer gene products, such as HER-2, MDM2, EGFR and cyclin dependent kinases.³⁷ Also, a normal gene product associated with cancer, such as telomerase which has been found at increased levels in 85 % of tumour cells tested.

A variety of different biotechnological techniques can be used to identify potential targets. Cancer gene discovery is predominantly carried out using genome-wide high-throughput systematic screening techniques which include gene copy number analysis, gene expression profiling and gene resequencing.³ RNA interference technologies³⁸ and proteomic techniques³⁹ are also becoming more widely used to identify potential targets for drug development.

1.2.3.2 Target Validation

Genomics and proteomics have identified thousands of potential targets.⁴⁰ However, the new mechanism-based approach to drug discovery has led to a decrease in the number of agents approved for cancer treatment, with only one compound being approved in 2008.⁸ This high failure rate can be attributed to incorrect target selection.⁴¹ For example, reversal of only a few of the large numbers of abnormalities present in the cancer genome has a profound effect on tumour growth. This is described as oncogenic addiction.⁴²

Discovery and development of a drug towards a particular target costs time and money. Once a target enters the pipeline of a pharmaceutical company, it can take approximately 12 years to develop at a cost of \$1 billion by the time it reaches the market.⁴³ Therefore, once a target has been identified it is essential to prove that the target is involved in neoplastic development and also that modulating the target can have a therapeutic effect, hence be a suitable target for development of a new drug.

A target may be validated, even so not all validated targets are deemed to be 'druggable'. The intrinsic pathway of apoptosis is blocked in many cancer cells, exposing many targets. Developing a drug towards these blocks appears attractive to induce programmed cell death. However, all the targets are 'undruggable' protein-protein interactions, resulting in the majority of chemotherapeutic agents for these targets being antisense compounds, rather than synthetic small molecules.⁴⁰

Target validation aims to prove that the knockdown of a target is associated with a phenotypic change.⁴⁴ Many molecular tools are available for target validation, including RNA interference, transgenic mice and chemogenomics and chemoproteomics using high throughput screening of compound libraries.⁴⁵

1.2.3.3 Hit Compound Identification

Following target validation the process of hit compound identification is initiated. Hit compounds are ideally 'drug-like', that is they have functional groups and similar properties to known drugs. Hit identification can be achieved by various methods. Large libraries of compounds, including small molecules and natural products, can be evaluated against targets in specific, high-throughput biochemical assays, such as the NCI's SMART screen, to allow a hit compound to be identified. This method has proved successful as it allows up to 10,000 compounds to be investigated daily.⁴⁶ Combinatorial chemistry has allowed the numbers of compounds in libraries to increase. Fragment-based lead discovery follows the same approach, but uses low molecular weight compounds, instead of drug-like molecules. This technique has increased the hit rate of lead discovery and also increased the ability to take hits successfully through optimisation into drug candidates.⁴⁷ Following screening of fragment libraries, combination of the fragments, can lead to a hit compound.⁴⁸ Alternatively hits can arise from literature compounds, where analogues can be selected on the basis of medicinal chemistry experience, intuition and trial and error.⁴⁹

Rational structure-based drug design can be employed to discover hit compounds. X-ray crystallography and NMR can provide detailed structural information to generate 3-D structures of targets, in addition to computational generation of models.⁵⁰ If the location of the binding pocket of a target is known, ligands can be generated to fit the target with the aid of molecular modelling programmes. Docking methods and structure based *in silico* HTS can be used, with the advantage of calculating the local binding energy.⁵¹

A hit compound is usually a small organic molecule with potent activity and selectivity. Good hit compounds are generally smaller, less lipophilic and contain less functionality than typical drugs.⁷ These features allow for modification during the lead optimisation stage in the drug discovery process. It has been shown that leads share many structural similarities to their corresponding drugs, therefore, hit identification along with lead optimisation are crucial steps in the drug discovery process.⁵²

1.2.3.4 Lead Compound Optimisation

Hit compounds that are potent in *in vitro* studies often need optimising to improve the pharmacokinetics, selectivity and solubility. In addition to reduce toxicity effects whilst retaining potency.⁵³ This stage of the drug discovery process is challenging and resource-intensive.

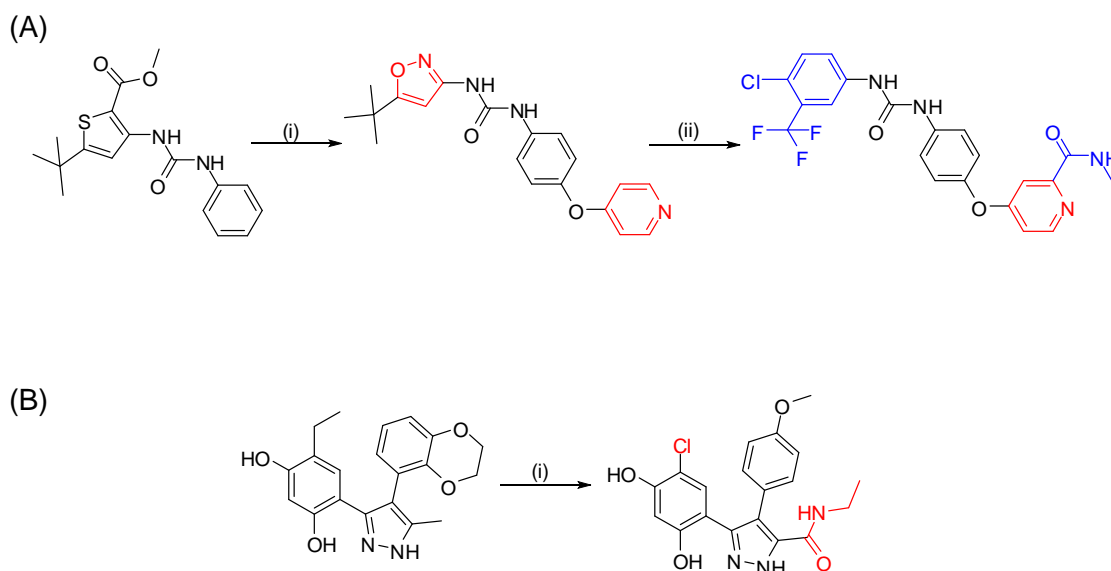
Compound optimisation can be pursued using a combination of experimental and computational studies. Computational studies using the 3-D ligand-protein

structure can accurately predict the binding affinity of the complex. Compound optimisation is carried out on the basis that increased affinity leads to increased potency.⁵⁴ Optimisation normally involves the synthesis of hundreds of analogues, however the combination of synthesis with informatics can lead to a significant reduction in time and cost.⁵³

Modifications in the structure of a lead, followed by *in vitro* screening, allow the generation of structure-activity relationship (SAR) information essential for developing a lead compound into a drug for clinical use.⁵⁵ Analysis of drugs and their corresponding leads has shown that the structural core of the lead remains and that drugs are usually more complex.⁵⁶ In addition leads are generally smaller, more polar, have higher solubility and are less flexible than drugs.⁵⁷ Scheme 1.1 outlines some of the modifications made to compounds progressing through the lead optimisation phase of drug discovery.⁷

A study calculating the variations between drugs and leads found that common changes were made to leads during the optimisation stage of drug development.⁵⁸

- Increase in molecular weight.
- Increase in lipophilicity.
- Relatively small increase in the number of hydrogen bond acceptors.
- Almost no change in the number of hydrogen bond donors.
- Almost no change in the number of rings within the structure.



Scheme 1.1. Modifications made to small molecule leads during optimisation into drugs. (A) Nexavar is a kinase inhibitor. The hit was discovered during HTS against CRAF protein kinase. (i) Combinatorial modification generated a more potent analogue ($IC_{50} = 0.23 \mu\text{M}$ compared with $17 \mu\text{M}$). (ii) Studies to increase *in vivo* activity lead to Nexavar. (B) VER49009 is a HSP90 inhibitor. The hit was identified using HTS against HSP90 ATPase and co-crystallised with the enzyme. (i) Structure based design was modified to give the potent inhibitor ($IC_{50} = 25 \text{ nM}$ compared with $9 \mu\text{M}$).⁷

Properties desirable for orally-available drugs have been outlined by Lipinski. His 'rule of fives' assesses physical and structural features of compounds to determine whether they could become orally active drugs.⁵⁹ These empirical rules predict information about the pharmacokinetics of a compound, specifically the Absorption, Distribution, Metabolism and Elimination (ADME), early in preclinical development.

Lipinski's rule considers the molecular weight of a compound, as evidence shows poorer intestinal permeability with larger molecular weight compounds.⁶⁰ It relies on the ClogP, which is the logarithm of the calculated

n-octanol/water partition coefficient, a measure of the compound's lipophilicity, and hence relates to its aqueous solubility, which affects the uptake and distribution of a drug.

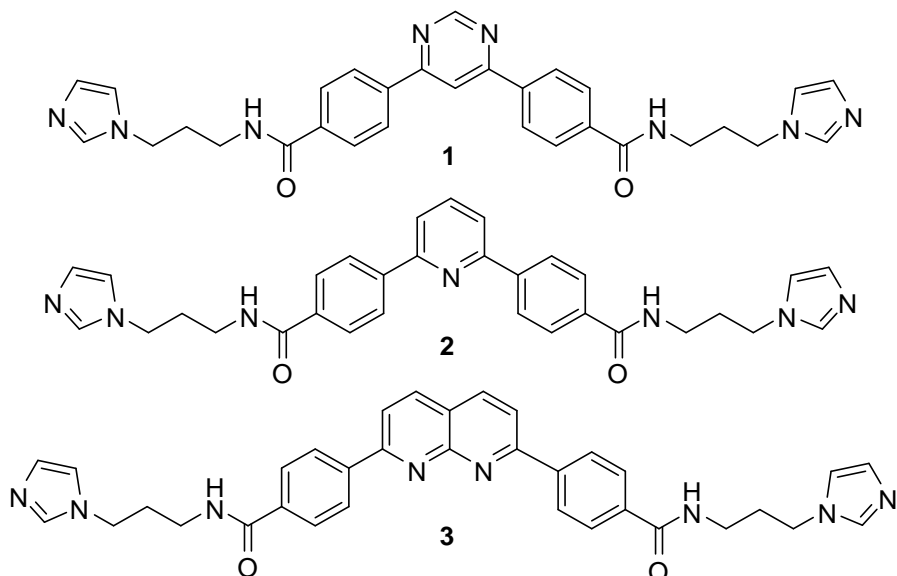
Finally the rule includes the number of hydrogen bond donors and acceptors. Excessive numbers of hydrogen bond donors can reduce permeability across a cell membrane.⁵⁹ Lipinski's rule of fives state that a drug should ideally have no more than 5 hydrogen bond donors and 10 hydrogen bond acceptors, ClogP \leq 5 and the compound should have a molecular mass < 500.

In addition to the rule of fives, the rule of threes has emerged for fragments during fragment based HTS (Section 1.2.3.3). This rule is based on Lipinski's rule and states for a fragment to be built into an orally-available drug, it must have a molecular weight < 300, ClogP \leq 3 and it should have no more than 3 hydrogen bond acceptors and three hydrogen bond donors.⁶¹

1.3 BIARYLHETEROCYCLIC HIT COMPOUNDS

A series of biarylheterocyclic compounds bearing tertiary and cyclic amino groups was discovered by Victoria Phillips.⁶² This work was designed to develop and evaluate molecular twist ligands as a novel strategy for structure-selective nucleic acid binding. Their DNA binding ability was assessed using thermal DNA 'melting' assays and compounds showed strong DNA binding affinity with ΔT_m ranging between 10-15 °C.⁶³ During this work three novel

compounds (**1**, **2**, **3**) bearing an imidazol-1-ylpropyl side chain were synthesised.⁶²



Unexpectedly, biophysical evaluation of these three novel compounds with DNA thermal melting assays revealed only modest duplex DNA binding ($\Delta T_m \leq 2.1$ °C). Further biological evaluation of the three compounds in cell-based chemosensitivity assays revealed potent activity against the drug-resistant (MMC and EO9) colon carcinoma cell line BE⁶⁴ (Table 1.1). It can therefore, be concluded that the observed potent activity is due to a different mechanism from DNA binding. Additional evaluation was carried out by the NCI.

The NCI 60-cell line screen confirmed activity with nM concentrations against some of the high priority tumour cell lines, non-small cell lung cancer and leukaemia sub-panels (Figure 1.10).

Table 1.1. DNA Binding and MTT Chemosensitivity Data⁶²

	$\Delta T_m / ^\circ\text{C}$ CT DNA	Chemosensitivity ^c / μM		
		A2780	BE	H460
1	2.1 ^a	0.079	0.644	15.2
2	0.6 ^b	0.014	0.064	2.18
3	1.0 ^b	0.010	0.016	1.43

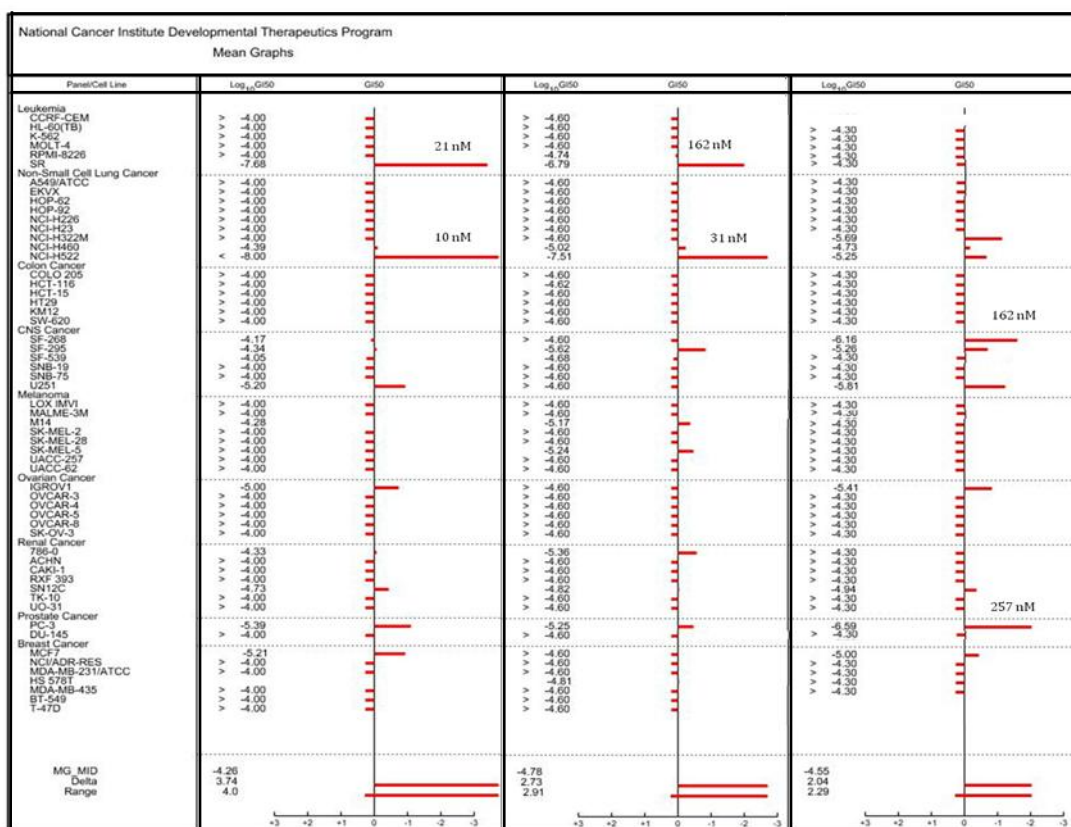
^a50 μMbp DNA, 10 μM ligand, 10 μM $\text{Na}_2\text{HPO}_4/\text{Na}_2\text{H}_2\text{PO}_4$, 1 mM Na_2EDTA , pH = 7.00 \pm 0.01.

^b50 μMbp DNA, 20 μM ligand. ^cIC₅₀ 96 h exposure.

Standard COMPARE analysis utilising the GI₅₀ data for each of the compounds as seed revealed some attractive properties of the hit compounds. These data demonstrated:

- no correlation with 175 FDA-approved anti-cancer drugs ($P \leq 0.365$);
- correlation with gene-like sequences of currently unknown function ($P \leq 0.77$)^{22, 23} in the molecular target and microarray panels;
- correlation with a complex marine natural product dehydrodidemnin (aplidine).

These data provide evidence that the hit compounds work by a novel mechanism of action targeting a gene-like sequence of unknown function. Isolation and subsequent identification of the target could offer an insight into the mechanism of action and scope for optimisation of the novel ligands.



1

2

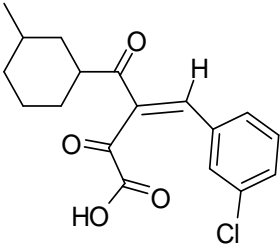
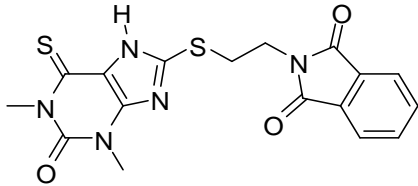
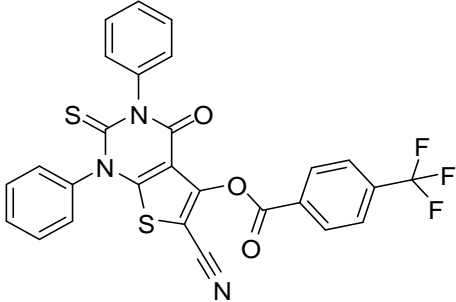
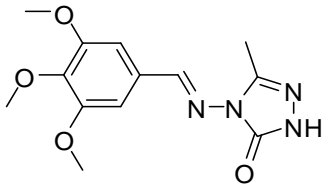
3

All Compounds ^a	1.0 ^b	0.748 ^d	0.498
Molecular Targets	0.77 ^c	0.615	0.524
Standard Agents	0.365	0.347	0.35

Figure 1.10. NCI GI₅₀ Mean Graphs and Results of COMPARE Analysis. ^aPearson correlation coefficient (P). ^b30 compounds, diverse structures. ^cGC11087, gene-like sequence of unknown function. ^dDehydrodidemnin (aplidine).

Further investigation into the NCI data for compound 1 was conducted by Dawen Rong,⁶⁵ who confirmed the correlation with apolidine, and unveiled new correlations (Table 1.2).

Table 1.2. Compare data for compound 1. Adapted from Rong 2009.⁶⁵

Database	NSC number	Correlation	Structure
Standard Agent (TGI)	325319	1.0	Aplidine
Mechanistic Set	636084	0.824	
Diversity Set	99671	0.825	
Synthetic Compounds	641873	0.932	
	637579	0.927	
	640586	0.913	
	634772	0.913	
	671133	0.878	
All Public Data	698068	0.803	

The compounds in the NCI databases showing correlation with compound **1** were analysed with the NCI 3D MIND tools. Each compound was mapped onto the complete SOM (Figure 1.4) to identify the functional cluster each of the compounds belong. Compound NSC 698068 was mapped in the F-region (unknown) on the 2003 version of the DTP complete SOM and the P-region (phosphatase- and kinase- mediated cell cycle regulation) in the earlier 1999 version. The similarities in structure suggest the hit compounds (**1**, **2**, **3**) may also cluster to this region due to a similar MOA, however, observation of the data vector for this compound suggested differences in chemosensitivity profiles. Compound 99671 was mapped to functional clusters J5 and J7. These cluster regions still have an unknown MOA. The data present within the cluster vectors showed similar chemosensitivity profiles.

Rong also mapped compounds found in each cluster onto the bit map to find structures with similar chemical properties. He synthesised several different analogues incorporating some of the structural features found within the 3D MIND data, for example replacing the amide with a sulphonamide, inverting the amide and using a non-planar conjugated ring system. Although the potency of the compounds was not improved, useful SAR data were achieved.⁶⁵

In addition to the compounds synthesised following investigation into the NCI 3D MIND and COMPARE research, Rong also synthesised compounds based around the common structure **4** (Figure 1.11). It was believed that the pharmacophore features of compounds **1**, **2** and **3** were still present in the

minimal structure **4** and this minimal structure also better adhered to Lipinski's rule.

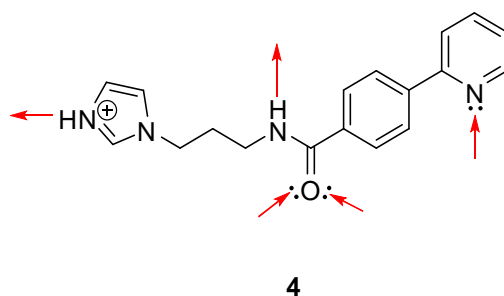
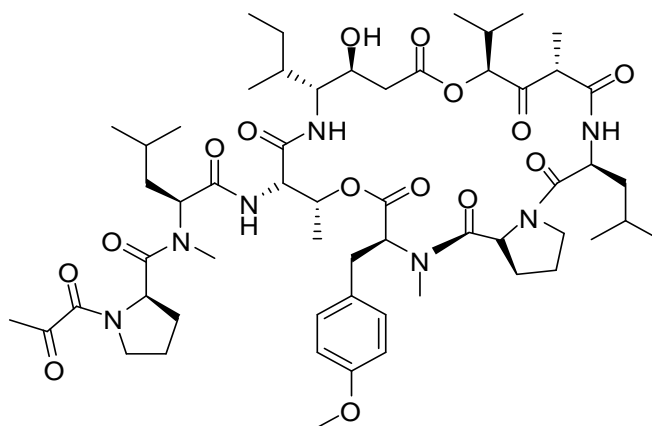


Figure 1.11. Pharmacophore features of compounds **1**, **2** and **3** outlined in red.⁶⁵

The minimal compounds synthesised retained activity, however, they showed reduced potency against selected cell lines ($IC_{50} = 0.9 \mu\text{M}$, compared with 0.01-0.079 μM for hit compounds **1**, **2** and **3** against the A2780 cell line).⁶⁵

1.4 APLIDINE

The NCI COMPARE data gave high correlations of hit compounds **1** ($P=1.0$), **2** ($P=0.748$) and **3** ($P=0.498$) with aplidine. Aplidine (plitidepsin / dehydrodidemnin B) is a marine antitumour natural product currently in phase II clinical trials for solid and haematological cancers, such as multiple myeloma, renal cancer and aggressive lymphomas.⁶⁶ The cyclic depsipeptide, originally isolated from the Mediterranean tunicate *Aplidum albican* and currently obtained by total synthesis,⁶⁷ shares structural similarities with didemnin B and other didemnins harvested from the tunicate *Trididemnum solidum*.⁶⁸ Aplidine has a cytotoxicity of < 10 nm against Ehrlich cells.⁶⁹



Aplidine

Aplidine has been described as a multifactor apoptosis inducer which has a low systemic toxicity and high specificity for tumour cells.⁶⁶ The mechanism of action of aplidine is only partially understood. Many mechanisms of action have been reported for the natural product, including blocking tumour cell proliferation, inhibiting cell cycle progression and suppressing tumour associated angiogenesis.⁷⁰

It is understood that aplidine is a potent inducer of apoptosis in a *p53*-independent and caspase-dependent manner.^{71, 72} One study suggests aplidine induces cellular oxidative stress producing ROS, increasing lipid and DNA oxidation, by GSH depletion. It also showed JNK activation is critical for aplidine-induced apoptosis. This is caused by two gene-expression independent mechanisms, Rac1 activation and MKP-1 down-regulation. Rac1 activation can be caused by GSH depletion.⁶⁷

Other research has shown apilidone treatment up-regulates phosphorylation of EGFR by a mechanism partly dependent on Src activation and sustained activation of JNK and p38 MAPK, due to a decrease in GSH levels and Src kinase, in breast and renal cell lines.⁷¹ In addition, apilidone causes induction of the mitochondrial apoptotic pathway, including the release of cytochrome *c*, activation of the caspase cascade and proteolytic cleavage of PKC δ generating a positive feedback loop which amplifies the apoptotic cascade.⁷³

Investigations into the mechanism of action of apilidone in melanoma cells (SK-MEL-28 and UACC-257) confirmed activation of Rac1 leading to JNK and p38 MAPK activation. However, at low concentrations of the drug, apilidone has a potent anti-proliferative effect caused by G₁ phase cell cycle arrest and G₂/M blockade, which is not seen in breast, renal and cervical carcinomas. This is supported by the fact cyclin A and cyclin B are at reduced levels in cells treated with apilidone.⁷⁴

Apilidone down-regulates the *flt-1* receptor for VEGF which is involved in angiogenesis. These data suggested the mechanism may involve blocking an autocrine loop relevant for cell growth and survival.⁷⁵ It has also been shown that apilidone inhibits the expression of multiple angiogenic genes.⁷⁶

Mapping apilidone onto the DTP complete SOM revealed it clustered into several different areas (N7, P1, P2, M2, M3 and M5). The areas represented are mitotic (M), membrane function (N), and phosphatases/kinase and oxidative stress (P). COMPARE analysis and data mining carried out by Rong suggested that the hit

compounds, **1**, **2**, and **3** would map in the P region of the SOM.⁶⁵ These combined results would imply that the MOA of the hit compounds may be related to protein kinase inhibition or oxidative stress.

1.5 TARGET DECONVOLUTION

A phenotypic approach to drug discovery is becoming more widely applied due to the limitations of target-orientated discovery. It involves HTS of large libraries of compounds against cell lines to identify a desired cellular response. Subsequently, the biological target(s) responsible for the change in phenotype is identified – target deconvolution. This approach extends beyond single target inhibition, which often leads to resistance, allowing signal transduction pathways to be explored.⁴⁴

There are many strategies which can lead to target deconvolution.⁷⁷ Directly exploiting a compounds affinity for its biological target(s) can be achieved by classical affinity chromatography, to purify a biological target using the immobilised compound, or by screening protein microarrays using the compound as a probe. Alternatively, indirect approaches involve looking for compound-induced changes in the protein, metabolite and mRNA expression profiles.

1.6 PROTEOMICS

Proteomics is a relatively new enabling technology that is becoming an integrated part of the drug discovery process.⁷⁸ It focuses primarily on understanding the structure, function, expression, cellular localisation, interacting partners and regulation of every protein produced from a whole genome.^{79, 80}

Proteomics seeks to profile every protein expressed in a target cell or tissue at any time. It has advantages over genomics as it is the proteins that directly assert the potential function of genes by means such as enzymatic catalysis, molecular signalling and physical interaction.⁸¹ In addition, a single gene may produce several different protein isoforms, each with different functions. Due to the differential splicing and translation of many genes in the genome, each human gene may encode multiple different proteins. Furthermore, proteomics can also identify the post-translational modifications of proteins⁷⁸ which have profound effects on the biological function and their cellular localisation.⁸² However, there is currently no equivalent to PCR for protein amplification, therefore detection of proteins in low abundance remains a problem.⁸³

Proteomic studies provide a wealth of information to aid understanding of biological processes, as proteins are involved in every biological pathway in living organisms.⁸⁴ Recent advances in proteomic research have identified target proteins involved in these pathways, leading to a better understanding of how new therapeutic agents can be established.⁸⁵ Cancers share a restricted

set of capabilities critical for their development and progression, 'The Hallmarks of Cancer' (Figure 1.9).³⁴ The protein expression pivotal to the hallmarks varies between patients with the disease and healthy individuals. These proteins can be considered as the key to finding new drug targets or biomarkers for detection and diagnosis of cancer.⁸⁶

During the early years of proteomics, protein profiling primarily relied on 2D Electrophoresis (2DE) to separate and detect proteins⁸⁷ and Edman degradation was the standard procedure for identification of the proteins from their amino acid sequence. Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) has become the standard technique for the separation of protein mixtures.⁸⁸ The technique separates denatured proteins on the basis of their molecular weight. Two-dimensional polyacrylamide gel electrophoresis separates proteins according to their isoelectric point, by isoelectric focussing, and their molecular weight in the second dimension by SDS-PAGE.⁸⁹ Edman degradation is the process of sequencing a protein or peptide from the amino terminus.⁹⁰ This was the most common approach to protein identification due to its reliability and automation, however, sensitivity remained poor and the process was slow.⁸⁴ Applications of mass spectrometry are now described as the 'workhorse' methodology of proteomics, allowing large, complex mixtures of proteins and peptides to be investigated.⁹¹

1.6.1 MASS SPECTROMETRY

Mass spectrometry is an analytical technique used to determine the structure, elemental composition and purity of a compound. The process relies on the formation of charged ions (positive or negative) in the gas phase which can be easily separated by their mass to charge ratio (m/z).⁹²

Early mass spectrometers were built to analyse small molecular weight compounds and used electron impact ionisation methods.⁹³ The harsh conditions used prevented the mass spectrometric analysis of larger biomolecules. Complex biological molecules such as proteins cannot be transmitted into the gaseous state required for mass spectrometry without degradation of their structure. Therefore, it was not until the introduction of soft ionisation techniques that biomolecules could be analysed using mass spectrometric techniques.⁹²

There are many different types of mass spectrometer, described by their ionisation method, mass analyser and detector. A mass spectrometer operates through three basic steps outlined in Figure 1.12.

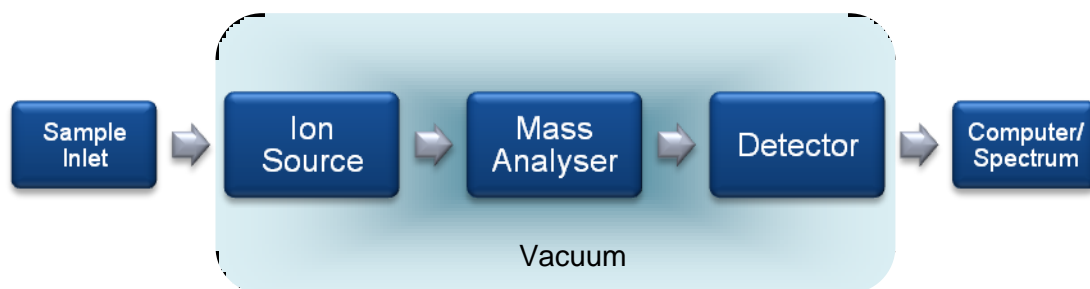


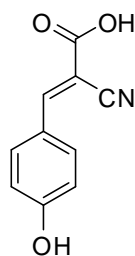
Figure 1.12. The basic components of a mass spectrometer.

A sample can enter a mass spectrometer by a variety of methods including HPLC, GC or direct injection. The first step then involves the formation of gaseous ions (positive or negative) within the ion source. There are several different ionisation methods available⁹⁴ and Matrix-Assisted Laser Desorption Ionisation (MALDI) is currently the most widely applied ionisation method used in protein research.⁹² In addition to proteomics, MALDI has been used to study other biomolecules, such as lipids⁹⁵ and DNA.⁹⁶

1.6.1.1 Matrix-Assisted Laser Desorption Ionisation

MALDI was first described by Franz Hillenkamp and Michael Karas in 1988⁹⁷ who demonstrated that alanine could be readily ionised using a laser beam of 266 nm when it was combined with tryptophan. Following their discovery, Kochi Tanaka established MALDI for large molecular weight biomolecules using a matrix of ultrafine cobalt powder in glycerol.⁹⁸ For this he was awarded the Nobel Prize for Chemistry in 2002. Many of the matrices that are used today are of an organic nature and these were first described by Karas and Hillenkamp who ionised proteins using a matrix of nicotinic acid.^{97, 99}

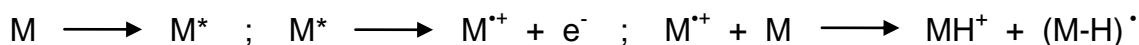
MALDI requires the sample to be mixed with a matrix prior to ionisation. The matrix strongly absorbs at a maximum wavelength matched to that of the laser used for the ionisation. α -Cyano-4-hydroxycinnamic acid (CHCA) is an organic compound commonly used as a matrix for MALDI.

 α -cyano-4-hydroxy cinnamic acid

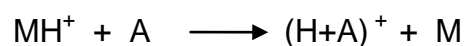
$$\lambda_{\max} = 337 \text{ nm}$$

An aqueous solution of CHCA is applied to a metal target plate and allowed to dry. The sample is then loaded in a diluted organic solvent and more matrix is applied. The analyte molecules adhere to the matrix molecules by hydrophilic interactions.¹⁰⁰ Upon evaporation of the solvent, the sample is concentrated onto the matrix to form a solid solution with isolated analyte molecules embedded within the matrix crystals.

The target plate is inserted into a vacuum lock and transferred to the ion source. A pulsed laser beam is applied to the homogeneous solid solution on the target with the same frequency as the chromophore of the matrix (337 nm) leading to vibrational excitation of the CHCA molecules. Desorption of the matrix and co-crystallised analyte puts them in the gas phase where the photo-excited matrix molecules (M) stabilise via proton transfer.



The charge on the matrix molecules is then transferred to the analyte molecules (A).



The analyte ions are accelerated using an electrostatic field to the mass spectrometer for analysis. The process is outlined in Figure 1.13.

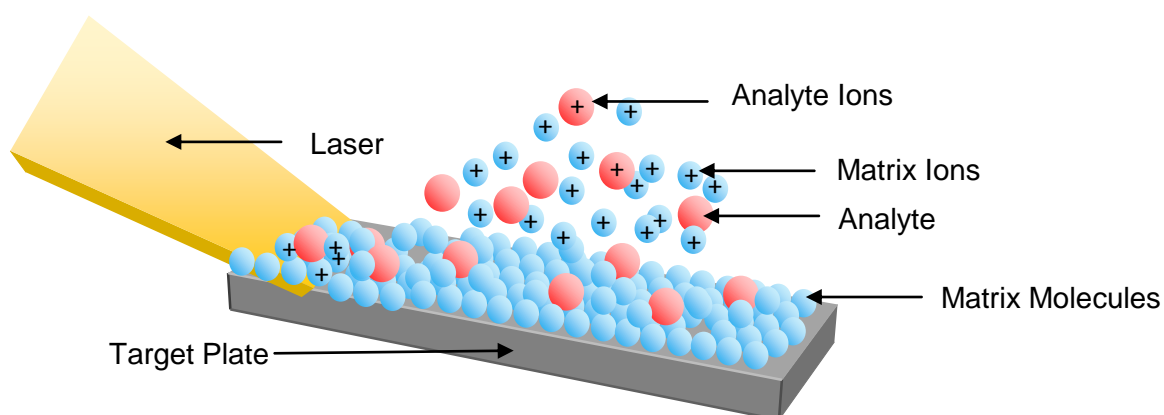


Figure 1.13. The MALDI ionisation process.

Once the gaseous ions have been formed the ions need to be separated based on their m/z ratio. A number of different mass analysers have been developed which can be used to separate the ions. All of the analysers work using static or dynamic electric and magnetic fields alone or in combination.⁹⁴

1.6.1.2 Time of Flight Mass Analysers

Time-of Flight (TOF) mass analysers were first introduced in 1946 by Stephens¹⁰¹ and have a relatively simple underlying principle. A basic TOF mass spectrometer consists of: a short source extraction region, where ions are produced and accelerated to a constant kinetic energy using electric fields; a drift region, which is a field-free region with a bounded extraction grid at ground potential; and a detector. The principle is that the ions cross into the drift region with a velocity that is inversely proportional to the square root of its mass.

$$v \propto \frac{1}{\sqrt{m}}$$

Therefore, lighter ions have a higher velocity and reach the detector in a faster time than heavier ions.¹⁰²

TOF-MS theoretically has no mass range limitations because the ions, however large, will still arrive at the detector in a particular time.¹⁰³ The resolution of the data generated is dominated by the error in flight time and the difference in pulsed action and energy aberration of the laser. The introduction of MALDI with a TOF mass analyser has simplified the TOF design and has a great effect in eliminating error due to time and spatial distribution.¹⁰² MALDI produces gaseous ions using a pulsed laser, therefore leading to a time distribution (Figure 1.14A). Ions are extracted from the source using a high velocity pulse after a time delay, hence they are accelerated and enter the drift region at the same time,¹⁰³ (Figure 1.14B). This is known as 'time-lag focussing' and allows the ions within the source to rearrange their position with respect to their initial velocities. Therefore, ions with higher velocities travelling into the drift region will gain less energy from the extraction field when it is switched on,¹⁰⁴ (Figure 1.14).

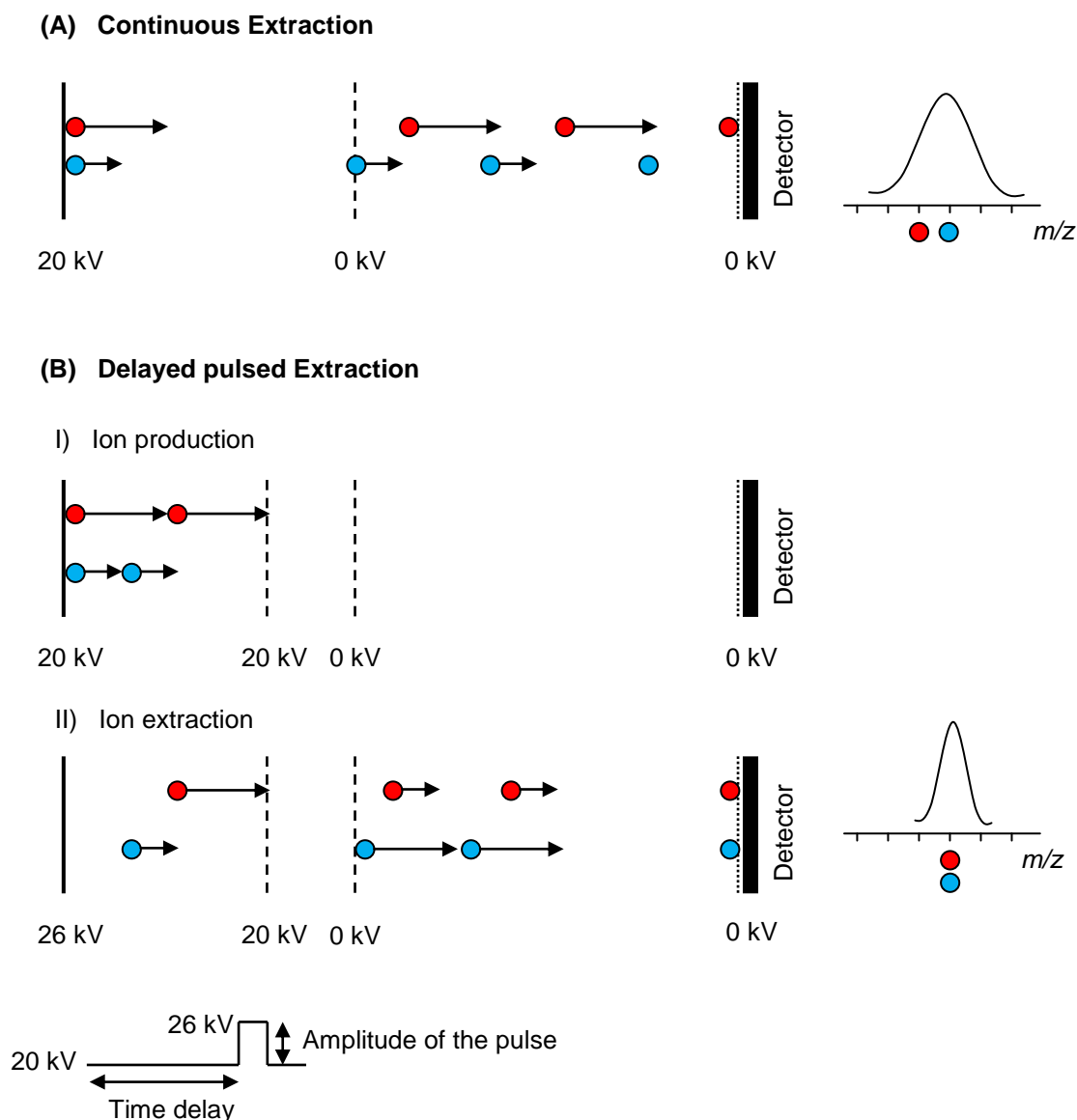


Figure 1.14. Continuous and delayed pulse extraction mode in a TOF analyser. Blue ions of a certain mass have the correct kinetic energy. Red ions have the same mass, however a higher kinetic energy. Delayed pulse extraction corrects the energy dispersion.⁹⁴

Another improvement to the TOF mass analyser came with the development of an electrostatic reflector called the reflectron.¹⁰⁵ This is basically an ion mirror which deflects the ions back through the flight tube.⁹⁴ Once ions reach the reflector they are deflected by an electric field and continue towards the detector.¹⁰³ This compensates for any difference in the initial energy and can

therefore separate ions with the same m/z with increased resolution. This is because ions with the same m/z may have differences in their kinetic energy. The reflectron allows the ions with a higher kinetic energy to move more deeply into the reflectron than ions with a lower energy, therefore correcting the time of flight (Figure 1.15).

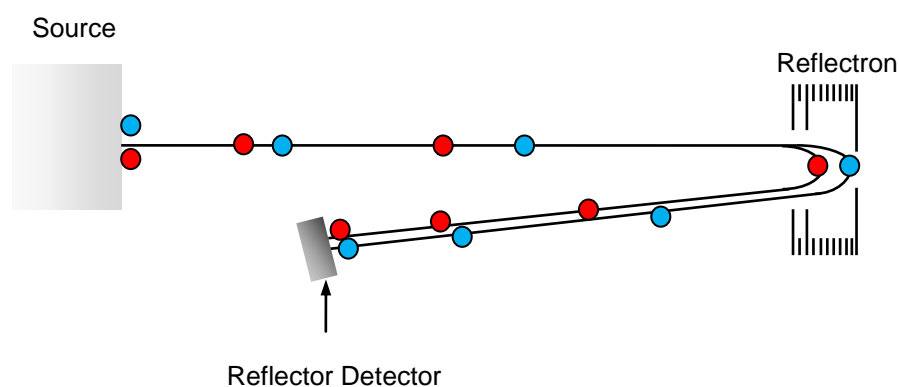


Figure 1.15. A reflectron. Blue and red ions have the same mass, however red ions have a kinetic energy which is too low and therefore enter the reflectron later but do not travel as deeply. Both ions leave the reflectron with the same kinetic energy as before and hence reach the detector at the same time.⁹⁴

Fragmentation of ions due to their high kinetic energy after acceleration may occur before leaving the source, known as in-source decay (ISD) fragments, or after leaving the source, which are known as post-source decay (PSD) fragments. ISD fragments can be separated from their parent ion on the basis of their m/z in the TOF analyser because they have the same kinetic energy and therefore different velocities on entering the drift tube. However, PSD fragments have the same velocity during their journey through the drift region and therefore, the same flight time in the absence of a field. The reflectron TOF MS

has therefore, allowed distinction between parent and PSD fragments as the fragment ions will not move as deeply into the reflectron as their parent ions.⁹⁴

1.6.1.3 Tandem Mass Spectrometry

Tandem mass spectrometry (MS/MS) incorporates multiple stages of mass analysis. This can be useful in peptide identification due to their fragmentation patterns which are representative of their amino acid sequence. A TOF reflectron mass analyser can be used to perform MS/MS. The best procedure combines a linear TOF and a reflectron TOF analyser.⁹⁴ During MS/MS analysis, parent ions can be identified in the linear TOF, whereas fragments can be detected using the reflectron TOF mass analyser. Collision induced dissociation (CID) is a process by which parent ions decay by collisions during passage through a cell filled with an unreactive gas, e.g. Ar. In MS/MS, CID allows fragmentation of ions selected by the first mass analyser and the fragmentation products to be analysed by the second detector. CID leads to an increase in the number of ions that fragment and also increases the number of fragmentation pathways.⁹⁴

1.6.1.4 Instrumentation

The Bruker Daltonics Ultraflex II is a MALDI tandem mass spectrometer which has been specifically designed for automated MS and MS/MS high throughput identification of proteins and peptides. The instrument uses a MALDI ionisation source and has two TOF mass analysers (Figure 1.16).¹⁰⁶

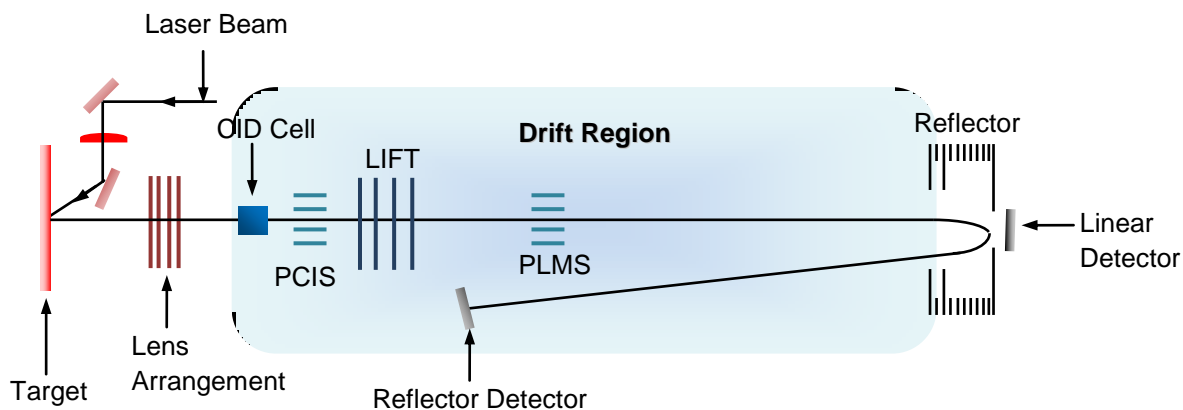


Figure 1.16. The Bruker Daltonics Ultraflex II

The instrument has an implemented 'LIFT' device which allows full fragmentation spectra to be produced in the MS/MS mode. In MS mode, the LIFT feature is disabled and therefore only the parent ion is detected. In MS/MS mode the LIFT raises the kinetic energy of the parent and fragment ions allowing the detection of the fragments with the parent. Fragment ions have ground potential in comparison to the parent. The 'lift' in kinetic energy allows both parent and its fragment ions to travel down the flight tube with the same velocity. Upon entering the reflectron the ions with lower kinetic energy – fragments – do not travel as deep and are detected by the reflectron detector in a shorter flight time than the parent.

Another feature of the instrument is that it has a precursor ion selector (PCIS) which enables selection a particular parent ion and its fragments, then separates them from other ions for MS/MS analysis. It consists of deflector plates which can be switched on and off to allow only ions with a certain velocity to pass through. A second deflector, known as the post-lift metastable

suppressor (PLMS), deflects any remaining intact precursor ions, preventing undesired fragmentation which may occur following acceleration.¹⁰⁷

Protein identification is achieved by associating the isolated protein with a gene and hence a biochemical function based on its sequence identity revealed following mass spectrometric analysis.¹⁰⁸

1.6.2 PROTEIN IDENTIFICATION

Protein identification is achieved through the analysis of peptides produced following enzymatic digestion of a protein into smaller peptides.⁸⁵ Enzymatic digestion with trypsin exclusively hydrolyses the peptide bond formed at the C-terminus of both arginine and lysine residues in a protein (provided the next amino acid in the sequence is not proline), resulting in a variety of shorter peptides.⁸³ The peptides produced, therefore have two protonation sites, the N-terminal amino group and the C-terminal lysine or arginine side chain. The peptide fragments produced each have a mass which is characteristic of their amino acid sequence and hence one particular protein which is known as the Peptide Mass Fingerprint (PMF).^{83, 109} The PMF produced can be compared against *in silico* digestion of protein sequences in databases to produce a list of likely matches.¹⁰⁸ Protein sequence databases are constantly updated with amino acid sequences generated by the translation of nucleotide sequences in their correct reading frame.¹⁰⁸

The PMF cannot always identify a protein, especially if the protein is part of a mixture,⁸³ or there is an amino acid substitution or post-translational modification.¹⁰⁸ Therefore further analysis using MS/MS can be conducted. Under CID conditions, peptides fragment to create patterns characteristic of the amino acid sequence. These fragmentation patterns are reproducible and predictable if a sequence is known.¹¹⁰

Peptide CID spectra are generally more useful for protein identification than PMFs, as they provide information about the peptide sequence in addition to the peptide molecular mass.¹¹¹ The amino acid sequence can be manually generated from the series of daughter ions identified in a MS/MS spectrum.¹¹² Manual interpretation and subsequent identification relies on identifying various daughter ions whose mass difference is equal to the exact mass of one amino acid residue.¹¹⁰

Amino-acid specific fragment ions are primarily produced from the cleavage of amide bonds in the peptide backbone during MS/MS analysis. Depending on the basicity of the amino acids within the peptide, the charge can be retained upon the *N*-terminus giving a acylium ion, known as a 'b' ion or by proton rearrangement be retained on *C*-terminus leading to 'y' ions^{85, 110} Fragmentation can also occur at other points along the peptide backbone (Figure 1.17).

The most frequently occurring ions are type a and b and these are generally seen when there is no basic amino acid or it is located near the *N*-terminus; y ions are observed if the basic amino acid is located at the *C*-terminus of the

peptide, as is the case of trypsin generated peptides. All three are generally observed with one or two dominating.¹¹³ The ions are thought to be produced following protonation of the amide nitrogen.¹¹⁴

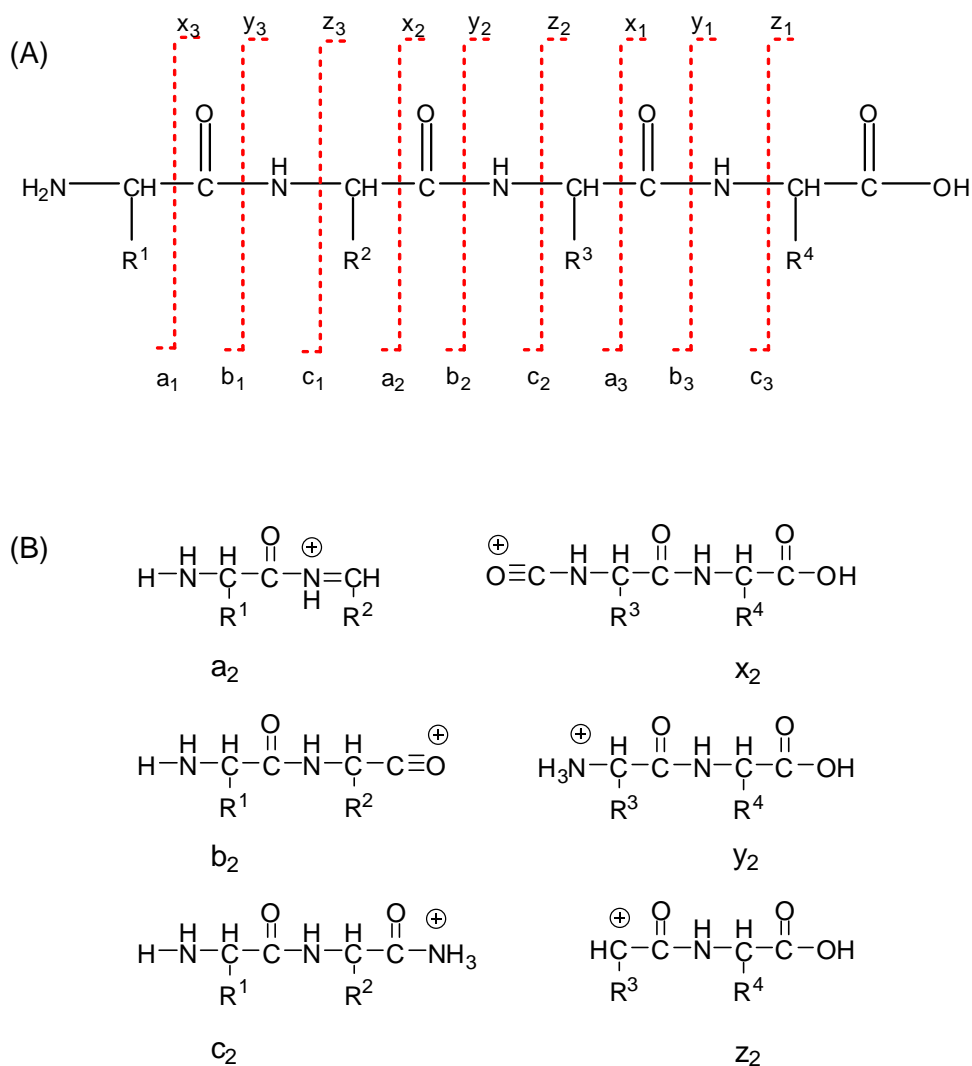


Figure 1.17. Fragmentation patterns of peptides showing: (A) the different cleavage positions in the peptide backbone; (B) the residual fragments produced following cleavage.

Several algorithms and computer programmes have been designed to search protein sequence databases and identify proteins. ‘Mascot’¹¹⁵ is a search engine which incorporates probability-based protein and peptide identification. It

is a development on the MOWSE molecular weight search engine,¹¹⁶ which was one of the first programmes for identifying proteins by PMF. Mascot can conduct three different types of searches, PMFs; sequence queries; and tandem MS ion searches.

For protein identification, MOWSE calculates the sequence molecular weights of all proteins in the database (DMW) and calculates the molecular weights of the peptides produced following complete enzymatic digests (FMW). The peptide mass calculations are based on the linear sequence and use the average isotopic masses of amide bonded amino acid residues¹¹⁶ described in Equation 1. A hit is scored when the criterion from Equation 2 is met where tolerance is user selective.¹¹⁶

$$1 \quad \sum_{n-1}^n \text{residue mass} + 18.0153$$

$$2 \quad \text{DMW} - \text{tolerance} - 1 < \text{FMW} < \text{DMW} + \text{tolerance} + 1$$

The programme generates a ranked 'hit' list of proteins and each protein identification is given a Mascot score, which is based on the absolute probability that the observed match is a random event. Specifically the scoring is based on the frequency of a fragment molecular weight being found in a protein of a given weight. Scores are calculated as: $-10 \log_{10}(p)$; where p is the absolute probability. Therefore a very low probability will give a high Mascot score. The match with the lowest probability is reported as the 'best' match. Good matches

have a probability of $p < 0.05$. Mascot searches can also take into account post-translational modifications and modifications due to chemical derivatisation.¹¹⁷

Proteomics coupled to mass spectrometry and database searching can be used in drug discovery to indicate novel targets by comparing the proteomes from normal and neoplastic tissues. Differences in each of the proteomes could identify potential targets for drug discovery. In addition to large scale proteome screening, chemical proteomics can be used to identify, enrich and isolate targets as a more direct method to target deconvolution.

1.6.3 CHEMICAL PROTEOMICS

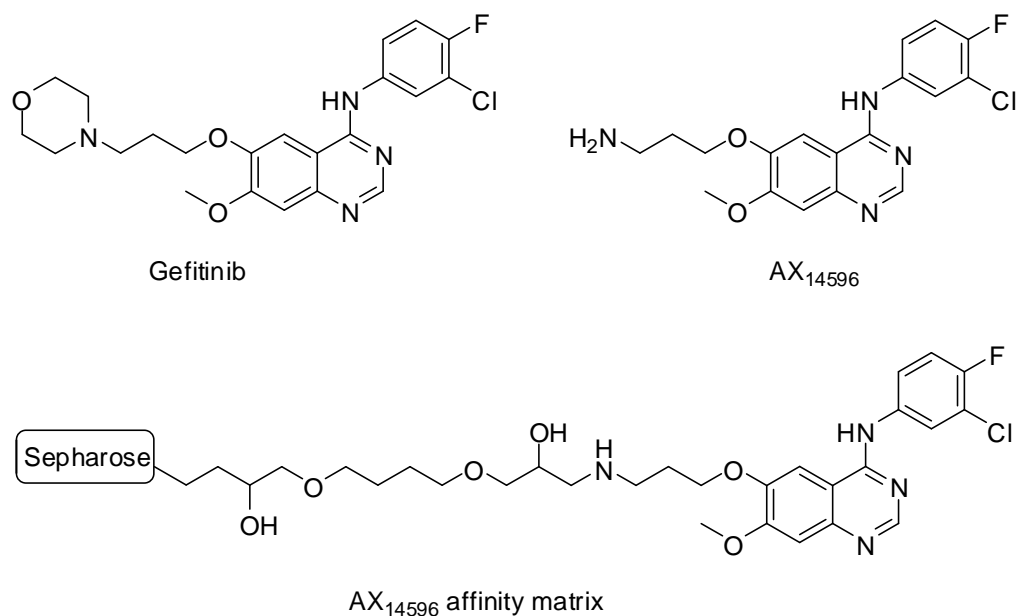
Chemical proteomics can take two different forms (Figure 1.18). The first is activity-based probe profiling (ABPP). This approach focuses on the enzymatic activity of a particular protein family. The second form is a compound-centric approach which aims to characterise the molecular mechanism of action of a small molecule inhibitor.¹¹⁸

The compound-centric approach can be described as the classical drug affinity chromatography¹¹⁸ and has proven a successful approach to studying inhibitor function. It utilises small molecular derivatives which can be covalently immobilised onto beads for affinity purification of cellular target proteins.¹¹⁹ Following direct immobilisation of a compound on to a solid support, a protein lysate can be incubated with the resin. Subsequent to the affinity capture and

several washing steps, the bound proteins can be eluted by denaturation or competitive elution using the free ligand for identification.³⁹

Brehmer *et al.*¹²⁰ used a chemical proteomic strategy to identify any additional targets of the EGFR kinase inhibitor Gefitinib used in the treatment of NSC lung cancer. They synthesised an analogue of Gefitinib incorporating a free amine in place of a morpholino group for direct immobilisation to a solid support.

The affinity matrices were incubated with total HeLa cell lysate. Non-binding proteins were removed from the column by washing and bound targets removed by competitive washing with free gefitinib and ATP. LC-MS/MS analysis of trypsin digested target proteins revealed more than twenty kinases and other cellular proteins.



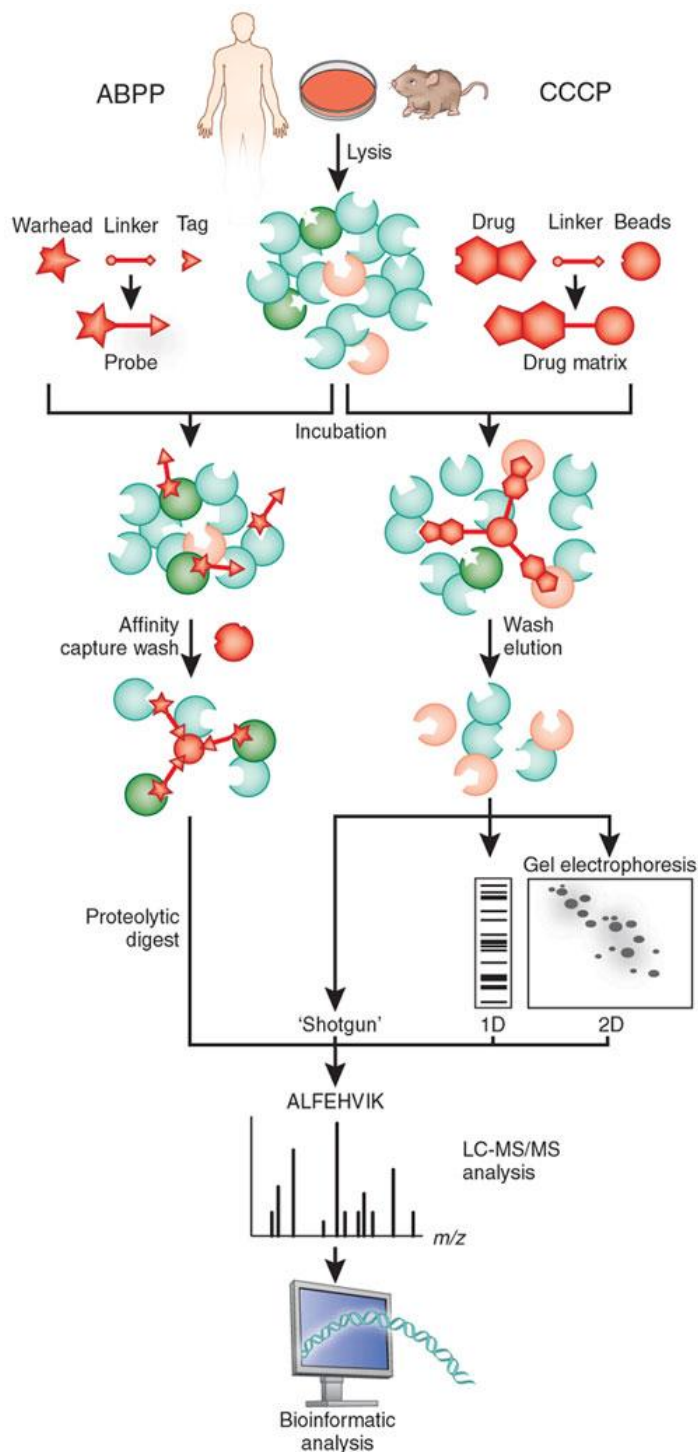


Figure 1.18. Comparison of activity-based probe profiling (ABPP) and compound-centric chemical proteomics (CCCP).¹¹⁸ ABPP forms a covalent bond between a warhead and a protein. The warhead can be an irreversible inhibitor attached by a linker to a reporter tag such as biotin. Bound proteins are then captured on an affinity matrix through the reporter tag using an affinity wash. CCCP uses a compound of interest bound to an inert matrix which is incubated with a biological sample for capturing proteins. Unbound proteins are washed from the matrix.

1.7 CONCLUSION

The three hit compounds synthesised by Phillips⁶² have COMPARE profiles indicative of a novel mechanism of action. The target of compounds, **1**, **2** and **3**, is present in sensitive cell lines from the NCI 60-cell line screen, in particular leukaemia and NSC lung cancer subpanels. COMPARE analysis with the molecular targets database identified that the protein target is from a 'gene-like' sequence of unknown function. Isolation and subsequent identification of the biological target, could present attractive possibilities for optimisation of the hit compounds and indicate a new molecular target for drug discovery.

Isolating the target will be attempted using an affinity-based proteomic technique and subsequently identified using MALDI-TOF mass spectrometry and Mascot. Isolation requires affinity probes of the hit compounds. Therefore, the hit compounds must be immobilised on a solid support. Immobilisation requires inclusion of an additional functional group into the hit compounds that can be used for reaction with the chromatography medium, or alternatively incorporating a moiety such as biotin, into the compounds for complexation with a streptavidin-based matrix.

Any alteration to the structures of the compounds may affect their cytotoxicity. Therefore, all compounds must be tested before affinity chromatography is carried out. During the affinity chromatography procedure, compounds of similar structure demonstrating no potency will also be required to act as controls. This will eliminate non-specific binding of cellular proteins.

1.8 AIMS

The aims of this work are:

- To synthesise bespoke matrices of the three hit compounds for affinity chromatography.
- To synthesise bespoke matrices of an 'inactive' compound to act as a control.
- To evaluate the biological responses of the modified hit compounds prior to attachment to a matrix.
- To use a proteomics-mass spectrometry programme to isolate, identify and characterise target proteins.

CHAPTER 2

DESIGN OF THE AFFINITY PROBES

2.1 INTRODUCTION

The hit compounds previously identified in this laboratory⁶² possess common structural features only differing in the core heterocyclic ring. Figure 2.1 highlights the differences (coloured) and similarities (black) between each of the structures. The similarities can also be described as the pharmacophore identified by Dawen Rong.⁶⁵

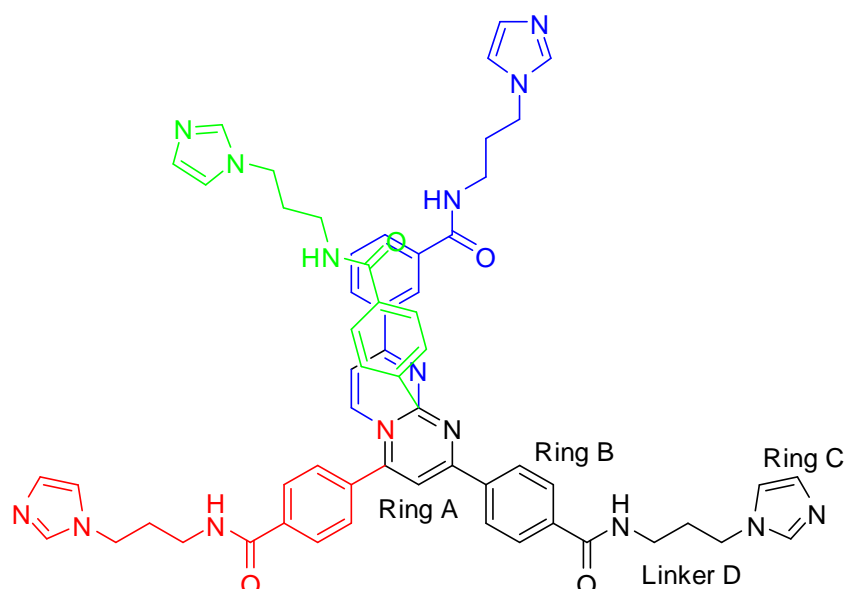


Figure 2.1. Superimposition of the three hit compounds **1** (red), **2** (green) and **3** (blue). The structure **4**, common to all three ligands is shown in black.⁶²

The preliminary SAR for the compounds⁶² showed:

- *N*-oxides of Ring A block cytotoxic activity;
- *Meta*-substitution of ring B results in a decrease in activity and selectivity;

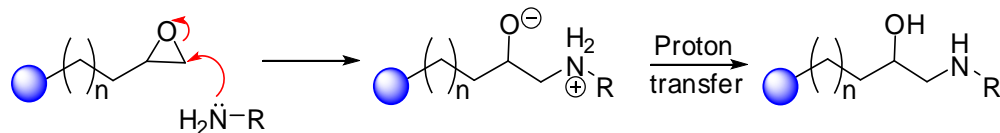
- Reduction in the length of linker D reduces activity in ligands;
- Saturated rings in the place of ring C reduce cytotoxicity, dialkyl amines are inactive.

These SAR data, in addition to the evidence that compound **4** is the minimal structure required for activity,⁶⁵ reveal a number of accessible sites for attachment of the hit compounds to a matrix for use in affinity chromatography. As the compounds are symmetrical, a linker, containing a suitable functional group for attachment to a solid support, could be attached at various points within half of the compound, leaving the other half unchanged and accessible to the biological target.

Evidence shows compound **4** to have a decreased potency compared with hit compound **1** (IC_{50} 2.6 μ M against the A2780 cell line for compound **4**⁶⁵ compared with 0.079 μ M for hit compound **1**⁶²), suggesting that the biaryl heterocyclic ring system, may be required for activity, or that the target has a symmetrical binding site.

2.2 CHROMATOGRAPHY MEDIUM

The hit compounds can be immobilised directly onto a chromatography medium by a suitable functional group, such as a carboxylic acid or an amine. Epoxy-activated beads can be used to attach the hit compounds to the solid support by an amine (Scheme 2.1).



Scheme 2.1. Mechanism of a hit compound containing a primary amine reacting with epoxy-activated chromatography media.

Alternatively, instead of immobilising the hit compounds directly to the solid support, a more ‘universal’ approach is to use streptavidin or avidin columns with biotin-derivatised compounds tethered to the resin.³⁹ This strategy utilises the high affinity of biotin for (strept)avidin. However, a biotin tag could change the binding characteristics and directly affect the activity of the hit compounds.

2.2.1 THE BIOTIN/STREPTAVIDIN BOND

Streptavidin, isolated from *Streptomyces avidinii*, is a 159 residue, tetrameric protein with a molecular weight of $4 \times 15,000$.¹²¹ The protein is a dimer of dimers with 222 (d2) symmetry.¹²² It shares homology with the protein avidin which is isolated from egg white. Both proteins exhibit an exceptionally-high affinity for biotin ($K_d \sim 10^{15} \text{ M}^{-1}$) which is the strongest non-covalent bond known in nature between a ligand and a protein.¹²³ In the unbound state, the binding pocket is occupied by five water molecules, which abandon the binding site upon addition of biotin to the solution.¹²² Each subunit binds one biotin molecule and the binding energy is derived from several interactions between the protein and biotin. There are seven direct hydrogen bonds, five of which lie deep within the binding pocket, preventing competition from solvent. These include two to

the single ureido oxygen, one from each of the ureido nitrogens and one to the sulphur in the biotin molecule.¹²⁴ In addition there are hydrophobic and Van der Waals contributions with tryptophan residues 79, 92, 108 and 120¹²⁵ (Figure 2.2).

Each monomer of the streptavidin protein folds into an eight-stranded antiparallel β -barrel. Biotin binds into the open end of the twisted barrel and a surface loop folds over the binding site. This binding loop consists of 45-52 residues.¹²² This could present a problem for utilising the biotin streptavidin bond in affinity chromatography, as biotin binds directly into the cleft of streptavidin. Therefore, the hit compounds must be designed to have the correct accessibility to the target during incubation steps.³⁹

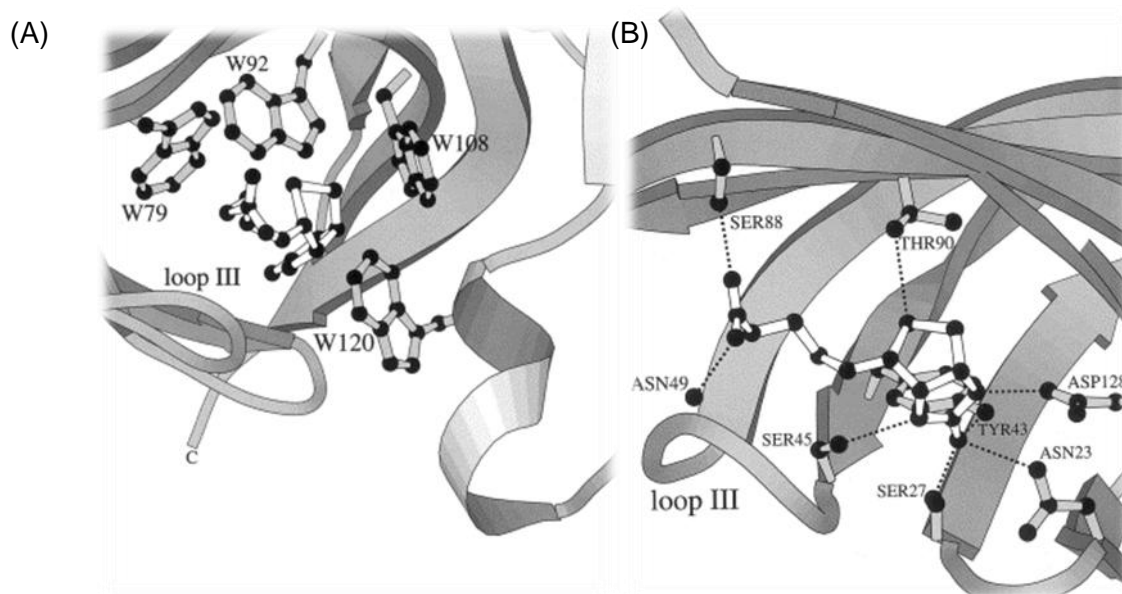
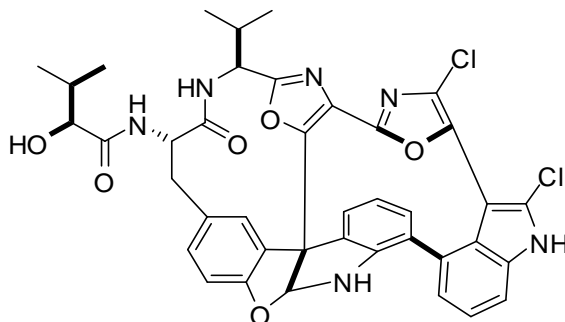


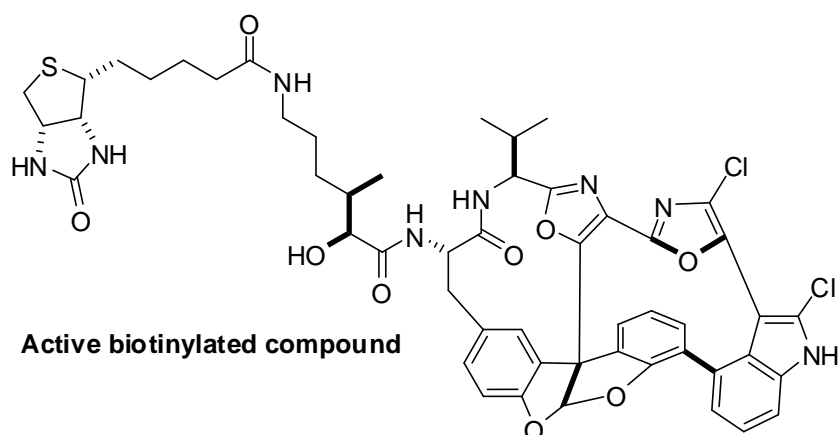
Figure 2.2. The binding of biotin (white) to streptavidin. The binding loop (loop III) bends over the biotin molecule. (A) Tryptophan residues involved in binding; (B) The hydrogen bonding system between ligand and protein.¹²⁵

Wang *et al.* employed a biotinylation approach to affinity capture to identify the binding protein of the antimetabolic agent Diazonamide A.¹²⁶

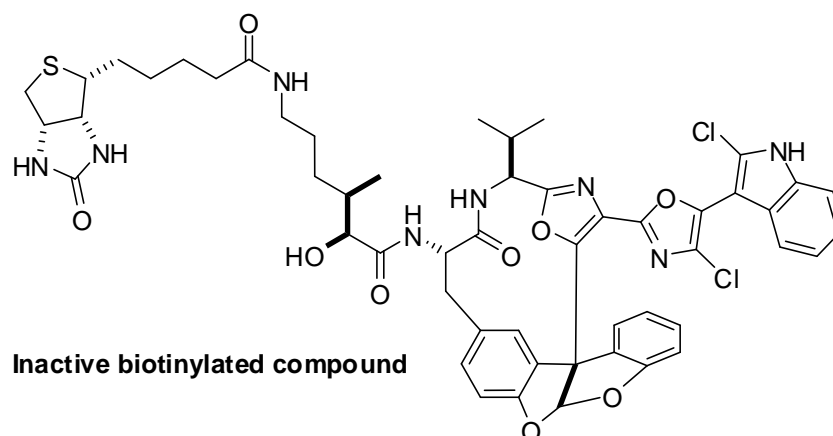


Diazonamide A

Cellular extracts of HeLa cells were probed with a biotinylated Diazonamide A and an inactive control. Mass spectrometry identified the cellular receptor that mediates the antimetabolic effects as ornithine aminotransferase (OAT). Biotin is attached to the compound by a three carbon linker, therefore a linker of no less than three carbon atoms is essential for attaching the hit compounds, **1**, **2** and **3** to affinity matrices.



Active biotinylated compound



2.3 BESPOKE MATRICES FOR AFFINITY CHROMATOGRAPHY

Synthesis of the bespoke matrices for affinity chromatography, following either of the methods described above, will involve addition of an extra functional group to the three hit compounds. However, the position of the functional group must be in such a place to ensure that the biological activity of the compounds is not compromised.

Taking into account the SAR already established for this group of compounds (Section 2.1), there are several sites for an extra functional group to be positioned, these are identified in Figure 2.3.

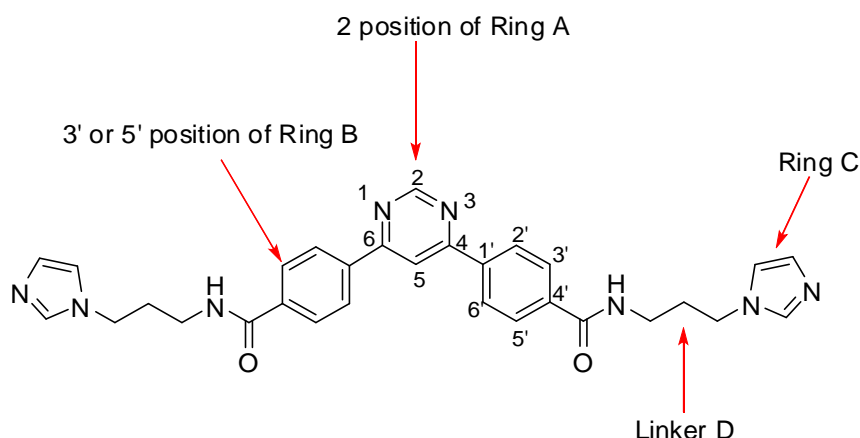


Figure 2.3. Compound 1, showing possible sites of incorporating an additional functional group.

The central ring (Ring A) could include an additional functional group, such as an amine. For compound 1, illustrated above, the position which will be the most synthetically accessible and least affect the 3D structure would be the pyrimidine-2 position. An alternative for all three hit compounds would be to position the functional group within one of the side chains, therefore disrupting symmetry. This could be placed either in Linker D or in place of Ring C. Ring B could be modified to carry a functional group at either of the unsubstituted positions (2'/6' or 3'/5').

The triaryl ring systems (ring A and rings B) of all three compounds are conjugated, but they do not adopt a rigid, planar structure. Instead they adopt a slightly rotated conformation (Figure 2.4C) due to steric hindrance effects between the hydrogen atoms *ortho* to the inter-annular bonds (Figure 2.4A). In contrast the bond between Ring A and Ring B cannot have a 90° torsion angle, as conjugation between the rings is lost due to a reduction in π -overlap (Figure

2.4B).⁶² SAR studies have shown the more planar the conjugated system, the higher the cytotoxicity of the compound.⁶⁵ *N*-oxidation of the nitrogen atom of the heterocyclic ring A reduces potency (IC_{50} 0.24 μ M compared with 0.079 μ M for hit compound **1**)⁶² suggesting that the nitrogen atoms may be involved in the binding to the protein target. In view of the SAR data from Rong concerning the effect of the planarity of the conjugated system, it would imply the N atoms are required for a conformational effect related to the protein:ligand binding complex. Therefore, the nitrogen atoms must be free to act as either hydrogen bond acceptors or to govern the adopted conformation which may be required for binding. In addition the 2'/6' positions of Ring B and the pyrimidine-5 position must also be kept clear to prevent conformational changes of the 3D structure of the hit compounds.

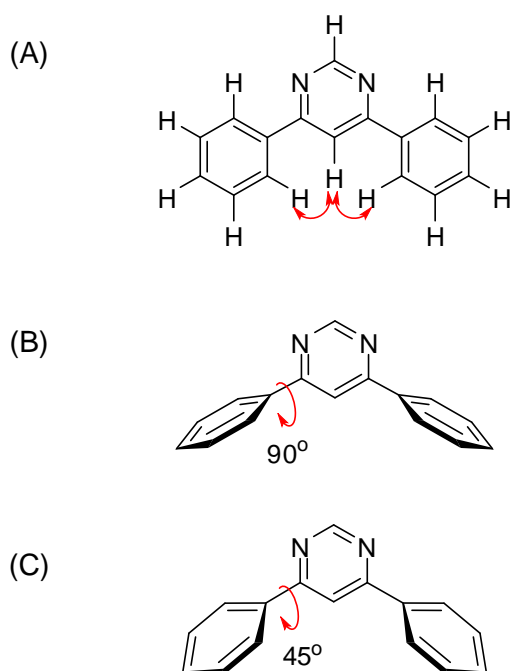


Figure 2.4. 4,6-Diphenylpyrimidine: (A) in a planar conformation, red arrows indicate the steric clash; (B) with a 90° torsion angle of the inter-annular bond, π overlap is reduced; (C) in the minimum energy conformation, a compromise of π overlap and steric clash.⁶²

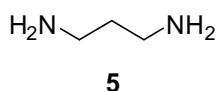
SAR data has shown that an increase in the length of Linker D, from a two carbon to a three carbon chain, causes a significant increase in potency of the compound series ($IC_{50} > 100 \mu M^{65}$ compared with $0.079 \mu M$ for hit compound **1**⁶²). Thus suggesting Linker D plays an important role in the binding of the hit compounds to the target. Therefore, at least one of the linker Ds must remain constant to ensure binding. In addition, the terminal imidazole group (Ring C) plays an essential role in the observed cytotoxic activity.^{62, 65} Replacing the imidazole groups with saturated heterocycles or alkyl amine groups resulted in a decrease in potency ($IC_{50} = 0.079 \mu M$ for compound **1** compared with $>100 \mu M$ for the pyrrolidinyl and dimethylamino equivalents).⁶² Taking into account the chemosensitivity results for compound **4**, it appears that only one of the imidazole groups is essential for binding to the biological target. Therefore, Ring C could be replaced with a nucleophilic group, such as a hydroxyl or primary amine group for attachment to biotin or directly to the epoxy spacer arm of the chromatography medium.

The biological target of the hit compounds is unknown and hence the binding mode. Therefore, the position of an additional functional group, required for attachment to either a solid support or biotin, must be in such a place as to ensure the final affinity probes maintain as much structural similarity to the hit compounds as possible. Thus all hydrogen bond donors and acceptors should remain uninterrupted, in at least half of the compound. Reviewing the SAR studies, it appears the triaryl ring system (Ring A and Rings B) is required for maximum activity. However, if one of the imidazole groups (Ring C) is omitted,

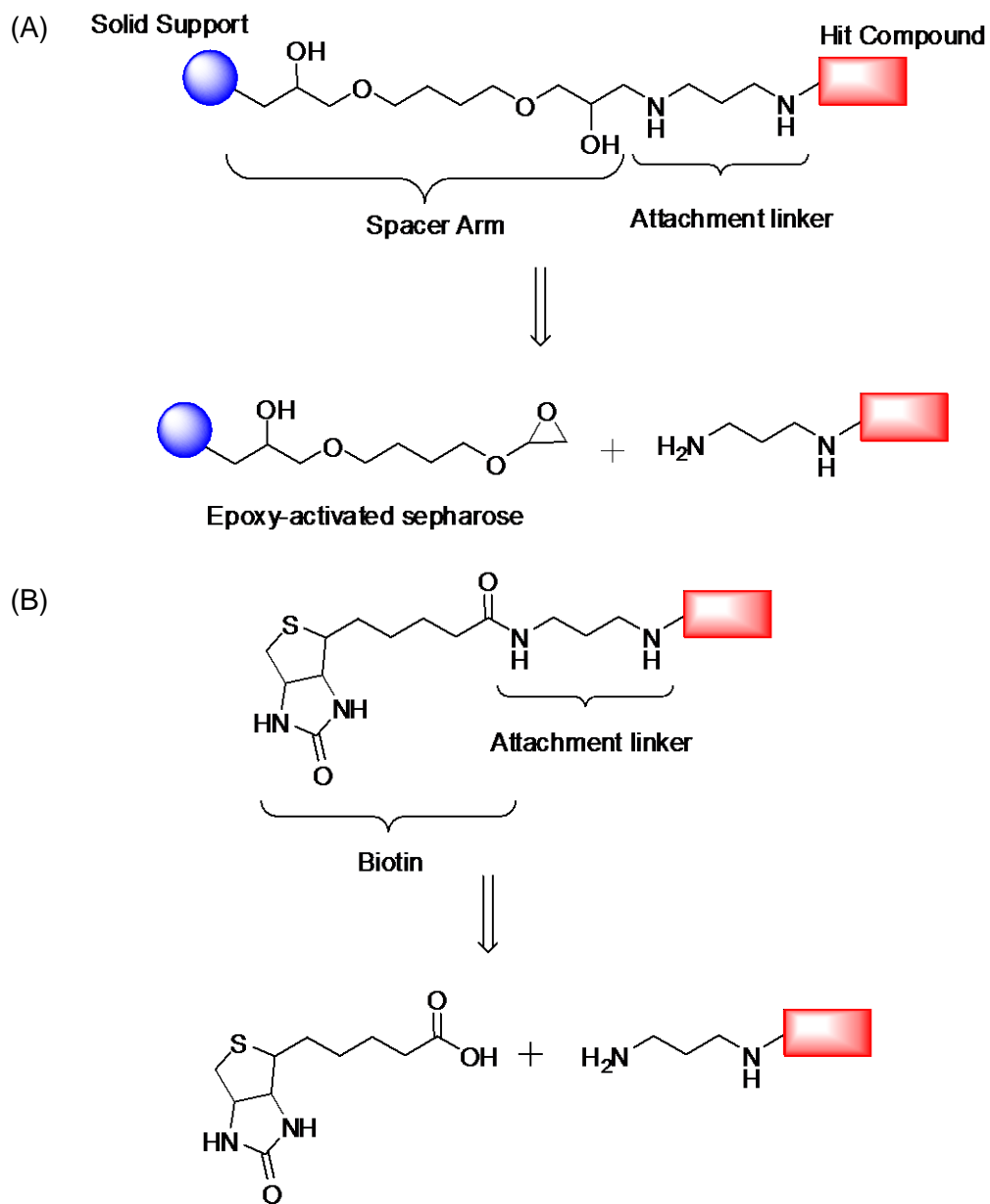
activity remains. Therefore addition of a nucleophilic group in place of Ring C, ought to be the best site for attachment to all three hit compounds.

2.4 ATTACHMENT LINKER

Attachment of the hit compounds to either biotin or directly to a solid support, requires the addition of a functional group to the hit compounds. There must also be a relatively long spacer arm attached to the solid support before binding of the hit compounds (whether attachment is direct or indirect through streptavidin-biotin). This would ensure the prevention of interference with ligand-protein recognition and capture. Commercially-available solid supports contain a long spacer arm (12-18 atoms), however a linker region would have to be incorporated to attach the functional group to the spacer arm of the solid support or to a biotin molecule. A suitable linker is 1,3-diamino propane, **5**.



1,3-Diaminopropane, **5**, is a bifunctional compound allowing both attachment to the hit compounds and to the chromatography medium or biotin (Scheme 2.2). Also, as the linker contains an aminopropyl group, it is possible to replace the whole of the side chain in one half of compounds preserving the original amide and carbon chain of Linker D.



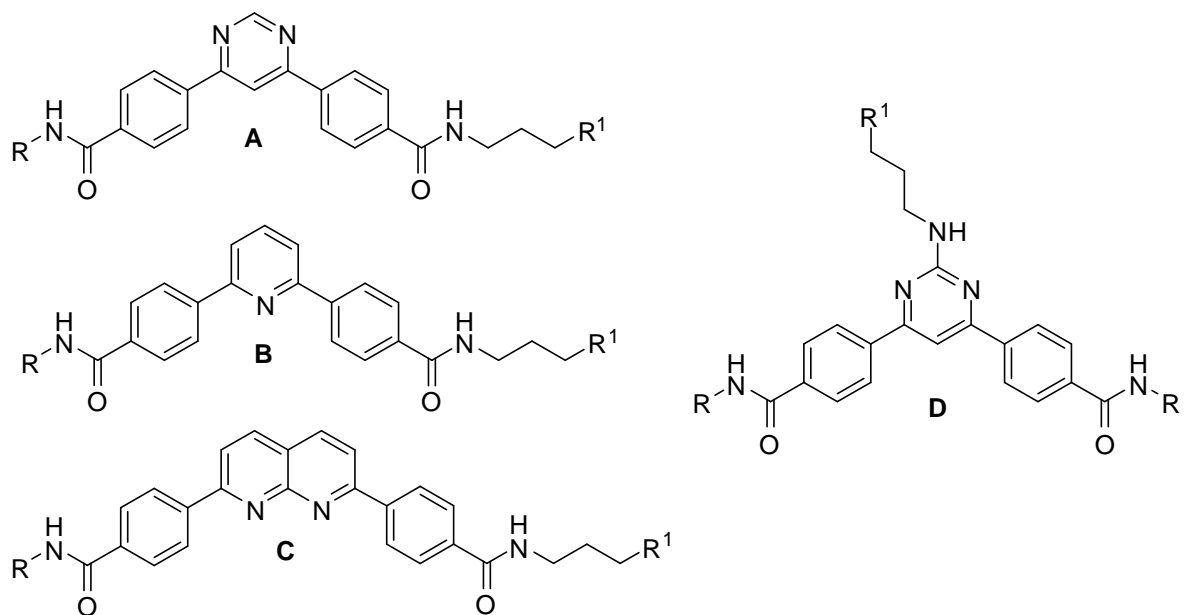
Scheme 2.2. Retrosynthetic analysis of the attachment of the linker to: (A) the spacer arm of the chromatography medium; (B) a biotin molecule.

2.5 PROTEOMIC STRATEGY

The preparation of affinity probes of the hit compounds with different attachment points of the linker should prove useful in identifying the target protein through affinity chromatography techniques. Using different attachment points should ensure that at least one of the immobilised hit compounds will maintain its ability to bind the target protein effectively. Also, compounds with free amines and compounds with biotin moieties will be synthesised to allow optimisation of methods for the chemoproteomics stage of the project.

Furthermore, compounds which are inactive towards cell lines sensitive to the original hit compounds will be synthesised for use as controls during isolation of the biological target. Since saturated cyclic amines show reduced toxicity, compounds with a terminal pyrrolidine group will also be synthesised.

The target compounds to be synthesised are shown in Figure 2.5.



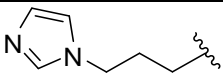
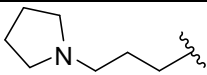
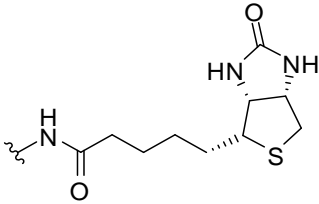
R			R ¹
A	6	7	
B	8	9	
C	10	11	NH ₂
D	12	13	
A	14	15	
B	16	17	
C	18	19	
D	20	21	

Figure 2.5. Target molecules for synthesis.

CHAPTER 3

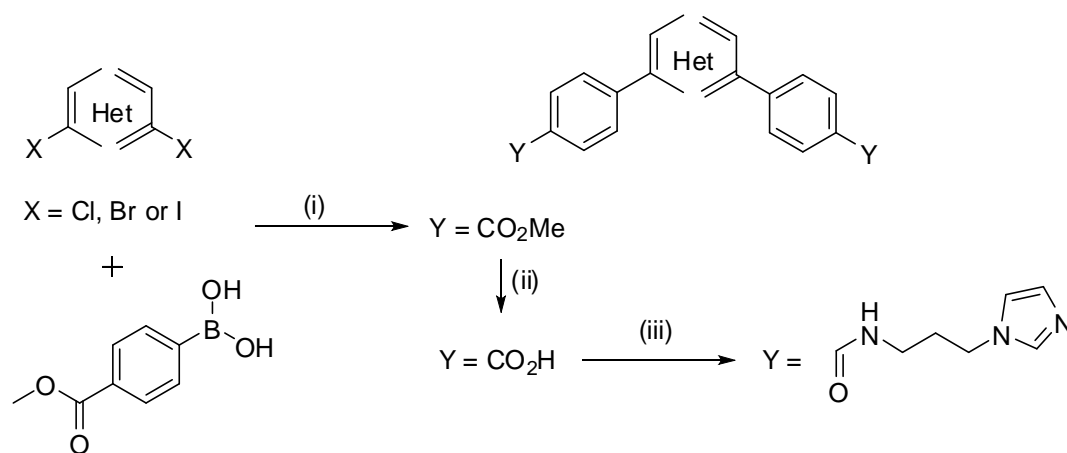
SYNTHESIS

3.1 INTRODUCTION

The target compounds for synthesis were defined in Chapter 2 and are summarised in Figure 2.5. The two main approaches to the target compounds are:

- Breaking the symmetry of the hit compounds and replacing one of the side chains with the linker 1,3-diaminopropane, **5 (A-C)**;
- Retaining the symmetry and attaching the linker at the pyrimidine-2 position (**D**).

The hit compounds, **1**, **2** and **3**, have previously been synthesised by a versatile three-step process outlined in Scheme 3.1.⁶²



Scheme 3.1. Synthesis of hit compounds, **1**, **2** and **3**.⁶² Het = pyrimidine, pyridine or naphthyridine.

Reagents and Conditions: (i) K₂CO₃, (Ph₃P)₄Pd, PhMe:MeOH (9:1), 72 h, Δ, 67-91%; (ii) NaOH (2 M), 12 h, Δ, then HCl, 90-100%; (iii) PyBOP, Et₃N, CH=NCH=CHN(CH₂)₃NH₂, CH₂Cl₂, 12 h, RT, 65-83%.

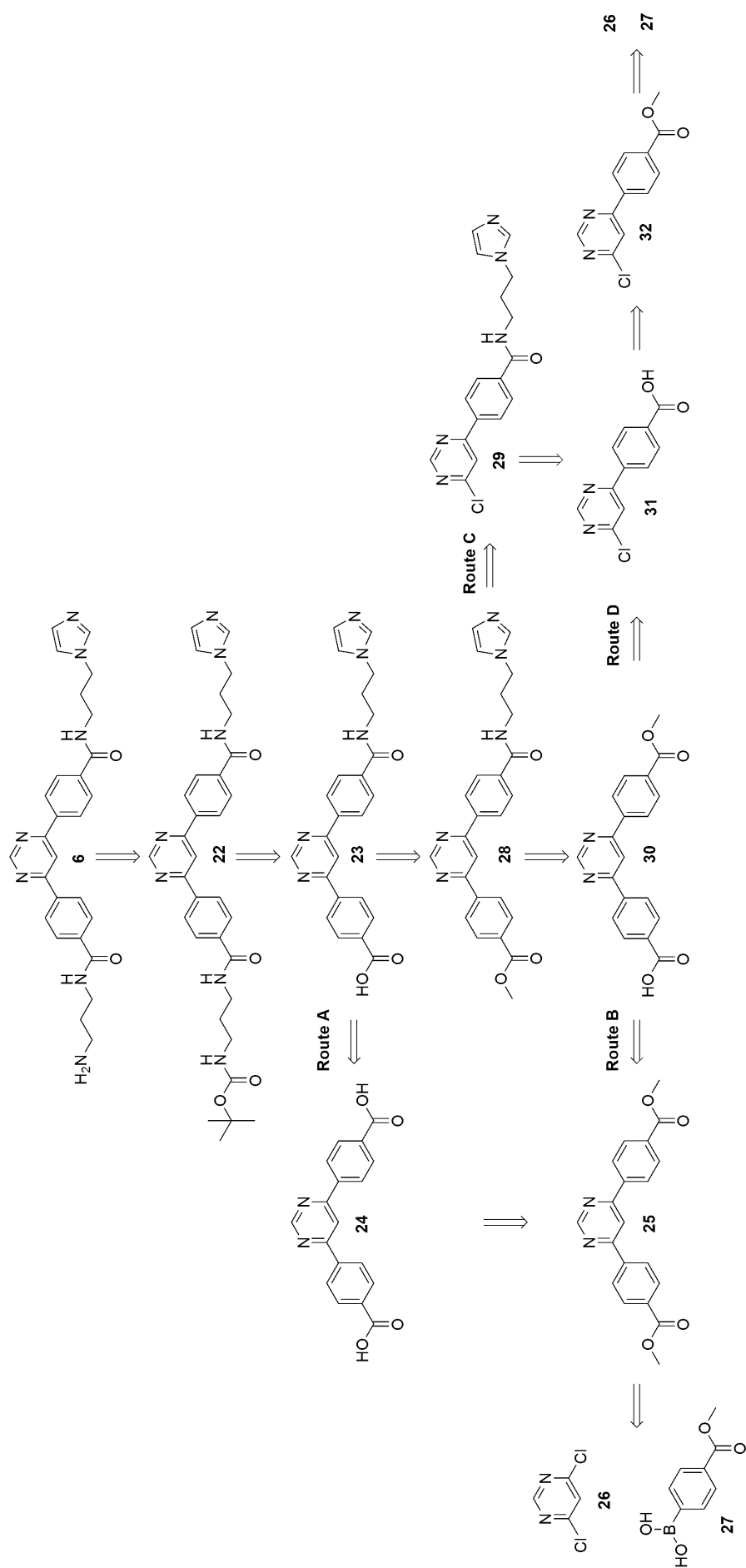
The three steps of the established synthesis are:

- Formation of the biaryl heterocyclic ring system to afford an aryl dimethyl ester;
- Hydrolysis of the esters;
- Formation of amides by attaching the relevant side chains.

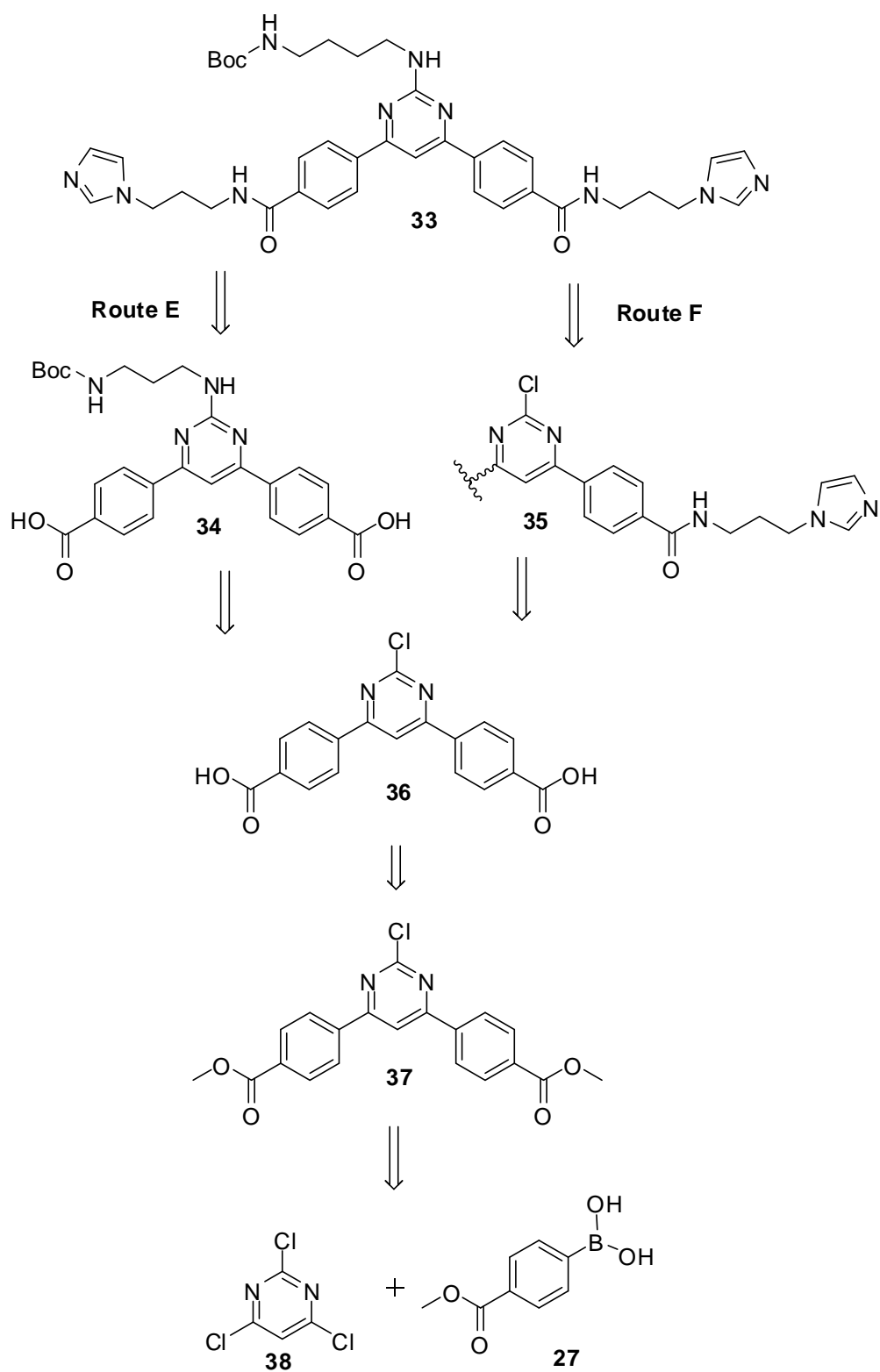
Each of the synthetic steps has been applied successfully in synthesising a wide range of similarly-structured analogues which has generated data contributing to the SAR.^{62, 63, 65, 127} Adapting these established methods would provide access to the target compounds, **6-21**.

Incorporation of the 1,3-diaminopropane linker, for attachment to a solid support or to biotin, could theoretically be achieved at any one of the three steps of the established synthesis. Numerous routes of synthesis are possible following both approaches and many of these have been explored. Scheme 3.2 shows a retrosynthetic analysis of non-symmetrical compound **6**, illustrating the different stages at which the symmetry of the compound can be broken. Route A involves breaking the symmetry during amide formation and Route B breaks the symmetry during hydrolysis. Both Route C and Route D involve breaking the symmetry in the first stage of the synthesis, formation of the biaryl heterocyclic ring system.

Scheme 3.3 shows a retrosynthetic analysis of compound **33**, again illustrating the steps at which the linker could be incorporated.



Scheme 3.2. Retrosynthetic analysis of compound **6**, showing the possible routes to non-symmetrical compounds.



Scheme 3.3. Retrosynthetic Analysis of compound **33** (Routes E and F)

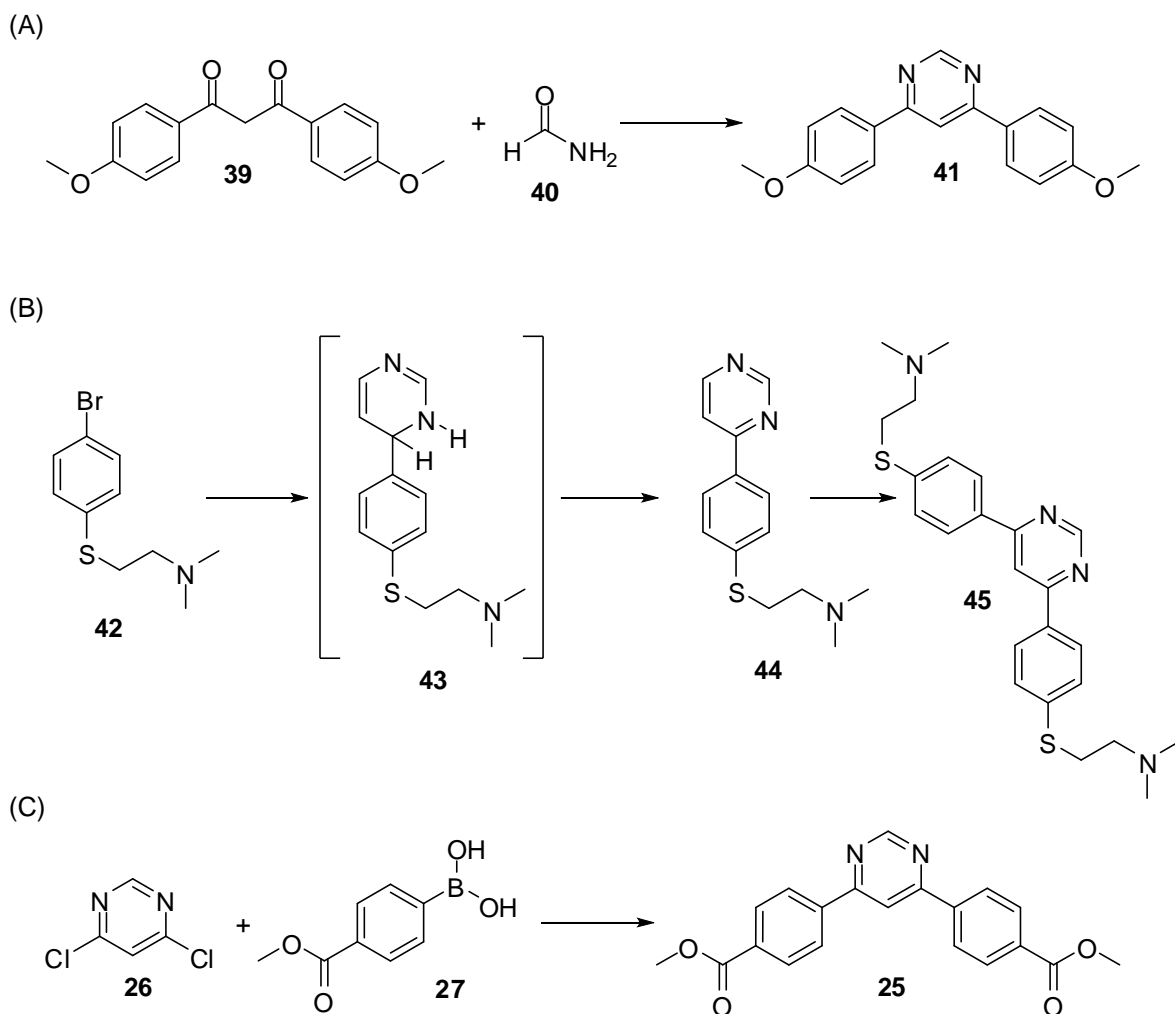
3.2 ESTABLISHED SYNTHESIS

3.2.1 STAGE 1 – FORMATION OF THE BIARYL HETEROCYCLIC RING SYSTEM

Biarylpyrimidines have previously been prepared by aromatic nucleophilic substitution reactions with 1,3-diketones, **39**, and formamide, **40**,¹²⁸ (Scheme 3.4A) and by sequential aryllithium additions to pyrimidine¹²⁹ (Scheme 3.4B). Previous work conducted in this laboratory has favoured Suzuki biaryl cross-coupling reactions to synthesise the biarylheterocyclic ring system.^{62, 63, 65, 127} (Scheme 3.4C). This reaction provides one of the most straightforward methodologies for various carbon–carbon bond formations.¹³⁰

3.2.1.1 The Suzuki Coupling Reaction

The Suzuki coupling reaction is extremely versatile resulting in the formation of a carbon-carbon bond between an aryl or vinyl boronic acid and an aryl, vinyl or alkyl halide through a metal-catalysed reaction.¹³⁰ The reaction was first described by Akira Suzuki's group in 1979.¹³¹ They reported that boranes react readily with 1-alkenyl halides in the presence of a catalytic amount of palladium and a base to give the corresponding conjugated dienes in good yields. More recently the Suzuki coupling reaction has had a huge number of applications, ranging from the synthesis of solar cells¹³² and carbon nanotubes¹³³ to unnatural amino acids¹³⁴ and versatile structural fragments for natural product synthesis.¹³⁵

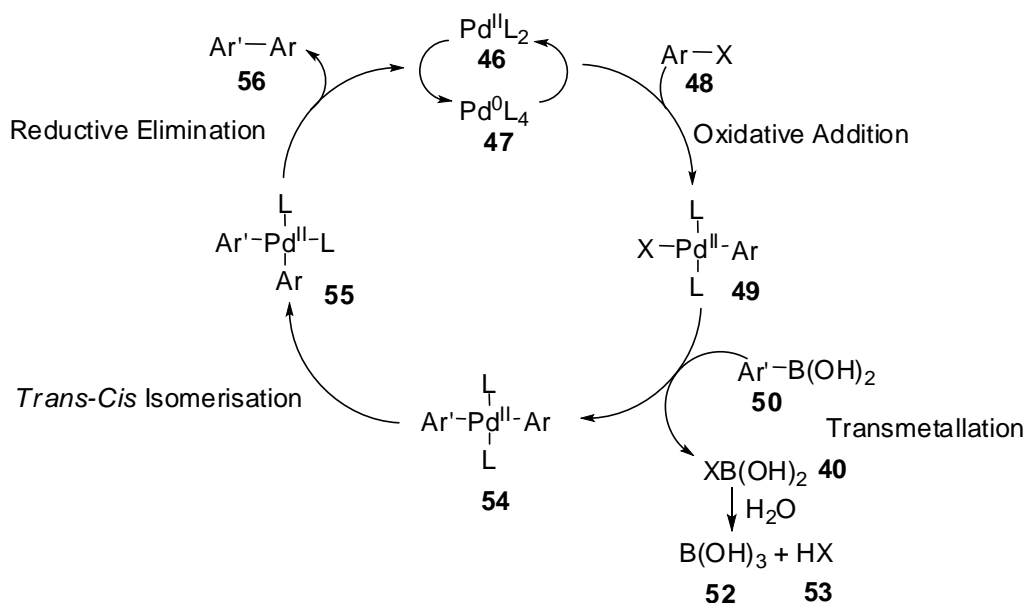


Scheme 3.4. Formation of biarylpyrimidines.

Reagents and Conditions: (A): 180-190 °C, 6 h, 62 %;¹²⁸ (B) *n*-BuLi in hexanes, pyrimidine, DDQ, Et₂O, -50 °C, 1.5 h, 40 %;¹²⁹ (C) K₂CO₃, Pd(Ph₃P)₄, PhMe:MeOH (9:1), Δ, 72 h, 90 %.⁶²

The mechanism of the Suzuki coupling is best viewed from the perspective of the palladium catalyst (Scheme 3.5) and involves oxidative addition, transmetalation, *trans-cis* isomerisation and reductive elimination. The catalyst used is palladium in its zerovalent state. Catalysts can be readily purchased containing palladium of this oxidation state, for example tetrakis(triphenyl phosphine) palladium (Pd(PPh₃)₄). Alternatively the catalyst can be added as a palladium (II) complex along with a ligand, such as palladium acetate with

triphenylphosphine, which reduces readily *in situ* to the active palladium (0) species. Several intermediates of the Suzuki coupling reactions have been isolated and identified by electrospray mass spectrometry and are identified in Scheme 3.5.¹³⁶



Scheme 3.5. Catalytic cycle of the Suzuki Coupling Reaction.¹³⁶

Solid-supported palladium catalysts are often used, which enable them to be recovered and recycled. This has led to suggestions that the Suzuki coupling reaction works *via* a heterogeneous mechanism¹³⁷ rather than the homogeneous mechanism more commonly described.^{138, 139}

Palladium is the most widely used transition metal in homogeneous catalysis and plays a major role in synthesis due to its electronegativity (2.2), which allows the formation of relatively strong Pd-C and Pd-H bonds and also polarised Pd-X bonds.¹⁴⁰ Palladium chemistry is dominated by two oxidation

states, 0 and +II, although oxidation states +I, +III and +IV are possible but seldom mentioned in literature.¹⁴⁰

An ideal catalyst is a complex that is stable in the resting state for storage but it becomes activated in solution by dissociation of a ligand. The most stable transition metal complexes satisfy the 18 electron rule; they have 18 electrons in their valence shells, hence a noble gas configuration. This is achieved by combining the electrons of the metal with those donated by coordinating ligands.¹⁴¹

Triphenylphosphine, as present in $\text{Pd}(\text{PPh}_3)_4$ **57**, is the most widely used coordinating ligand as it is relatively inexpensive and any contamination with phosphine oxide can be readily removed by recrystallisation from ethanol.¹⁴² $\text{Pd}(\text{PPh}_3)_4$ is widely used to catalyse Suzuki reactions. The palladium (0) metal, with 10 valence electrons in its 4d orbital, is complexed to four PPh_3 ligands, each donating 2 electrons from their lone pairs, leading to a palladium (0) four ligand complex, **57**, containing 18 electrons in its valence shell (Figure 3.2).

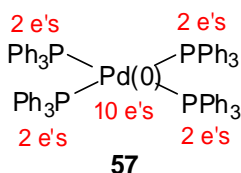
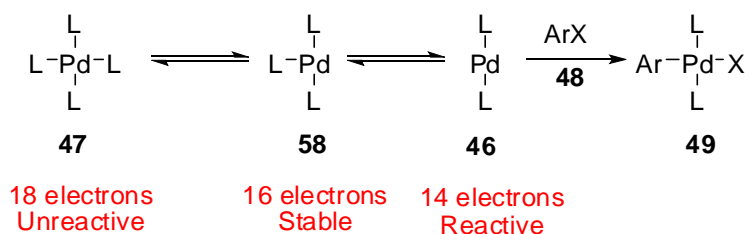


Figure 3.1. $\text{Pd}(\text{PPh}_3)_4$ complex, showing the number of electrons of the palladium (0) metal and its coordinating ligands.¹⁴¹

Each of the triphenylphosphine ligands, in $\text{Pd}(\text{PPh}_3)_4$, has a lone pair of electrons on the phosphine in a filled sp^3 type orbital. This overlaps with a vacant d_{sp} orbital of palladium. The result is the formation of a conventional 2 electron σ -bond, leading to an increase in the electron density of the palladium atom. Ligands without a lone pair of electrons or a filled π -orbital interact with transition metals by breaking a σ -bond. This is known as oxidative addition as the palladium oxidation state is raised by two. This is the result of adding two extra ligands bearing a formal negative charge.¹⁴¹

As a solid, $\text{Pd}(\text{PPh}_3)_4$, **57**, is an unreactive, stable complex, however, in solution, $\text{Pd}(\text{PPh}_3)_4$ loses two ligands affording a 14 electron species, **46**, which is very reactive. The oxidative addition stage of the Suzuki reaction is thought to take place on this coordinatively unsaturated species¹⁴¹ (Scheme 3.6).



Scheme 3.6. Ligands dissociation from the $\text{Pd}(0)(\text{PPh}_3)_4$ complex, leading to oxidative addition to the aryl halide.¹⁴¹ (L = PPh_3)

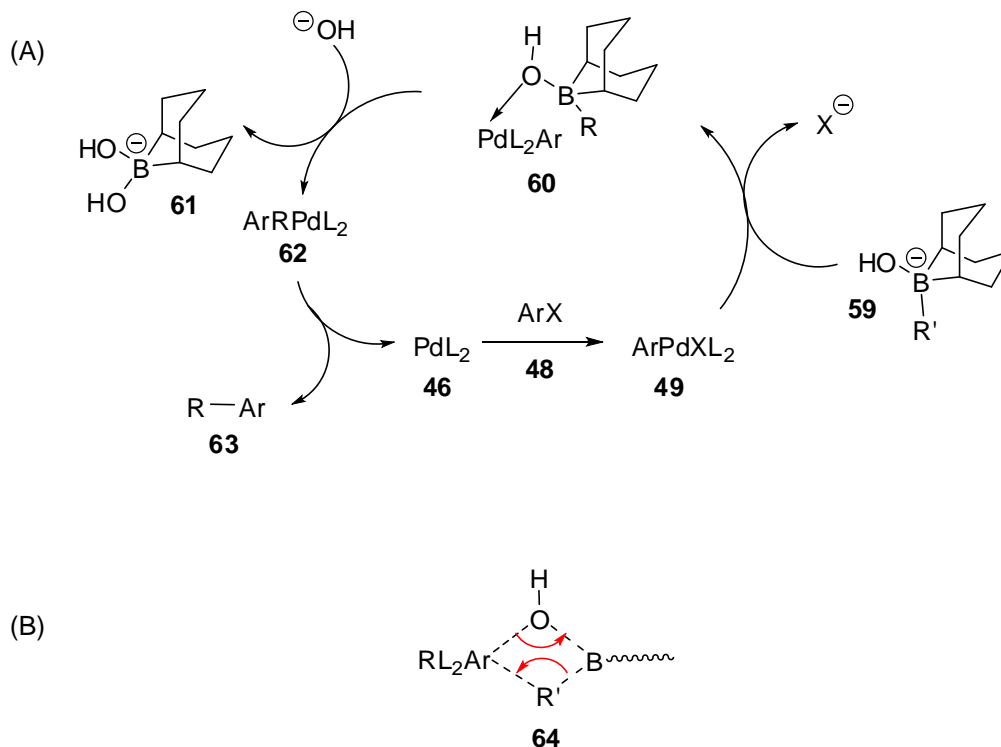
Oxidative addition of the electrophilic aryl halide, **48**, to a palladium (0) complex, **46**, affords a stable palladium (II) complex, **49**.¹³⁰ The palladium (0) complex has vacant coordination sites and acts as a nucleophile towards the aryl halide, **48**. This palladium complex undergoes oxidative addition via concerted insertion

into the aryl halide σ -bond.¹⁴³ The result is a σ -bond between the halide and the metal and a σ -bond between the metal and the aryl group. As the electrons in these bonds are shared between the metal and the newly complexed ligands, the palladium is reduced to a Pd(II)L₂RX, **49**, complex where the palladium has 16 electrons.

This is often the rate determining step in the catalytic cycle. Aryl halides with electron-withdrawing groups are much more reactive towards the oxidative addition stage than those bearing electron donating groups.¹³⁰ Alternatively the rate-determining step can be the transmetalation stage, this is dependent on the reaction conditions. A larger halide results in a slower transmetalation due to steric hindrance.¹⁴⁴

Following oxidative addition, transmetalation of an aryl group from boron, **50**, to palladium, **49**, affords a diaryl organopalladium species, **54** (Scheme 3.5). Transmetalation involves the exchange of ligands between two metal centres. In the Suzuki reaction exchange of the halide within the palladium complex, **49**, and the aryl group from the boronic acid occurs, **50**. Organoboranes do not undergo cross-coupling without the addition of a base due to the low nucleophilicity of the borane reagents.¹⁴⁵ Matos *et al.* have carried out detailed mechanistic studies of the role of a hydroxide base in the coupling reaction of trialkyl borates and aryl halides and have suggested a more detailed mechanism than is normally depicted, incorporating the role of base¹⁴⁵ (Scheme 3.7). The organoborane compounds are predominantly present as their hydroxyborate complexes, **59**, and it is this complex which reacts with **49** to

form a hydroxo-bridged intermediate, **60**, which facilitates the transmetalation through transition state **64**.



Scheme 3.7. (A) Suzuki coupling reaction incorporating the role of a base; (B) Transmetalation transition state. Adapted from Matos *et al.*¹⁴⁵

During the oxidative addition stage, a *trans* isomer of the palladium (II) square planar complex, **49**, is formed, through the addition of the palladium (0) catalyst, **47**, to the aryl halide, **48**. This isomeric form is retained through the transmetalation stage, **54**. For the reaction to proceed to the reductive elimination, *trans-cis* isomerism is required to give **55**. The mechanism for this stage is uncertain, however it has been suggested that isomerisation occurs by at least four concurrent bimolecular pathways, two of which are autocatalytic and the other two are solvent assisted.¹⁴³

Finally reductive elimination of the complex, **55**, eliminates the aryl partners regenerating the reactive palladium (0) catalyst, **47**, which can re-enter the catalytic cycle (Scheme 3.5).¹⁴⁶ Reductive elimination is driven by the close proximity of the two aryl groups in the *cis* coordination sites of the metal. The aryl-aryl bond is stronger than the coordination with the metal, so the two aryl groups form a new carbon-carbon bond and are released from their coordination through a concerted reaction. This results in concurrent reduction of the palladium (II) complex, **46**, and reassociation of the PPh₃ ligands to the palladium (0), **47**.

3.2.1.1.1 Variants on the Suzuki Coupling Reaction

There are several palladium-catalysed cross coupling reactions described for the formation of carbon-carbon bonds, summarised in Table 3.1. All of the couplings follow the same mechanism as that for the Suzuki reaction, oxidative addition, transmetalation, *trans-cis* isomerisation and reductive elimination.

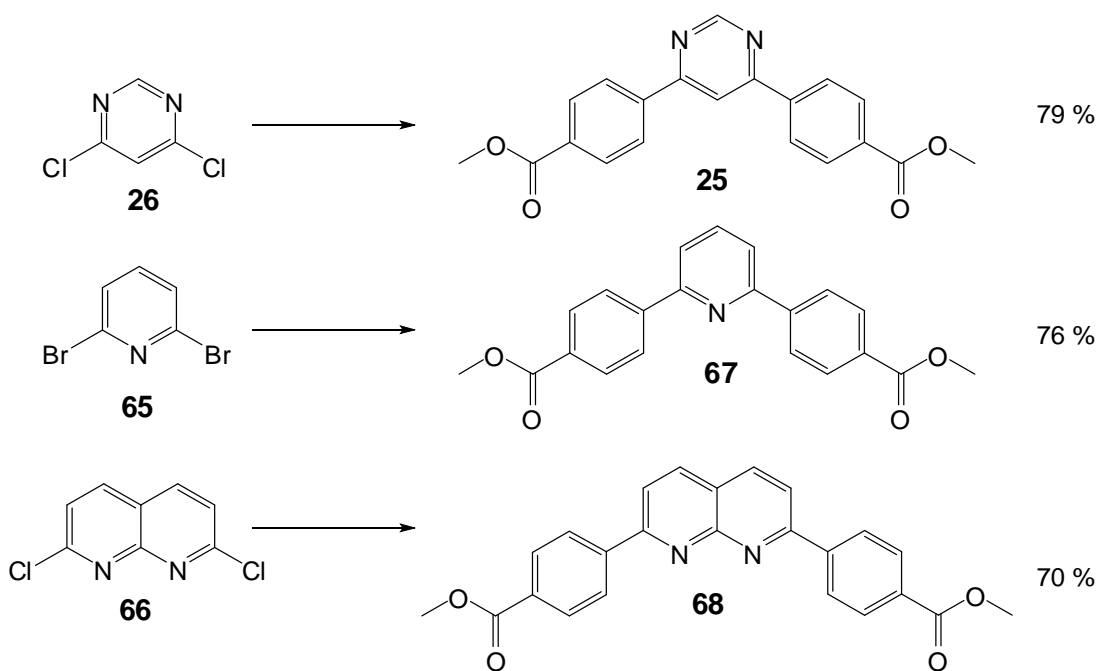
Table 3.1. Alternative palladium-catalysed cross coupling reactions.

Reaction	Scheme	Halide	Nucleophile
Heck ¹⁴⁷	$\text{R-X} + \text{R}'\text{-CH=CH}_2 \xrightarrow[\text{Pd}^0]{\text{Base}} \text{R}'\text{-CH=CH-R}$	sp ²	sp ²
Stille ¹⁴⁸	$\text{R-X} + \text{R}'\text{-Sn(R')}_3 \xrightarrow{\text{Pd}^0} \text{R-R}'$	sp ² , sp ³	sp, sp ² , sp ³
Hiyama ¹⁴⁹	$\text{R-X} + \text{R}'\text{-Si(R')}_3 \xrightarrow[\text{Pd}^0]{\text{F}^- \text{ or Base}} \text{R-R}'$	sp ² , sp ³	sp ²
Sonogashira ¹⁵⁰	$\text{R-X} + \text{R}'\text{-C}\equiv\text{C-H} \xrightarrow[\text{Pd}^0, \text{Cu}^+]{\text{Base}} \text{R}'\text{-C}\equiv\text{C-R}$	sp ² , sp ³	sp
Negishi ¹⁵¹	$\text{R-X} + \text{R}'\text{-Zn-X}' \xrightarrow{\text{Pd}^0} \text{R-R}'$	sp ² , sp ³	sp, sp ² , sp ³

3.2.1.1.2 Suzuki Reactions Carried Out in This Project

Suzuki reactions have been used in this project to synthesise the biaryl heterocyclic ring systems by coupling 4-methoxycarbonyl boronic acid, **27**, with either 4,6-dichloropyrimidine, **26**, 2,6-dibromopyridine, **65**, or 2,7-dichloronaphthyridine, **66**, yielding the respective products in high purity (Scheme 3.8). 4,6-Dichloropyrimidine and 2,6-dibromopyridine were available

commercially and the 2,7-dichloronaphthyridine, **66**, was synthesised following the method from Newkome *et al.*¹⁵² Each of the reactions was carried out in a solvent system of toluene and methanol (9:1), to allow for a high reflux temperature and to increase the solubility of the potassium carbonate base. As the palladium catalyst used ($\text{Pd}(\text{PPh}_3)_4$) is readily oxidised with air contact, the reactions were carried out under an inert atmosphere and the solvents were thoroughly degassed with either nitrogen or argon prior to use. Degassing with argon resulted in higher product yield due to the greater density compared with nitrogen.

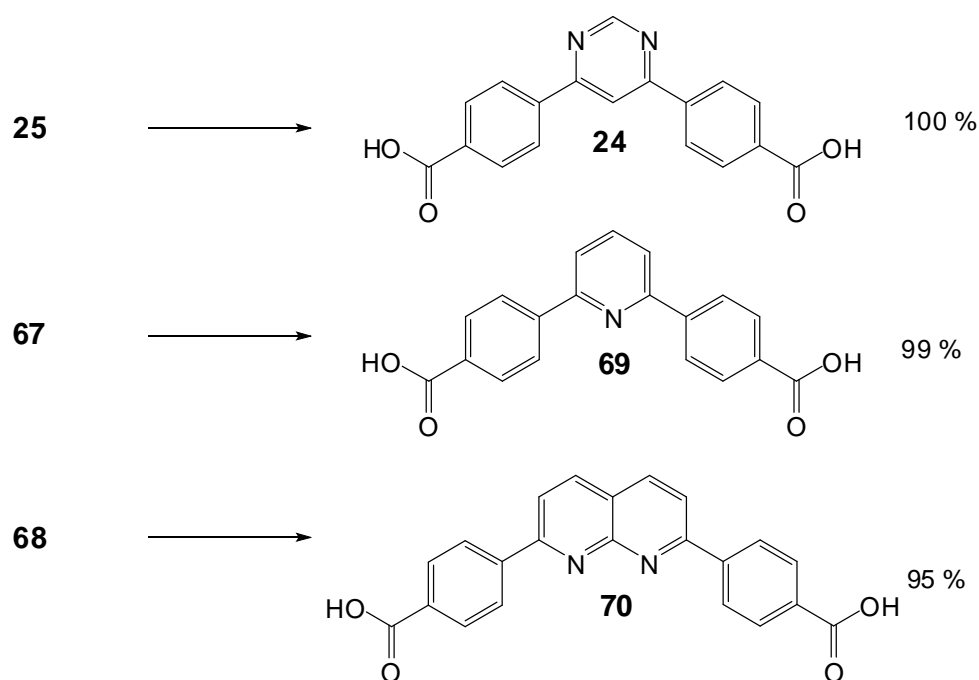


Scheme 3.8. Suzuki reactions of symmetrical *bis*-4-methoxycarbonylheterocycles.

Reagents and Conditions: **27**, K_2CO_3 , $\text{Pd}(\text{PPh}_3)_4$, $\text{PhMe}:\text{MeOH}$ (9:1), Δ , 48 h.

3.2.2 STAGE 2 – HYDROLYSIS OF THE METHYL ESTERS

Free acids of the symmetrical diesters, **25**, **67** and **68**, were prepared by basic hydrolysis of the methyl esters. The compounds were heated under reflux in 2 M sodium hydroxide solution and then acidified with 4 M hydrochloric acid, affording products in up to 100 % yields (Scheme 3.9).



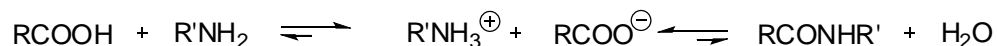
Scheme 3.9. Basic hydrolysis of symmetrical *bis*-4-methoxycarbonylheterocycles.

Reagents and Conditions: NaOH (2 M), Δ , 24 h, then HCl (4 M).

3.2.3 STAGE 3 – FORMATION OF AMIDE BONDS

Amide condensations are equilibrium reactions. Direct reaction of amines and carboxylic acids results in an acid base reaction, leading to the formation of

stable salts (Scheme 3.10) which reduces the possibility of nucleophilic attack of the amine, although the reaction can proceed at very high temperatures.¹⁵³



Scheme 3.10. Reaction of carboxylic acids with amines

Therefore, activation of the carboxylic acid is required, which converts the carboxylic acid into a reactive acylating agent, bearing a good leaving group, which can undergo a nucleophilic acylation reaction with the amine affording an amide. Carboxylic acids can be activated as acyl halides, acyl azides, anhydrides and esters. Activated esters can be formed through the reaction with alcohols and phenols containing an electron-withdrawing substituent, leading to an increased electrophilicity at the carbonyl centre. Examples of alcohols and phenols used to form activated esters are shown in Figure 3.3.

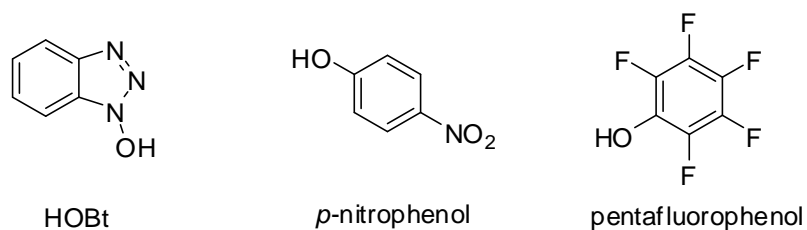
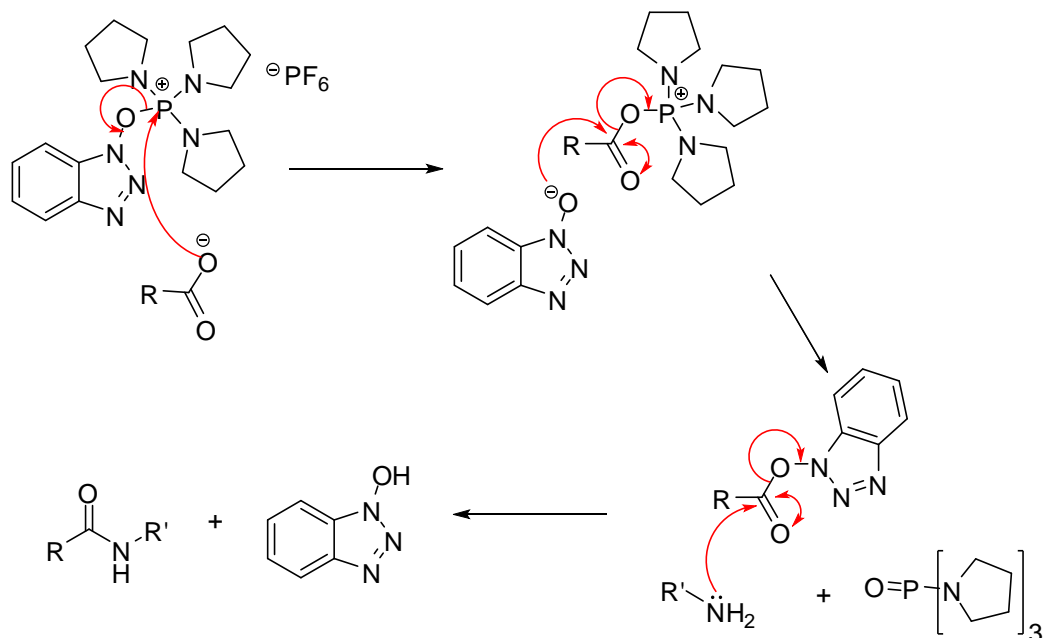


Figure 3.2. Alcohols and phenols used in the synthesis of activated esters.

Catalysts incorporating an electron withdrawing alcohol, can generate activated esters *in situ*. Benzotriazole-1-yl-oxytrispyrrolidinophosphoniumhexafluorophosphate (PyBOP) contains HOBt and can activate a carboxylic acid providing

a good leaving group for nucleophilic acylating reactions with amines (Scheme 3.11).



Scheme 3.11. PyBOP mediated coupling.¹⁵⁴

PyBOP coupling is used widely for solid phase peptide synthesis and has been extensively used for the synthesis of compounds similar to the target compounds in high yields.^{62, 63} For reactions carried out in this project, the carboxylic acid was converted to the respective carboxylate anion with triethylamine prior to coupling, which also increased solubility in the dry DCM solvent. Completion of the reaction resulted in precipitation of the amides which were collected by filtration.

3.3 SYNTHESIS OF NON-SYMMETRICAL COMPOUNDS

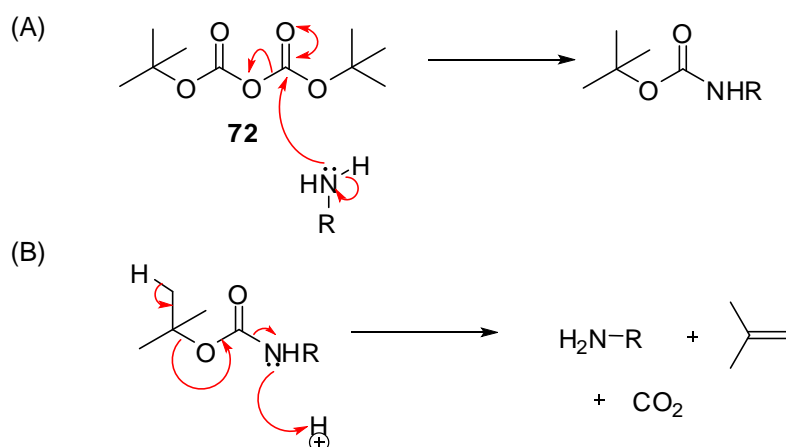
Non-symmetrical compounds can be produced using the routes shown in Scheme 3.2. Many of these routes have been explored with varying success. Prior to synthesis, the Boc-protected attachment linker **71** was synthesised.

3.3.1 SYNTHESIS OF THE ATTACHMENT LINKER, **71**

Although, 1,3-diaminopropane, **5**, provides an attractive linker for joining the hit compounds to biotin or a solid support, as it is a bifunctional compound, reactions could lead to the formation of polymers. Therefore, one of the amine groups was protected before use in synthesis.

A Boc group provides good protecting group for one of the amines, as it is stable in the basic conditions required for each step of the established synthesis (Scheme 3.1), and can be easily removed with weak acid. *N*-Boc-1,3-diaminopropane, **71**, has been synthesised previously in very high yields (93 %) from di-*tert*-butyldicarbonate, **72**, and the unprotected diamine, **5**, in THF.¹⁵⁵ This synthesis has been repeated and produced high yields of product with only minor diprotected impurities, which were easily removed by distillation (160 °C, 0.1 mm Hg).

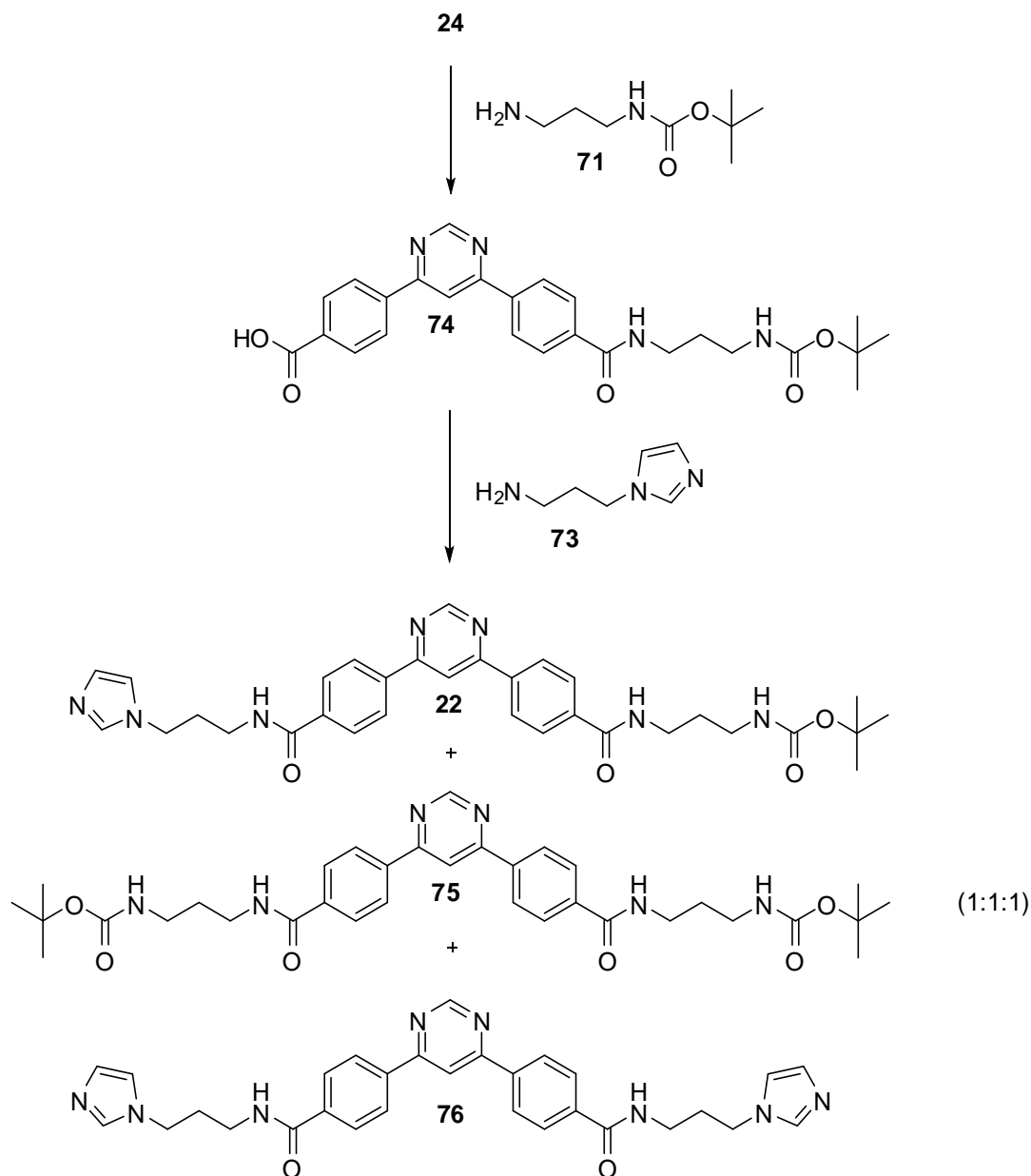
The Boc protecting group is introduced by reaction of the amine with di-*tert*-butyl dicarbonate by a nucleophilic acyl substitution reaction and later removed by acidolysis (Scheme 3.12).



Scheme 3.12. Reaction mechanism of (A) Boc protection of a primary amine; (B) removal of a Boc protecting group using acid.

3.3.2 ROUTE A – BREAKING SYMMETRY DURING AMIDE FORMATION

Diacid **24** was synthesised using the established synthesis (Section, 3.2.1.1.2 and 3.2.2). To break the symmetry during amide formation, a ‘one-pot’ reaction approach was attempted to produce **22**. One equivalent of PyBOP, Et₃N and *N*-Boc-1,3-diaminopropane, **71**, were added and left to stir for 18 h. This was followed by the addition a further equivalent of PyBOP and Et₃N with 1-(3-aminopropyl)imidazole, **73** (Scheme 3.13). It was speculated that amide formation to one of the carboxylic acid groups would affect the pK_a of the second acid due to the conjugation between the functional groups. Therefore, making initial reaction on another diacid molecule, **24** more favourable than a second reaction on a mono-amide **74**.



Scheme 3.13. One-pot synthesis of non-symmetrical amides.

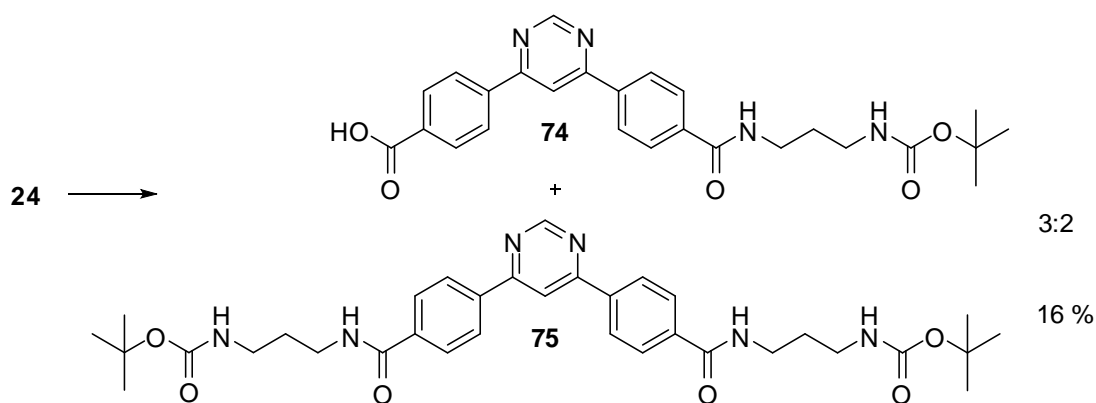
Reagents and Conditions: PyBOP, 1 eq. Et_3N , dry DCM, RT, 18 h.

The Boc protected side chain was added first as higher selectivity in coupling was observed. NMR data for the final product showed signs of the non-symmetrical compound, **22**, but the mass spectrum demonstrated that both possible dimers, **75** and **76**, and the non-symmetrical product, **22**, were present.

Attempts to separate the three compounds were made using flash chromatography, but the *R_f* values were too similar for separation. The solubility of all three compounds in various solvents was investigated, in the hope that this could lead to separation. However, all three compounds exhibited similar solubility in a wide range of solvents. Final attempts to separate the compounds involved the use of reverse phase HPLC, using a mobile phase of acetonitrile and water. The solute gradient was adjusted to try to achieve maximum separation, however, all attempts to separate the compounds failed.

The unsuccessful separation of the non-symmetrical compound, **22**, from the one-pot reaction mixture (Scheme 3.13) led to attempts to isolate the 4-(4-carboxyphenyl)-6-(4-[3-*N*-Boc-aminopropyl]carboxyamidophenyl) pyrimidine, **74**, prior to a second amide coupling with a different side chain, **73** (Scheme 3.14). Reaction was carried out using an excess of diacid, **24** (10 eq.). Unreacted starting material was recovered by trituration with hot methanol, leaving the product in solution.

The preparation appeared successful on observation of the crude ¹H NMR spectrum. Separation of the crude reaction mixture using flash chromatography, resulted in isolation of the side products HOBt, tris-(1-pyrrolidiny)-phosphine oxide, Et₃N and the disubstituted product, **75**. However, the mono-substituted product, **74**, failed to be collected in pure form due to very poor yields. However, mass spectra and ¹H NMR indicated the presence of the desired compound.



Scheme 3.14. Synthesis of 4-(4-carboxyphenyl)-6-(4-[3-*N*-Boc-aminopropyl]carboxamido phenyl)pyrimidine, **74**.

Reagents and Conditions: **71**, Et_3N , PyBOP, dry DCM, RT, 18 h.

3.3.3 ROUTE B – BREAKING SYMMETRY DURING ESTER HYDROLYSIS

Controlled partial hydrolysis of dimethyl esters, **25**, **67**, **68** was attempted using different bases and solvent systems. It was predicted that the use of a weak base would result in hydrolysis of one of the methyl esters.

Partial hydrolysis of 2,6-*bis*-(4-methoxycarbonylphenyl) pyridine, **67**, was tried using the different bases and reaction conditions summarised in Table 3.2. The reactions were monitored by thin layer chromatography.

The monohydrolysis of 2,6-*bis*-(4-methoxycarbonylphenyl) pyridine, **67**, appeared to successful, using either 3 eq. of potassium carbonate as a base or 30 eq. of sodium hydroxide. The ^1H NMR spectrum (Figure 3.4) showed a

methyl peak of three hydrogens and non-equivalence of the two aromatic phenyl rings, which are seen a multiplet in the diester, **67** and two doublets in the diacid, **69**. However, mass spectra showed these compounds to be mixtures of the starting material, **67**, and the diacid, **69**, in equimolar quantities.

Table 3.2. Reagents and conditions for the attempted partial hydrolysis of 2,6-bis-(4-methoxycarbonylphenyl) pyridine, **67**.

Base	Equivalents	Solvent	Temperature	Time	Result
NaOH	1.5	H ₂ O	50°C	18 h	Total recovery of starting material
NaOH	30	H ₂ O:MeOH (9:1)	50°C	24 h	Mixture of the diacid and diester (1:1)
NaOH	5	H ₂ O:MeOH (50:50)	RT	18 h	Total recovery of starting material
KOH	4	H ₂ O:MeOH (9:1)	50°C	24 h	Diacid
K ₂ CO ₃	3	H ₂ O:MeOH (50:50)	40°C	18 h	Mixture of the diacid and diester (1:1)

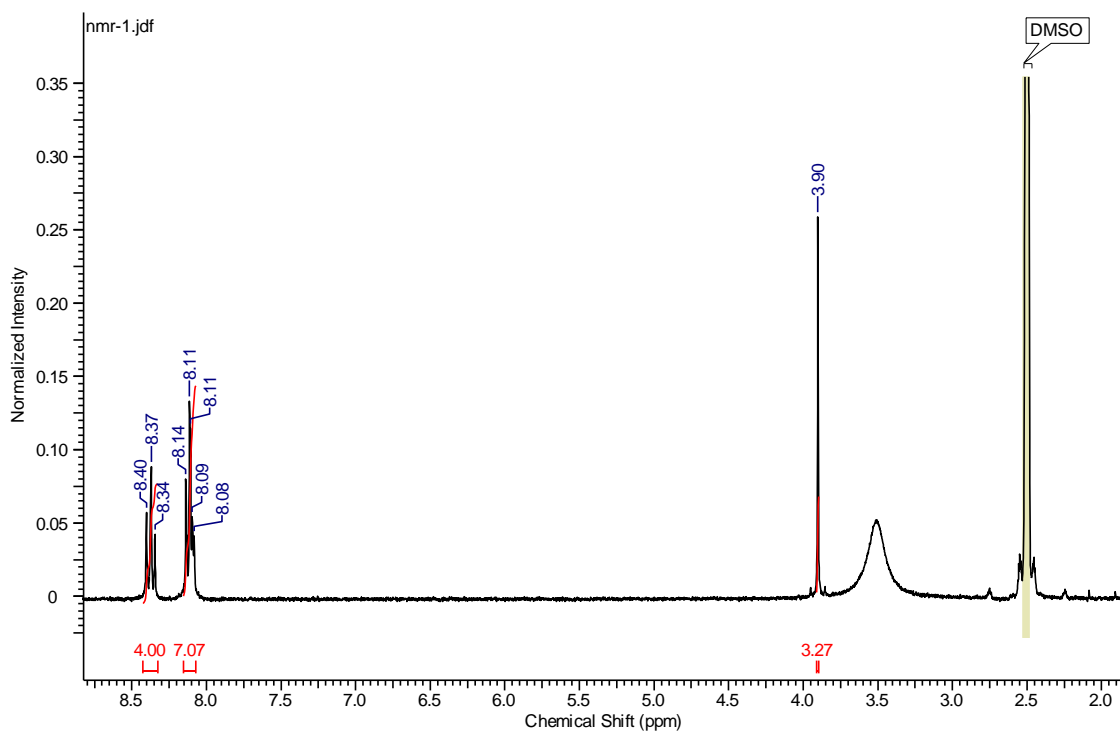


Figure 3.3. ^1H NMR spectrum in DMSO of the hydrolysis of **67** using K_2CO_3 as a base.

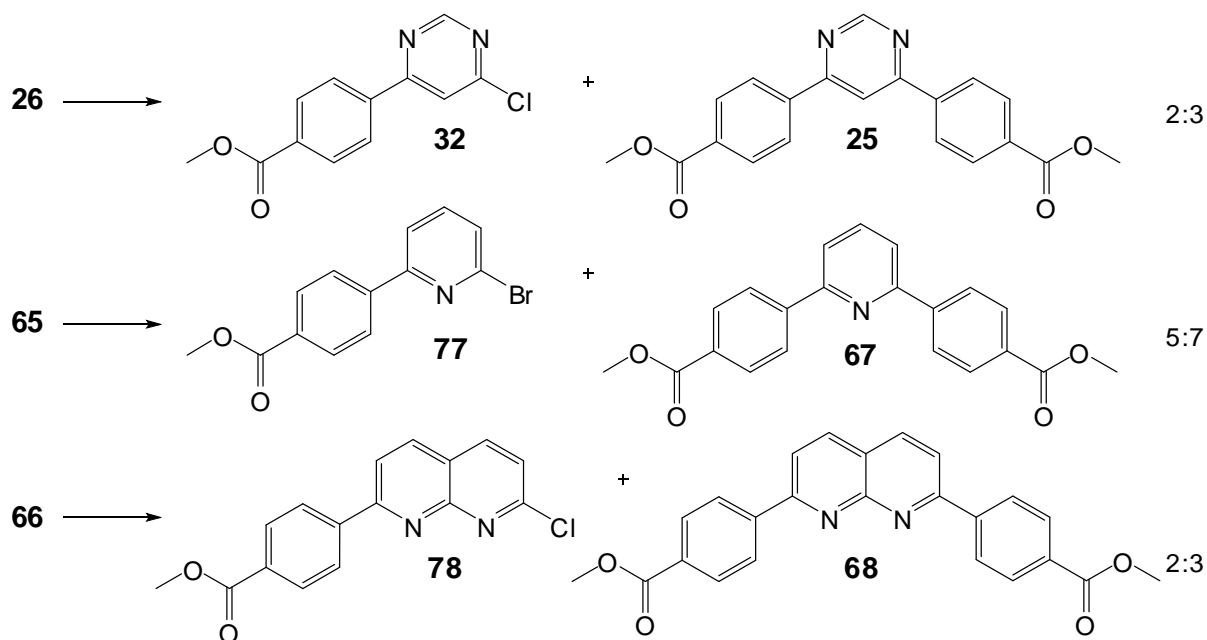
3.3.4 ROUTE C AND ROUTE D – BREAKING SYMMETRY DURING FORMATION OF THE BIARYLHETEROCYCLIC RING SYSTEM

The previous two routes to synthesis of non-symmetrical compounds encountered problems. Therefore, breaking the symmetry during the formation of the heterocyclic ring system was explored. This route involved coupling boronic acid **27** with one equivalent of dihalide **26** followed by hydrolysis of the methyl ester. Two routes are possible following the hydrolysis, attachment of the first side chain, amide **73**, prior to a second Suzuki coupling (Route C) or the second Suzuki coupling of **27** to monoacid **31** immediately (Route D).

3.3.4.1 Sequential Suzuki Coupling

Routes C and D for the synthesis of target compounds **6-13**, required sequential Suzuki couplings to be carried out, using heterocycles bearing more than one halide atom, to afford non-symmetrical products.

Following the established conditions described for symmetrical compounds, **25**, **67** and **68** (Section 3.2.1.1.2) using one equivalent of 4-methoxyphenyl boronic acid, **27**, with dichloropyrimidine was unsuccessful. The reactions afforded mixtures of the starting material, monosubstituted, **32**, and the disubstituted, **25**, products (Scheme 3.15). The solvent volume and temperature were reduced in order to force precipitation of the product following the first coupling (Table 3.3). However, these conditions also yielded mixtures of products.



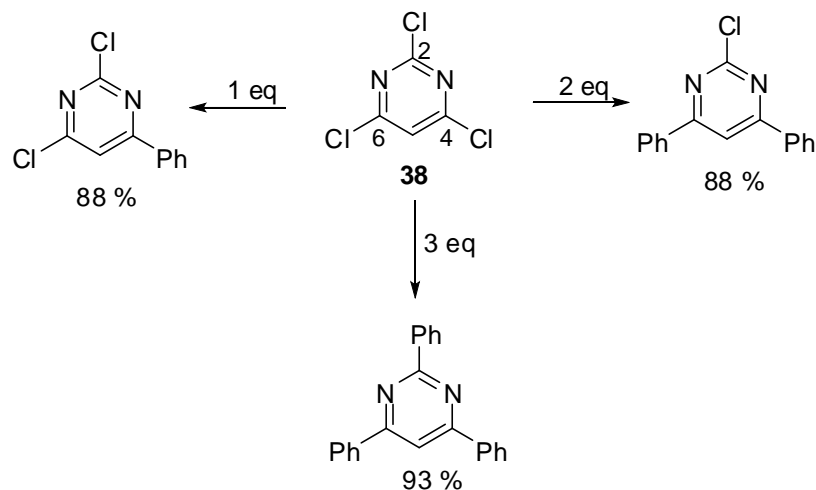
Scheme 3.15. Suzuki reactions of symmetrical *bis*-4-methoxycarbonyl heterocycles.

Reagents and Conditions: **27**, K_2CO_3 , $Pd(PPh_3)_4$, $PhMe:MeOH$ (9:1), Δ , 48 h.

Table 3.3. Conditions for attempted sequential Suzuki coupling of 4,6-dichloropyrimidine, **26**, and 4-methoxycarbonylphenyl boronic acid, **27**. ^aEstablished conditions.

Solvent Volume	Temperature	Result
^a 200 ml	^a 120 °C	Mixture of mono and disubstituted ester (2:3)
150 ml	90 °C	Mixture of mono and disubstituted ester (2:3)
100 ml	90 °C	Mixture of mono and disubstituted ester (1:1)
75 ml	90 °C	Mixture of mono and disubstituted ester (8:7)

The selectivity of arylation of 2,4,6-trichloropyrimidine, **38**, via Suzuki coupling reactions has been reported by Schomaker *et al.*¹⁵⁶ Substitution with one, two or three equivalents of an arylboronic acid led to the production of the mono-, di- and tri-substituted products respectively in very high yields (Scheme 3.16). This is rationalised as the consequence of regio-selectivity of the substrate coordinating with the catalyst and indicates an apparent order of reactivity, specifically position 4 > position 6 > position 2.



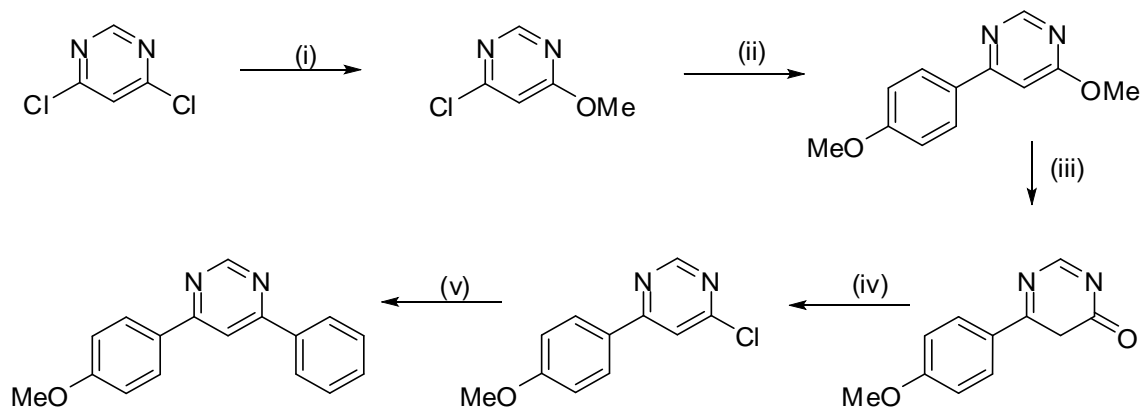
Scheme 3.16. Sequential coupling of biarylpyrimidines.¹⁵⁶

Reagents and Conditions: $PhB(OH)_2$, Na_2CO_3 (aq), $Pd(OAc)_2$, PPh_3 , glyme, Δ , 18-24 h.

The reaction conditions vary from those described in Section 3.2.1.1.2 for the disubstitution of the heterocyclic dihalides, in respect of the base (aqueous Na_2CO_3 instead of K_2CO_3), solvent (glyme instead of PhMe:MeOH) and catalyst ($Pd(OAc)_2$ and PPh_3 instead of $Pd(PPh_3)_4$). The same group also investigated the coupling of substituted phenylboronic acids and trichloropyrimidines. The reactions proved very successful, selectively providing high yields of mono-, di- and tri-substituted products.¹⁵⁷

Synthesis of substituted non-symmetrical 4,6-dichloropyrimidines has been investigated by Zhou *et al.*¹⁵⁸ The procedure involves nucleophilic displacement of one of the chlorine atoms for a methoxy group, allowing conventional Suzuki oxidative addition to the remaining carbon-chlorine bond of the heterocycle. The methoxy group can then be converted back to the chloride by the treatment with HBr-AcOH and $POCl_3$. A second Suzuki coupling can introduce a different

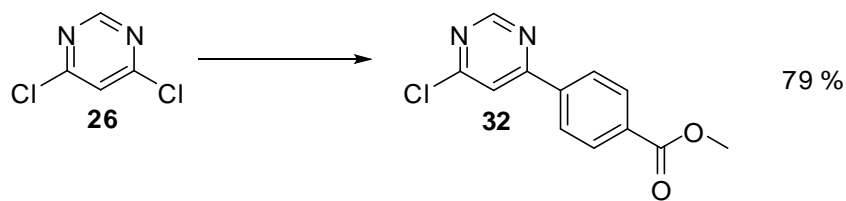
substituted phenyl boronic acid yielding non-symmetrical compounds. (Scheme 3.17).



Scheme 3.17. Synthesis of non-symmetrical 4,6-diaryl pyrimidines.¹⁵⁸

Reagents and Conditions: (i) 1 eq. NaOMe (25% w/w in MeOH), MeOH, RT, 10 min, 95%; (ii) 1.4 eq. MeOC₆H₄B(OH)₂, 0.1 eq. Pd(PPh₃)₄, 2 eq. Na₂CO₃, PhMe, 90 °C, 12 h; (iii) HBr-AcOH (1:3), 80 °C, then satd NaHCO₃; (iv) POCl₃, 30 min, 100 °C, then satd NaHCO₃, 90%; (v) 1.4 eq. PhB(OH)₂, 0.1 eq. Pd(PPh₃)₄, 2 eq. Na₂CO₃, PhMe, 90 °C, 12 h, 80%.

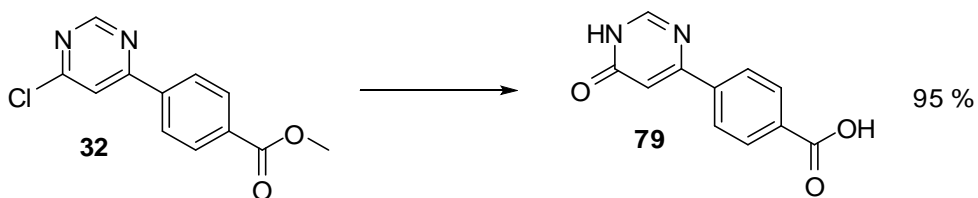
Reactions were carried out following Delia's method as a one step synthesis was more attractive than Zhou's multi-step process. 4,6-Dichloropyrimidine, **26**, and one equivalent of 4-methoxycarbonylphenyl boronic acid, **27**, were reacted following the conditions defined in Scheme 3.16. This provided relatively high yields of product **32** (Scheme 3.18).



Scheme 3.18. Selective arylation of dichloropyrimidine

Reagents and Conditions: **27**, aq. Na_2CO_3 , $\text{Pd}(\text{OAc})_2$, PPh_3 , glyme, Δ , 18-24 h.

Subsequent ester hydrolysis of the non-symmetrical compound **32** to afford the carboxylic acid **31**, prior to a second Suzuki coupling with **27**, was carried out using the reaction conditions previously described (Section 3.2.2). Mass spectra results revealed substitution of the chloride along with the hydrolysis of the methyl esters, **79** (Scheme 3.19).

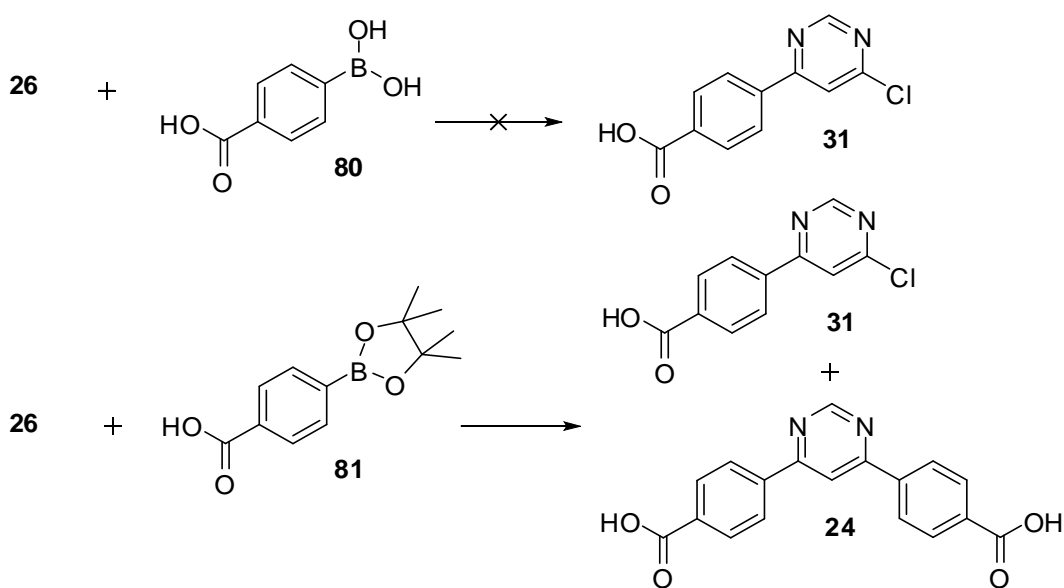


Scheme 3.19. Hydrolysis of non-symmetrical ester **32**.

Reagents and Conditions: NaOH (2 M), Δ , 18 h, then HCl (4 M).

Sequential Suzuki coupling reactions were performed using 4-carboxyphenyl boronic acid, **80**, and 4,6-dichloropyrimidine, **26**, to synthesise the carboxylic acid **31** directly. The reactions were unsuccessful possibly due to the boronic acid insolubility in the glyme:water solvent system.

To overcome the solubility problems, the couplings were carried out using 4-carboxyphenyl boronic acid pinacol ester, **81**, which is soluble in the solvent system required for sequential couplings. The desired mono- and di-phenylcarboxylic compounds were recovered; however they were in equimolar quantities suggesting there was no selectivity.

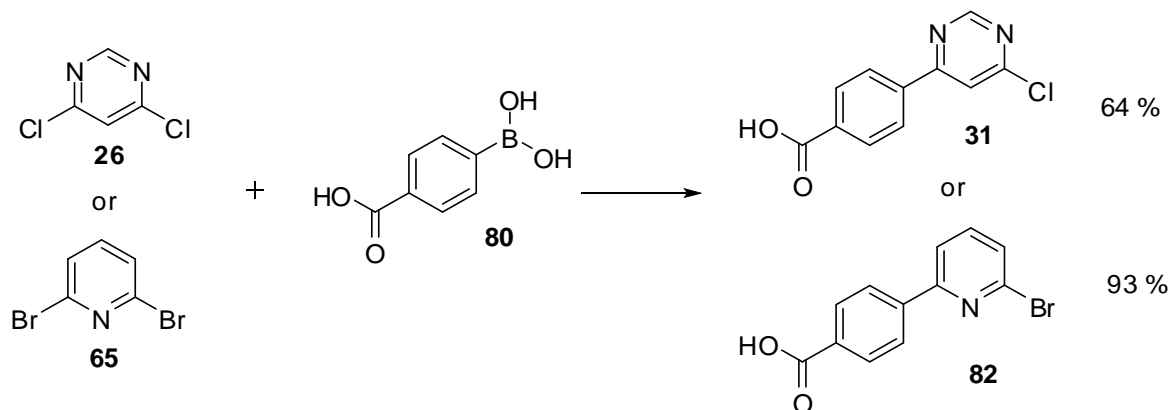


Scheme 3.20. Attempted synthesis of 4-chloro-6-(4-carboxyphenyl) pyrimidine, **31**.

Reagents and Conditions: aq. Na_2CO_3 , $\text{Pd}(\text{OAc})_2$, PPh_3 , glyme, Δ , 18-24 h.

Gong *et al.*¹⁵⁹ have synthesised 4-chloro-6-(4-carboxyphenyl) pyrimidine, **31**, from 4,6-dichloropyrimidine, **26**, and 4-carboxyphenyl boronic acid, **80**, via Suzuki couplings. The synthesis uses $\text{Pd}(\text{PPh}_3)_4$ as a catalyst, Na_2CO_3 as a base and a solvent system of 50 % aqueous acetonitrile (Scheme 3.21).¹⁵⁹ Reactions carried out following this method showed increased selectivity with only minor amounts of the dicarboxylic acid compound. Purification was achieved by washing the crude compound through a plug of silica with a 10 %

methanol in chloroform solution. The same method was used for the synthesis of 2-bromo-6-(4-carboxyphenyl) pyridine, **82**.

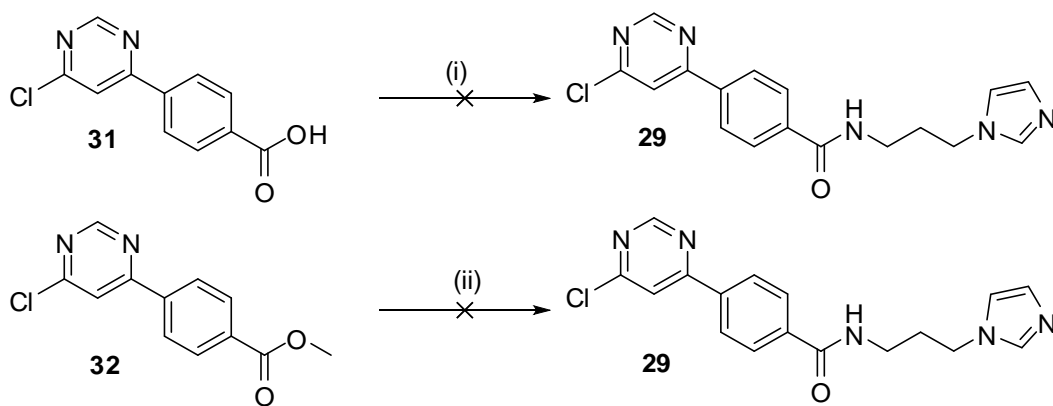


Scheme 3.21. Synthesis of 4-chloro-6-(4-carboxyphenyl) pyrimidine, **31**, and 2-bromo-6-(4-carboxyphenyl) pyridine, **82**.

Reagents and Conditions: Na_2CO_3 (0.4M):MeCN (1:1), 0.1 eq. $\text{Pd}(\text{PPh}_3)_4$, 90 °C, Ar, 24 h.

3.3.4.2 Route C

Route C (Scheme 3.22) involved the addition of the side chain prior to a second Suzuki coupling. Attempts at amide formation were made using **31** and **73** following the established PyBOP coupling method (Section 3.2.4). A crude ^1H NMR spectrum of the reaction mixture showed only starting material. The reaction was left for 48 h, but coupling still did not occur. As a second approach, 4-chloro-6-(4-methoxycarbonyl)pyrimidine **32** was heated in the presence of excess side chain, **73**. The excess side chain was removed by evaporation at reduced pressure, however, ^1H NMR spectroscopy revealed large quantities of **73** which could not be separated by flash chromatography or recrystallisation.

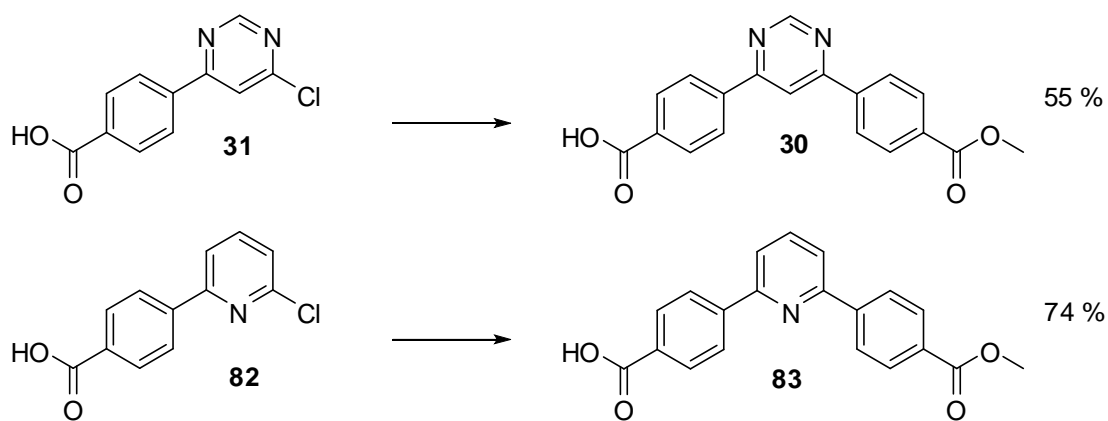


Scheme 3.22. Attempted synthesis of 4-chloro-(4-[3-imidazol-1-ylpropyl]amidophenyl)pyrimidine, **29**.

Reagents and Conditions: (i) **73**, PyBOP, Et_3N , dry DCM, RT, 48 h; (ii) **73**, Δ 48 h.

3.3.4.3 Route D

Following synthesis of carboxylic acids **31** and **82**, a second Suzuki coupling was performed, using one equivalent of 4-methoxyphenyl boronic acid, **27**. The conditions described by Gong *et al* could not be followed as this would result in hydrolysis of the methyl esters. Therefore the reagents and conditions previously used for the symmetrical compounds^{62, 63} were used (Scheme 3.23). Reactions produced high yields of products, when left for long periods of time (72 h) and were degassed for one hour prior to addition of the catalyst. However, when the reactions were left for 18 hours and only degassed for short periods of time, large amounts of starting material remained. This is possibly due to the poor solubility of the carboxylic acid compounds within the solvent system. Solubility was improved by increasing the proportion of methanol within the solvent system to 20 % compared with toluene. Starting material which did remain was removed by trituration with hot methanol.

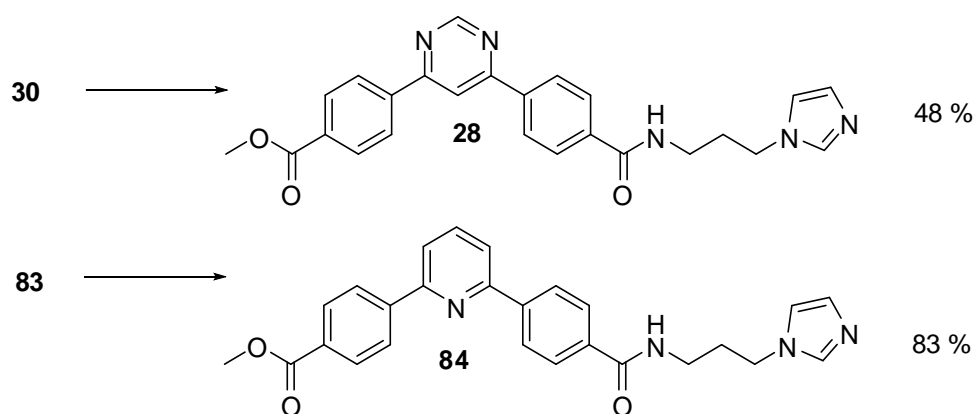


Scheme 3.23. Synthesis of non-symmetrical biarylheterocycles

Reagents and Conditions: **72**, K_2CO_3 , $Pd(PPh_3)_4$, $PhMe:MeOH$ (8:2), Δ , 72 h.

3.3.4.4 Acylation with Imidazole Side Chains

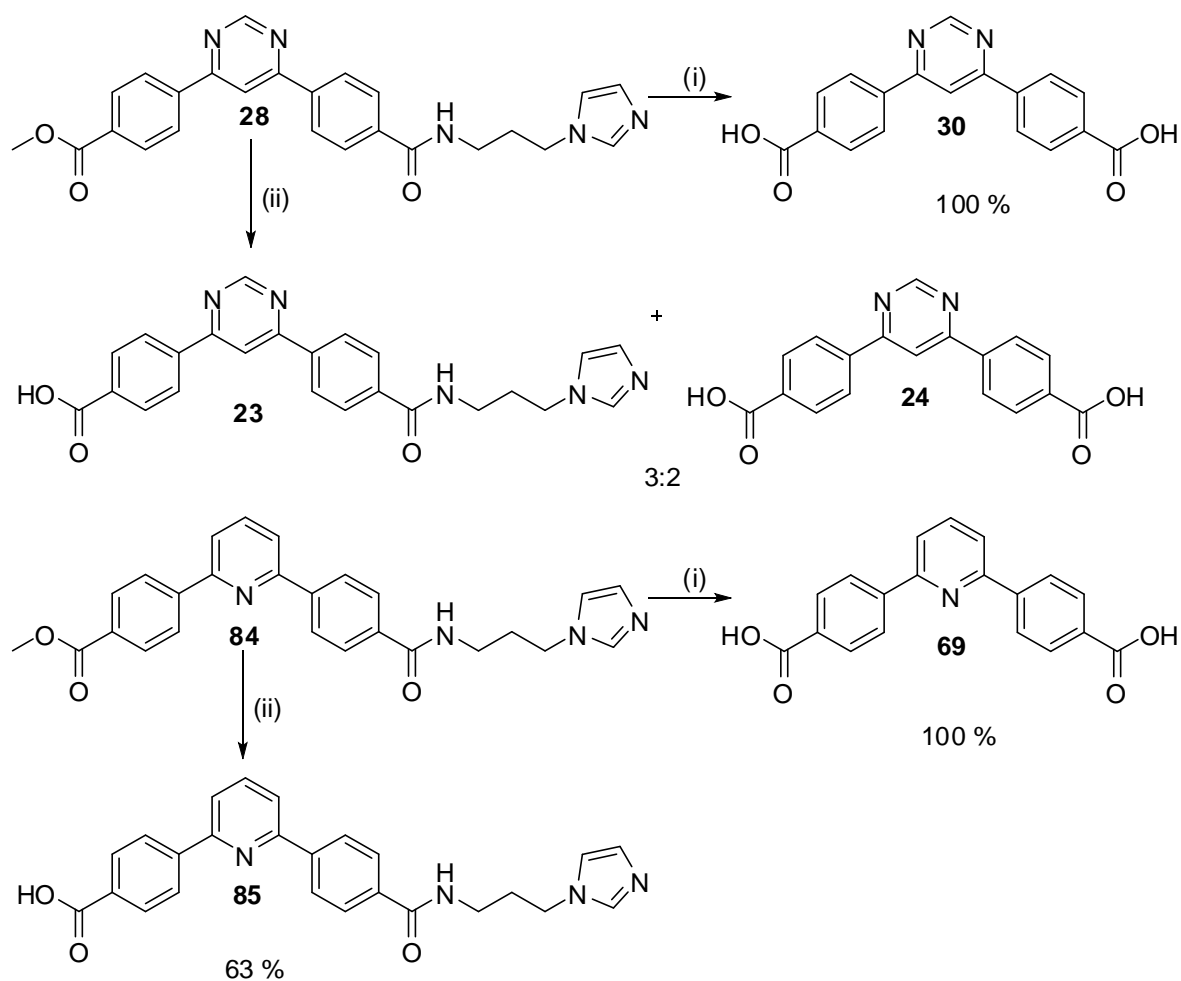
Side chains were coupled to the non-symmetrical compounds **30** and **83**, using the reaction conditions established in the synthesis of the hit compounds, **1**, **2** and **3**. Reactions proved successful yielding pure product (Scheme 3.24).



Scheme 3.24. PyBop Couplings of non-symmetrical diarylheterocycles.

Reagents and Conditions: **73**, Et_3N , PyBOP, dry DCM, RT, 18 h.

Hydrolysis of the methyl esters, following the conditions outlined in Section 3.2.2, of compounds **28** and **84**, led to the formation of the carboxylic acids, but the amide was also hydrolysed (Scheme 3.25). Using milder conditions of 10 equivalents of NaOH at 50 °C for 2 h (compared with 2 M NaOH, Δ , 24 h) led to hydrolysis of the ester while retaining the amide at the pyridine compound **84**, however, the pyrimidine compound **28**, gave a mixture of products.



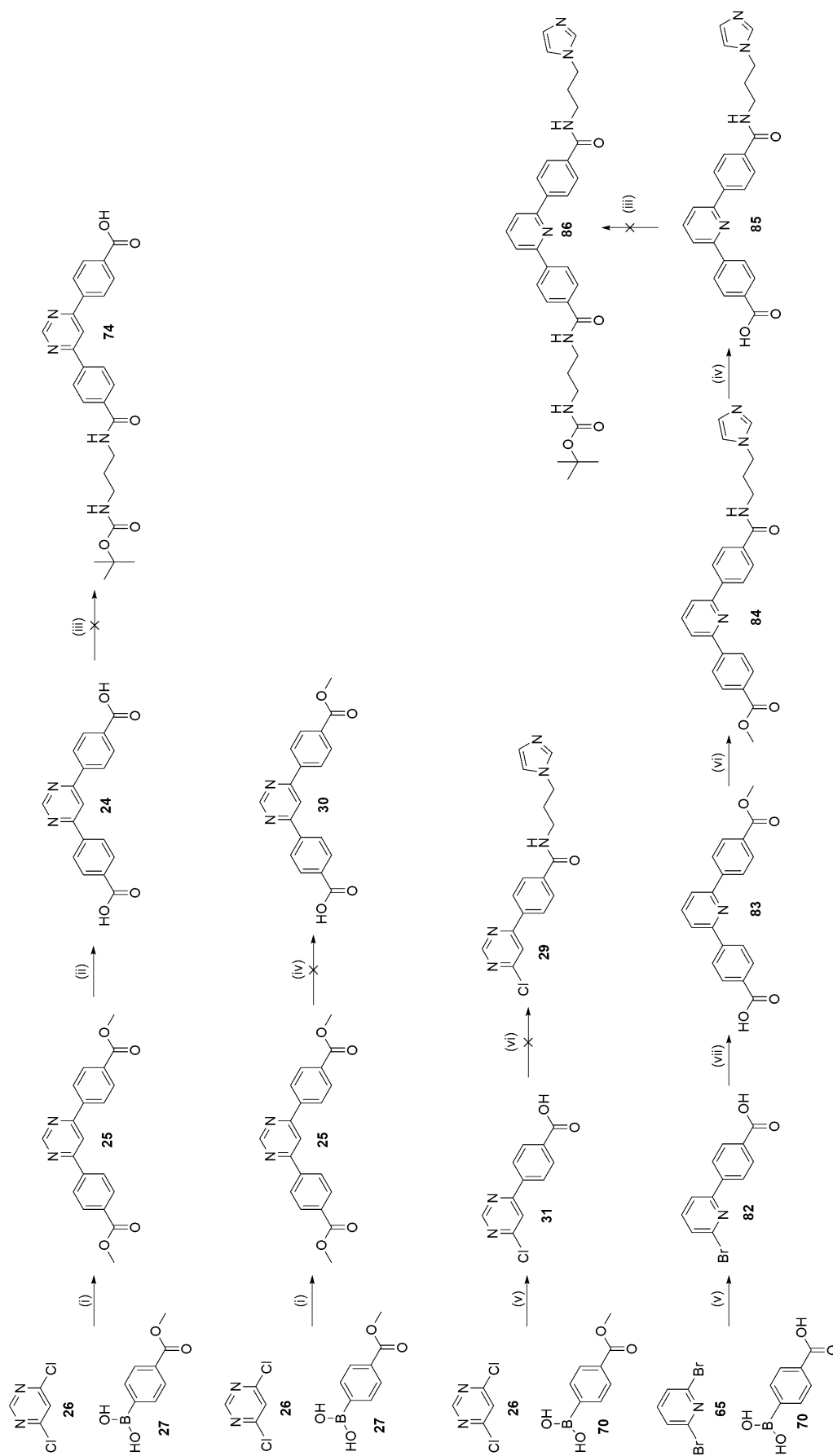
Scheme 3.25. Hydrolysis of esters bearing amides.

Reagents and Conditions: (i) NaOH (2M), Δ , 18 h; (ii) NaOH (10 eq), 50 °C, 2 h.

Attempts to couple the Boc protected side chain, **71**, to **85** resulted in large quantities of starting material remaining.

3.3.5 CONCLUSIONS

Routes to non-symmetrical compounds displayed several difficulties, although Route D (Scheme 3.2) provided the most attractive way to synthesis. Scheme 3.26 outlines the progress made during the synthesis.



Scheme 3.26. Synthesis of non-symmetrical compounds.

Reagents and Conditions: (i) K_2CO_3 , $Pd(PPh_3)_4$, $PhMe:MeOH$ (9:1), Δ , 48 h, 79 %; (ii) $NaOH$ (2M), Δ , 24 h, 100 %; (iii) **71**, Et_3N , $PyBOP$, dry DCM , RT, 18

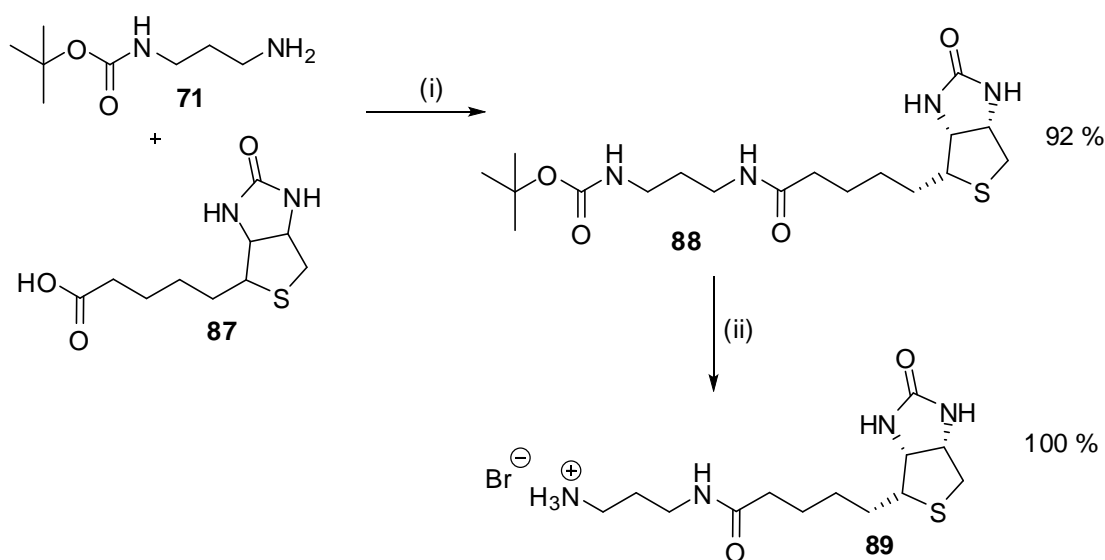
h; (iv) $NaOH$ (10 eq), 50 °C, 2 h, 80 %; (v) Na_2CO_3 (0.4M): $MeCN$ (1:1), 0.1 eq. $Pd(PPh_3)_4$, 90 °C, Ar, 24 h 64-93 %; (vi) **73**, $PyBOP$, Et_3N , dry DCM , RT, 18

h, 100 %; (vii) **16**, K_2CO_3 , $Pd(PPh_3)_4$, $PhMe:MeOH$ (8:2), Δ , 72 h, 80 %.

3.4 SYNTHESIS OF NON-SYMMETRICAL BIOTINYLATED COMPOUNDS

As discussed in Chapter 2, biotinylated compounds offer an attractive possibility for attachment to streptavidin activated sepharose beads. Due to the problems experienced with synthesising non-symmetrical compounds **22** and **86** bearing linker **5** following Route D (Scheme 3.2), it was speculated that synthesis of a biotinylated compound would prove more straight forward.

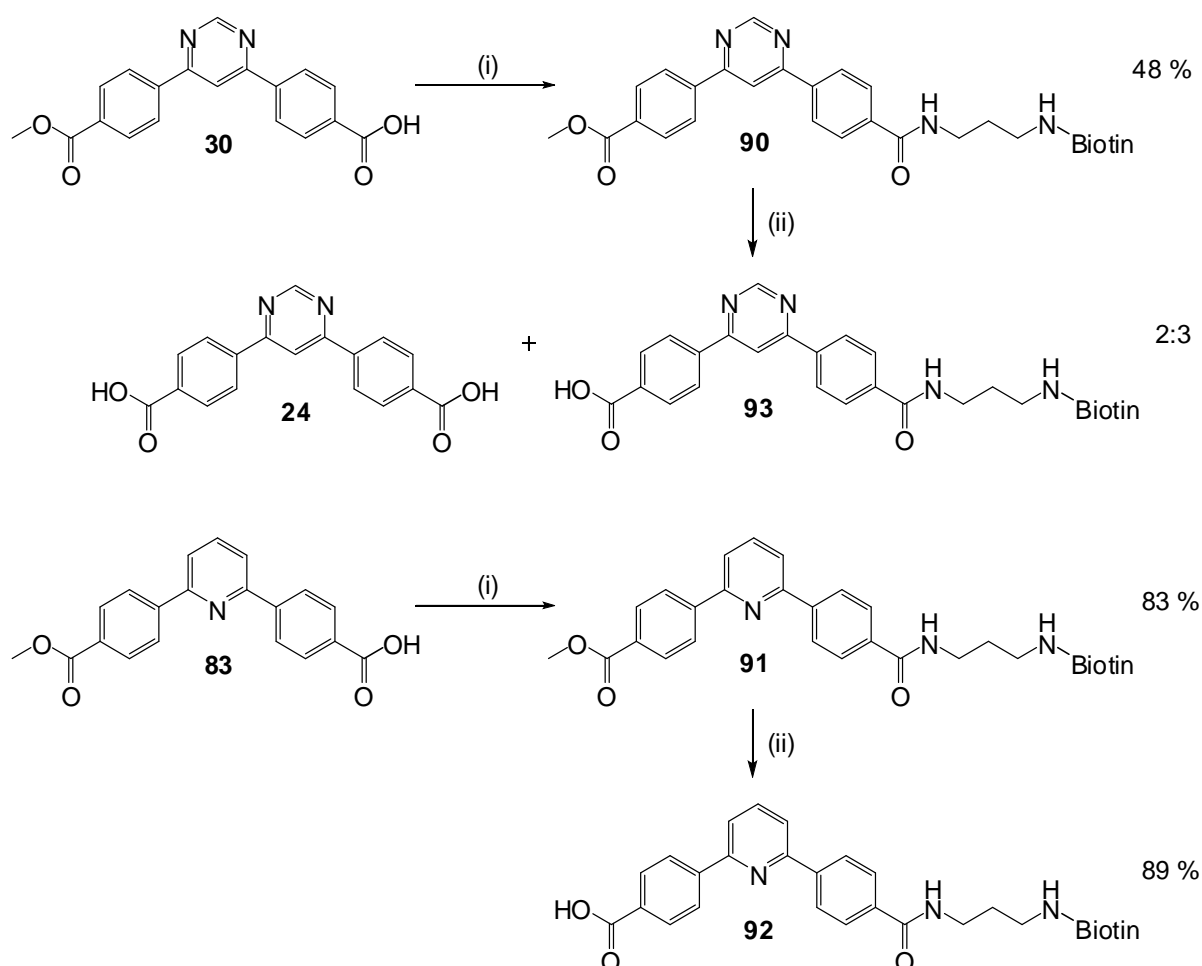
Biotin, **87** contains a carboxylic acid functional group and was attached to the Boc-protected linker, **71**, described in Section 3.3.1, by PyBOP coupling (Scheme 3.27). The coupling produced high yields of *N*-Boc-1-amino-3-biotinylamidopropane, **88**, which could be easily deprotected using concentrated HBr in ethanol. Excess HBr was evaporated under reduced pressure yielding the ammonium hydrobromide salt, **89**.



Scheme 3.27. Synthesis of biotinylated linker **89**.

Reagents and Conditions: (i) PyBOP, Et₃N, DMF, RT, 18 h; (ii) HBr, EtOH.

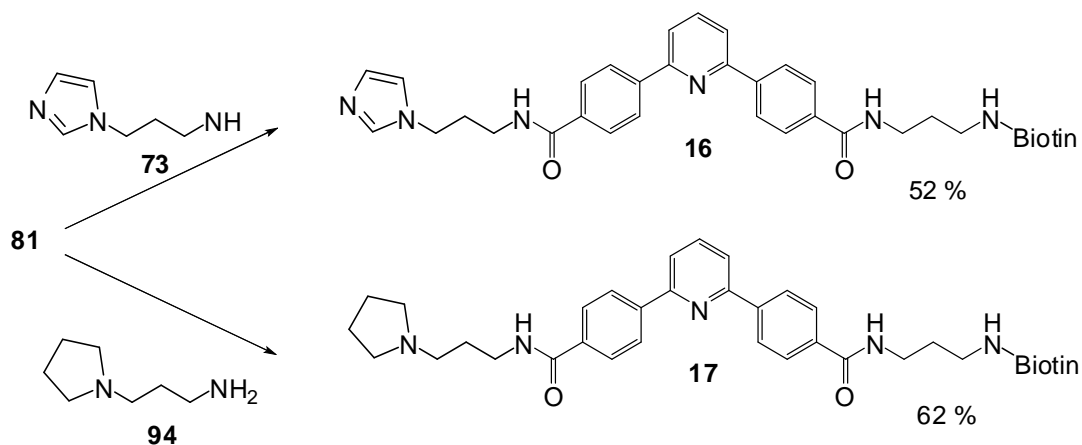
The biotin side chain, **89**, was directly coupled to 2-(4-methoxyphenyl)-6-(4-carboxyphenyl) pyridine, **83**, and 4-(4-methoxyphenyl)-6-(4-carboxyphenyl) pyrimidine, **30**. The solvent used for these couplings was DMF in which all the starting materials were soluble, allowing shorter reaction times. Subsequent hydrolysis of compound **90** and **91**, gave the corresponding carboxylic acid **92** for the pyridine compound while the amide remained intact, however the pyrimidine compound led to a mixture of diacid **24** and the mono-amide **93** (Scheme 3.28).



Scheme 3.28. Synthesis of non-symmetrical biotinylated compounds.

Reagents and Conditions: (i) **89**, PyBOP, Et_3N , DMF, RT, 18 h; (ii), NaOH (10 eq), 50 °C, 2 h.

1-(3-Aminopropyl)imidazole, **73**, or 1-(3-aminopropyl)pyrrolidine, **94**, were coupled to the non-symmetrical pyridine acid, **92**, to give the final affinity probes **16** and **17** (Scheme 3.29).



Scheme 3.29. Synthesis of non-symmetrical biotinylated affinity probes.

Reagents and Conditions: PyBOP, Et_3N , DMF, RT, 18 h.

3.4.1 CONCLUSIONS

Synthesis of biotinylated non-symmetrical compounds yielded two final products, **16** and **17**, which could be used as affinity probes for chemical proteomics to identify the binding proteins.

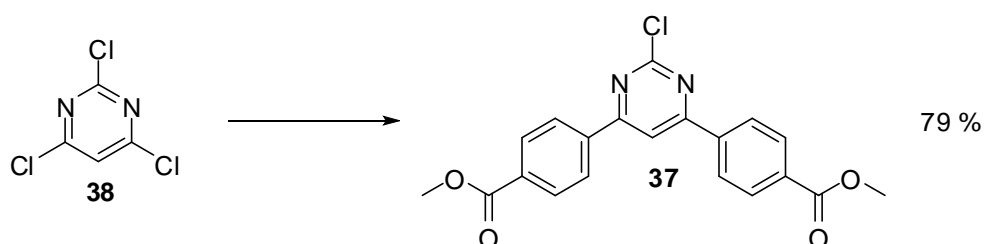
3.5 ATTACHMENT OF A LINKER TO THE PYRIMIDINE RING

An alternative approach to bespoke matrices for use in affinity chromatography involves attachment of the linker at the pyrimidine-2 position of hit compound **1**. Incorporation of linker **5**, could be achieved prior to any of the stages of the established synthesis outlined in Scheme 3.1. A retrosynthetic analysis of

compound **33**, is presented in Scheme 3.3, showing the different routes available (Routes E & F).

3.5.1 FORMATION OF THE TRIARYL HETEROCYCLIC RING SYSTEM

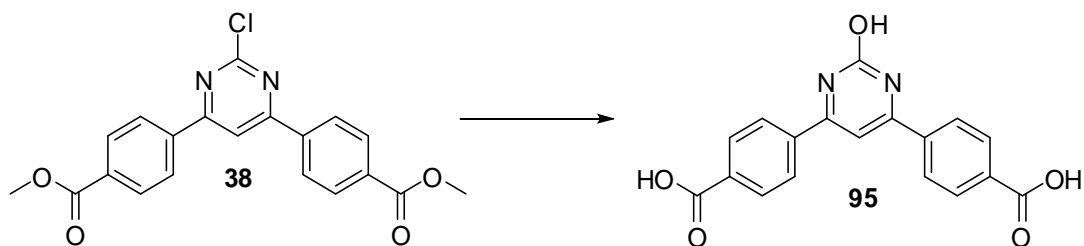
Sequential Suzuki coupling reactions following the Delia *et al.* method described in Section 3.3.4.1 successfully afforded product in high yields. The product precipitated from the glyme-water solvent system on completion of the reaction allowing straightforward purification by recrystallisation from ethanol.



Scheme 3.30. Selective arylation of trichloropyrimidine

Reagents and Conditions: **27**, aq. Na_2CO_3 , $\text{Pd}(\text{OAc})_2$, PPh_3 , glyme, Δ , 18-24 h.

Attempted hydrolysis of **37** resulted in the substitution of the 2-chloro group of the compound along with the hydrolysis of the methyl esters as seen during the synthesis of non-symmetrical compound **79**, Scheme 3.19 (Scheme 3.31). Therefore, attempts to attach the linker after the Suzuki coupling prior to hydrolysis of the methyl esters were made.



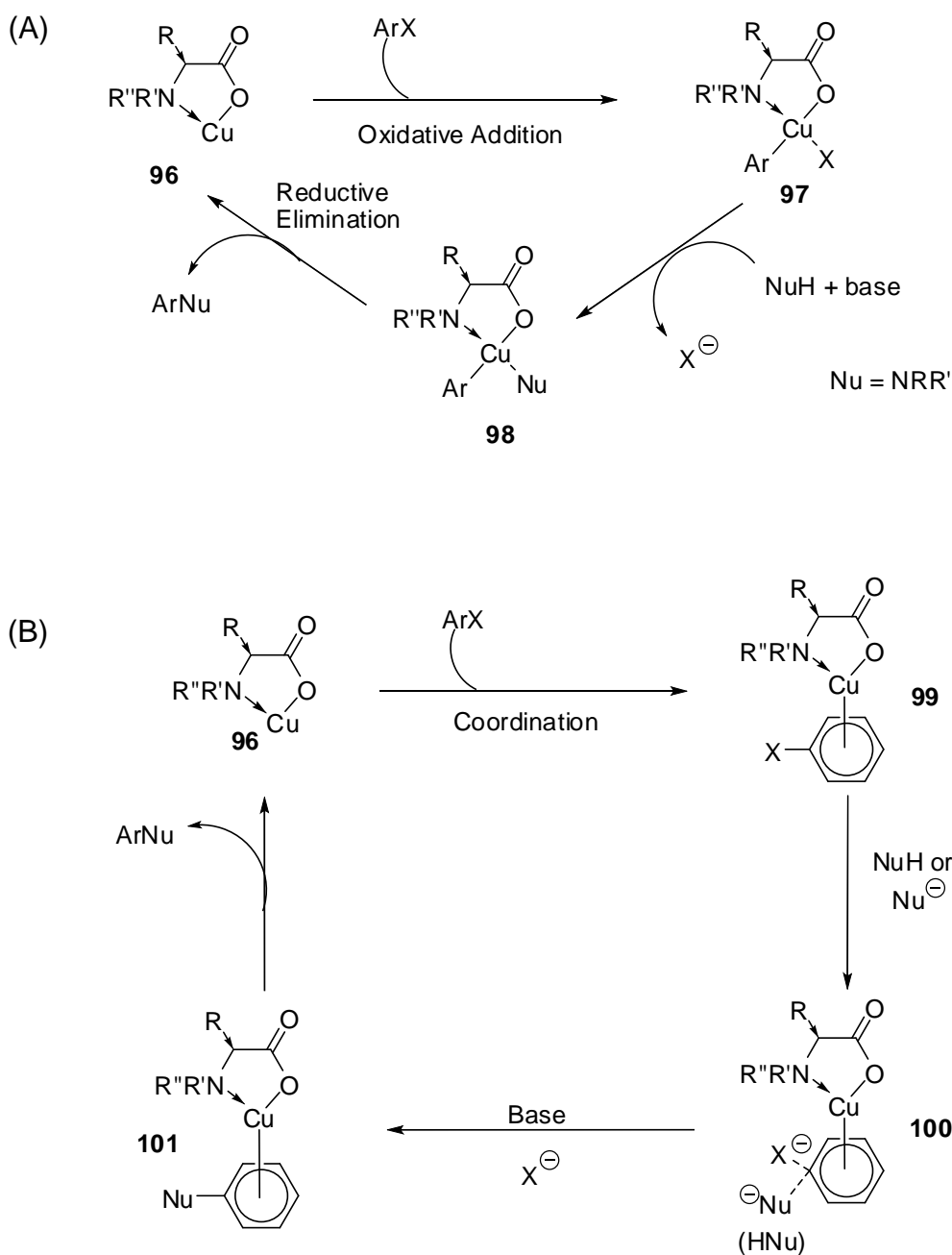
Scheme 3.31. Hydrolysis of **37**.

Reagents and Conditions: NaOH (2 M), Δ , 18 h.

3.5.2 ATTACHMENT OF THE LINKER

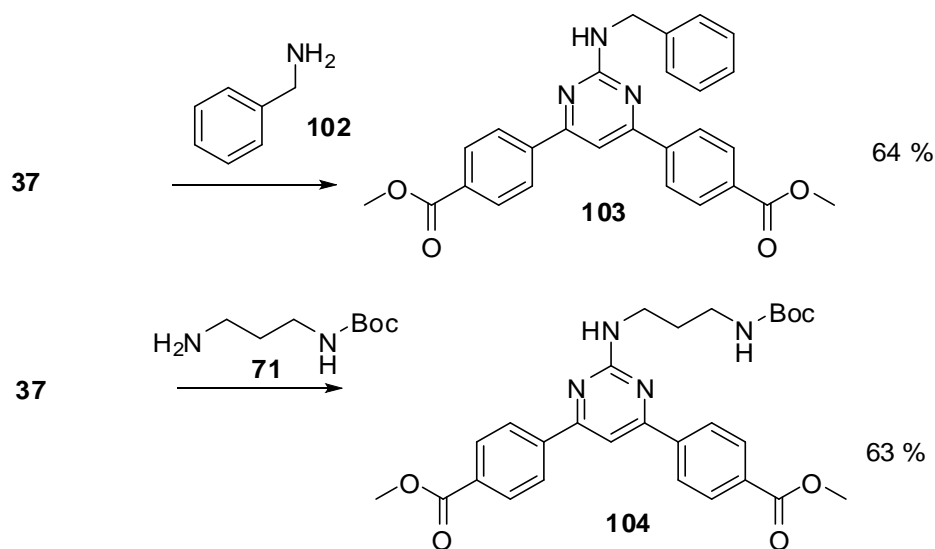
Copper-catalysed coupling reactions of aryl halides and primary amines can occur using *L*-proline as a promoter in Ullmann type coupling reactions.¹⁶⁰ The mechanism of the Ullmann reactions has been extensively studied and many mechanisms have been proposed. One of these includes the oxidative addition and reductive elimination cycle seen with palladium catalysed cross-couplings.¹⁶¹ However, this mechanism seems unlikely as copper (III) complexes are rarely seen. Research by Paine *et al* proved that the catalytic species is the cuprous ion.¹⁶²

Two different reaction mechanisms have been proposed by Zhang and co-workers.¹⁶⁰ (Scheme 3.32). The first (Scheme 3.32A) suggests that as Cu (I) complexes with the amino acid promoter, **96**, it becomes more reactive towards oxidative addition and stabilises the intermediate **97**. This catalytic cycle can explain the order of I > Br > Cl with respect to the ease of halogen displacement and the increased reactivity of electron deficient aryl halides.



to nucleophilic attack of the arylamines. This scheme could explain the substituent effect of the aryl halides.

The reaction provided an attractive route for attachment of the linker, **71**, to the hit compounds. The reaction was piloted using compound **37** and benzylamine, **102**. The reaction produced pure product, **103**, after extraction and recrystallisation in relatively good yields (64 %). Due to the success of the reaction it was carried out with *N*-Boc-1,3-diaminopropane, **71**, which also produced pure compound, **104**, in a yield of 63 %.

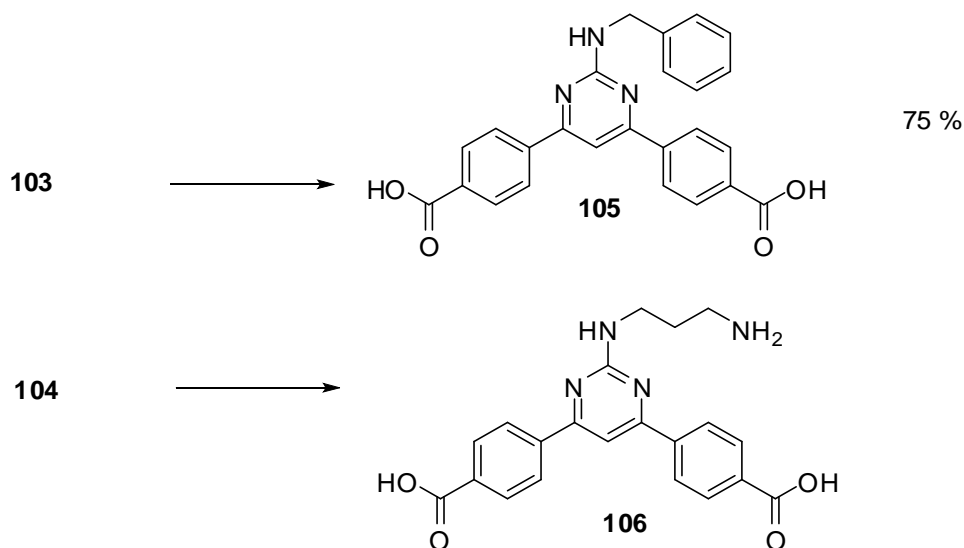


Scheme 3.33. Attachment of the linker to a pyrimidine ring.

Reagents and Conditions: K_2CO_3 , *CuI*, *L-Pro*, *DMSO*, 80 °C, 12 h.

Compounds **103** and **104** were hydrolysed as described in Section 3.2.2. The benzylamino compound, **103**, was successfully hydrolysed, precipitating out of solution affording pure product in 75 % yield. 1H NMR spectral data of the ester hydrolysis of compound **104** indicated that hydrolysis was successful, but the Boc-group was lost during the extraction process. Slow neutralisation of the

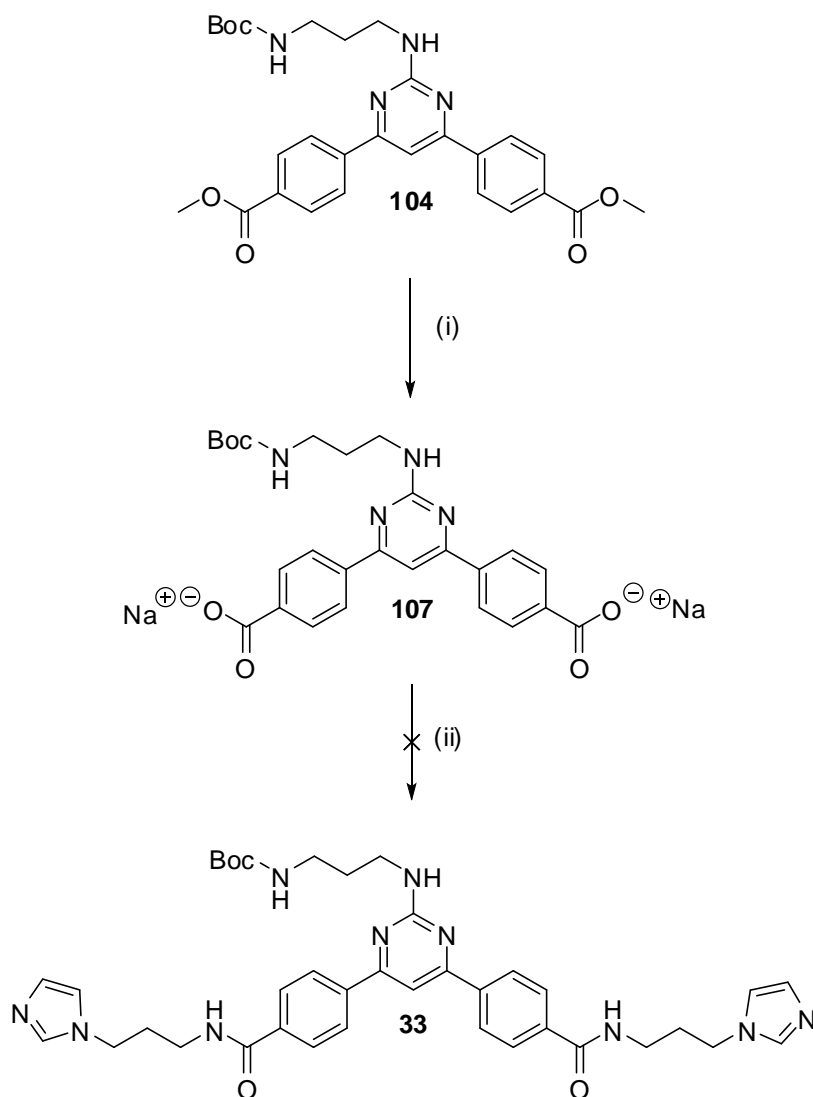
product with dilute acetic acid still led to premature removal of the Boc group, **106**.



Scheme 3.34. Hydrolysis of diesters **103** and **104**.

Reagents and Conditions: NaOH (2M), Δ , 24 h, then HCl (4 M).

Ester hydrolysis of compound **104** was repeated using 3 equivalents of sodium hydroxide in water. The water was evaporated under reduced pressure to leave the sodium salt, **107**. Peptide coupling of **107** with **73** was carried out following the method outlined in Scheme 3.24. However, the couplings did not proceed, leaving large quantities of starting material, even when left for long periods of time. The hydrolysis was repeated using potassium hydroxide as a base for isolation of the potassium salt, however, the same difficulties occurred during coupling.

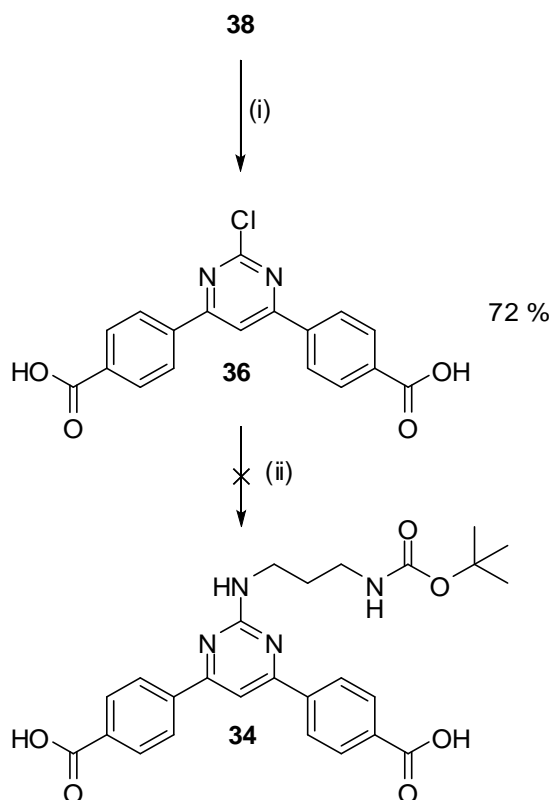


Scheme 3.35. Attempted synthesis of **33**.

Reagents and Conditions: (i) NaOH (2M), Δ , 24 h; (ii) **73**, PyBOP, Et₃N, DMF, RT, 48 h.

Sequential Suzuki coupling reactions were performed using 4-carboxyphenyl boronic acid pinacol ester, **81**, and 2,4,6-trichloropyrimidine, **38**, to synthesise the dicarboxylic acid, **36**, directly prior to attachment of the linker (Scheme 3.36). The reactions afforded the product in high yields (72 %). Attachment of the linker was attempted using the conditions described in Scheme 3.33, but the

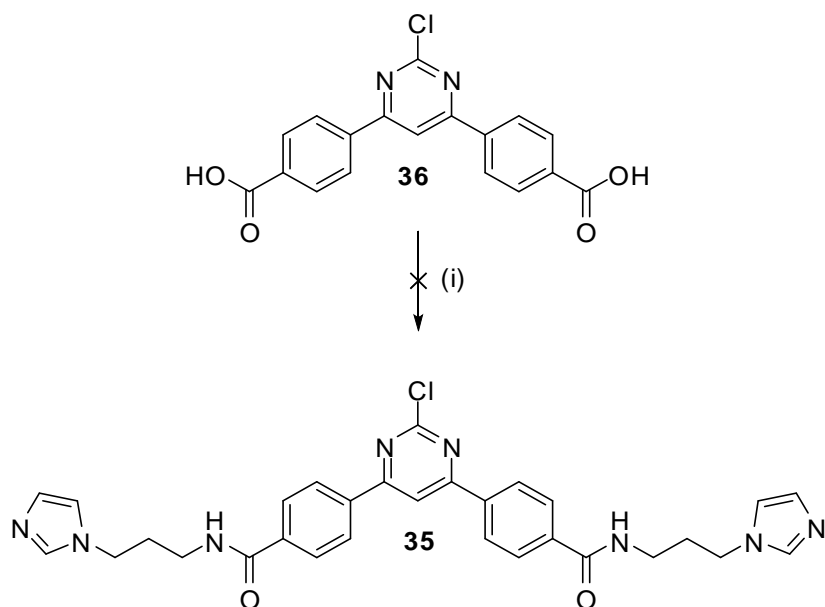
product **34** could not be isolated from the DMSO solvent. The reaction was attempted using DMF, but did not proceed, leaving large quantities of starting material.



Scheme 3.36. Attempted synthesis of **34**.

Reagents and Conditions: (i) **81**, aq. Na_2CO_3 , $\text{Pd}(\text{OAc})_2$, PPh_3 , glyme, Δ , 18-24 h; (ii) **71**, K_2CO_3 , CuI , *L-Pro*, DMSO, 80 °C, 12 h.

Following Route F attempts to couple **36** to **73** were made. Observation of the ^1H NMR spectrum revealed peaks representing the starting materials, suggesting that coupling had not occurred.

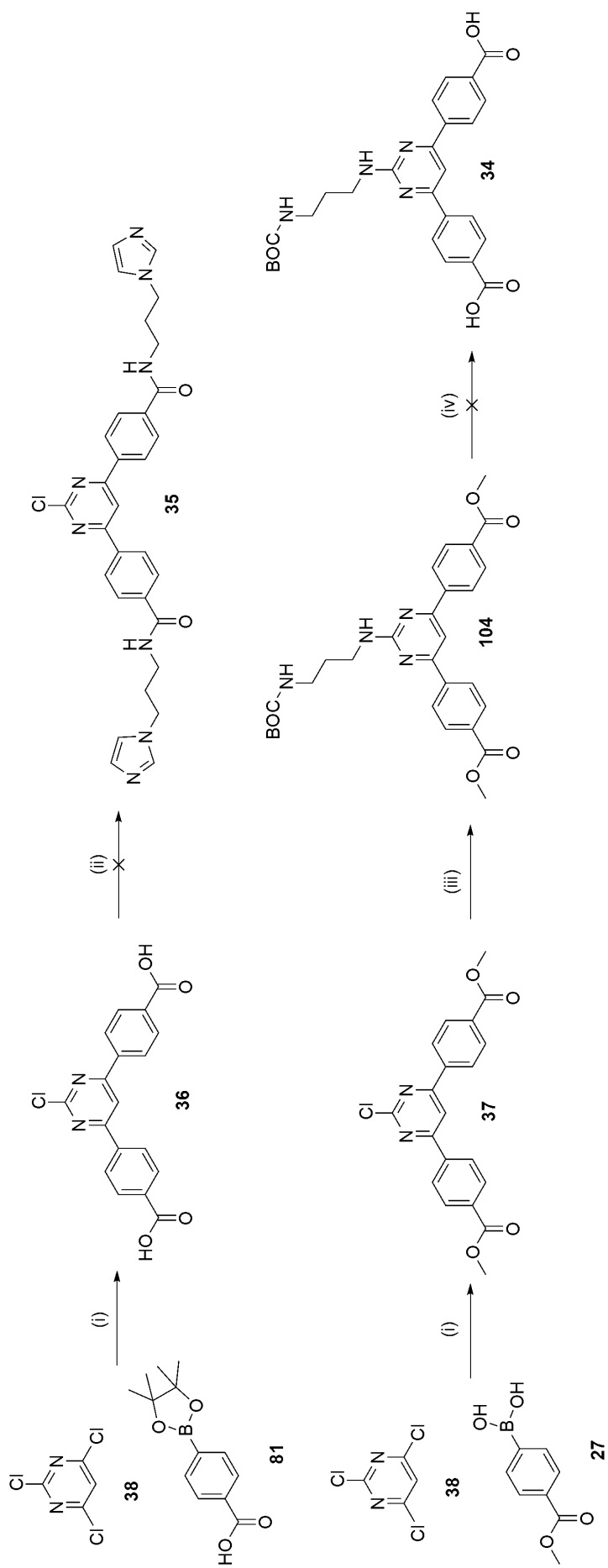


Scheme 3.37. Attempted synthesis of **35**.

Reagents and Conditions: (i) **73**, PyBOP, Et_3N , DMF, RT, 48 h.

3.5.3 CONCLUSION

Routes to synthesis of compounds with the linker attached at the pyrimidine-2 position have progressed. However, attachment of linker **5** using the Ullman reaction¹⁶⁰ has led to problems during separation due to the DMSO solvent required. Using a different catalyst to couple the linker may lead to the formation of compound **33**. Scheme 3.37 summarises the reactions carried out during the attempted synthesis of compounds with the linker attached at the pyrimidine-2 position.



Scheme 3.28. Attempted Synthesis of symmetrical compounds.

Reagents and Conditions: (i) *aq.* Na_2CO_3 , $\text{Pd}(\text{OAc})_2$, PPh_3 , glyme, Δ , 18-24 h; (ii) **73**, PyBOP , Et_3N , dry DCM, RT, 18 h; (iii) **60**, CuI , K_2CO_3 , CuI , L-Pro, DMSO, 80 °C, 12 h; (iv) NaOH (2M), Δ , 24 h.

3.6 CONCLUSION

Several routes of synthesis have been explored for the development of affinity probes for use in chemical proteomics. The two compounds taken forward for use in affinity chromatography were biotinylated compounds **16** and **17**. Compound **16** is predicted to be active against the sensitive human ovarian carcinoma A2780 cell line but not against the human colon adenocarcinoma HT29. Compound **17** is predicted to be inactive against both of the above cell lines.

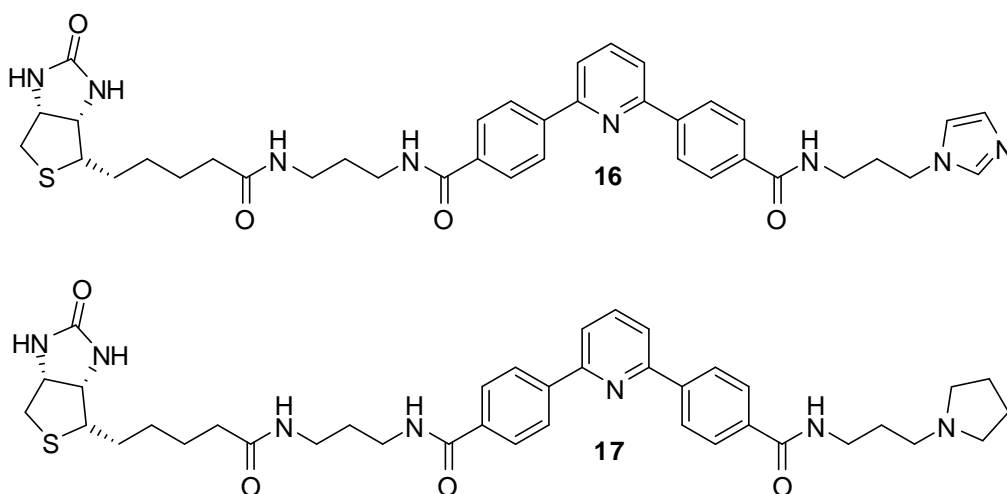


Figure 3.4. Affinity probes ready for chemical proteomics.

CHAPTER 4

BIOLOGICAL EVALUATION AND PROTEOMICS

4.1 INTRODUCTION

The two affinity probes, **16** and **17**, that progressed to use in chemical proteomics are shown in Figure 4.1.

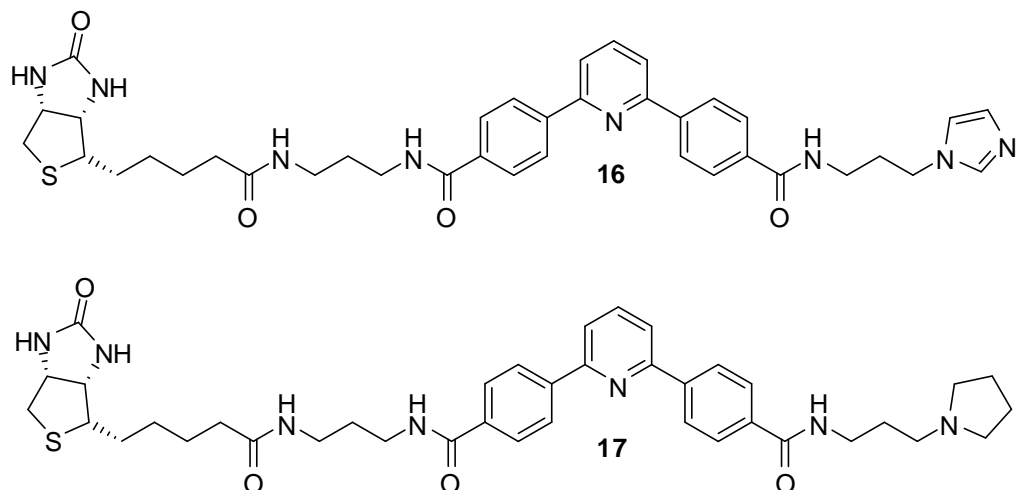


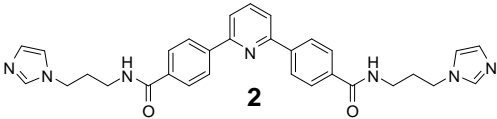
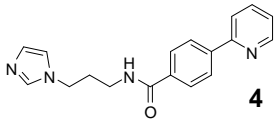
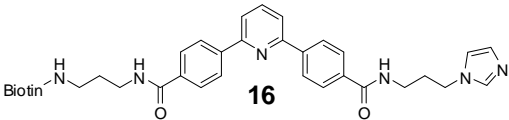
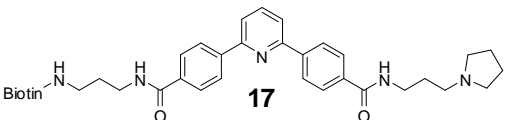
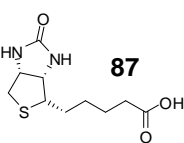
Figure 4.1. Affinity probes used for chemical proteomics.

Prior to the proteomics stage, the *in vitro* chemosensitivity of both affinity probes was measured to ensure that the distinction between the 'active', **16**, and 'inactive', **17**, analogues was retained.

4.2 *IN VITRO* CHEMOSENSITIVITY

The biotinylated compounds **16** and **17**, along with controls of biotin and hit compound **2**, were tested against the sensitive A2780 human ovarian adenocarcinoma cell line and the insensitive HT29 colorectal adenocarcinoma cell line. The chemosensitivity data are presented in Table 4.1.

Table 4.1. IC₅₀ values for compounds **2**, **4**, **16** and **17** against two cell lines.

Compound	IC ₅₀ (μM)	
	A2780	HT29
 2	0.006 ± 0.001	35.0 ± 0.1
 4	2.6 ± 0.2	47.8 ± 0.1
 16	6.25 ± 0.09	>100
 17	35.0 ± 0.2	6.25 ± 0.16
 87	>100	>100

Compound **16** showed a thousand-fold decrease in activity compared with the corresponding hit compound **2**. However, compared with the half compound **4** there was only a 2.4 fold decrease, suggesting that the full structure, including both imidazole rings, is required for potent activity. Biotin possesses no cytotoxicity and as such ought not to affect the activity observed for biotinylated compounds. The 2.4 fold decrease in activity between the half compound **4** and compound **16** could be explained by the ability of biotinylated compounds to enter the cells. The activity observed against the A2780 cell line was not seen with the insensitive HT29 cell line ($IC_{50} > 100 \mu M$), supporting the NCI COMPARE data (Figure 1.10).

Biotinylated compound **17** showed a 5.6 fold increase in IC_{50} compared with its imidazole equivalent **16**, supporting the SAR data that a saturated heterocyclic ring (such as pyrrolidine) is less active than an unsaturated ring.⁶² Surprisingly, the control probe **17** showed relatively potent activity with the 'insensitive' cell line HT29, which conflicts with the previous SAR data. It could be rationalised that the control compound has a different biological target leading to a different mechanism of action, or that biotinylation has resulted in a more potent compound still binding the same target.

Compound **16** has shown the expected distinction in chemosensitivity against the two cell lines, activity against the sensitive A2780 cell line ($IC_{50} = 6.25 \mu M$) and no activity against the insensitive HT29 cell line ($IC_{50} > 100 \mu M$). Therefore the compound is suitable for use as an affinity probe in chemical proteomic analysis. Control compound **17** has shown unexpected cytotoxicity against the

insensitive cell line HT29 ($IC_{50} = 6.25 \mu\text{M}$). However, potency is decreased, as expected, against the sensitive A2780 cell line when compared with compound **16**.

4.3 PROTEOMICS

The compound-centric approach to chemical proteomics, often referred to as classical affinity chromatography¹¹⁸ (Section 1.3.3), involves ‘fishing’ for a cellular target protein from an entire cell lysate or tissue extract.¹⁶³ Affinity capture with probes **16** and **17**, utilises the strong affinity of biotin for streptavidin. This allows for a simple isolation of the biological target following incubation of the probes with cell lysates.

Prior to affinity capture, large numbers of cells ($1.5 \times 10^7 - 6.1 \times 10^8$) were cultured to ensure that low abundance proteins were represented at detectable levels in the cell lysates. Cells were sub-cultured regularly to ensure they were kept in the exponential growth phase and were trypsinised, pelleted and stored at $-80 \text{ }^\circ\text{C}$ until there was sufficient for affinity chromatography.

As the abundance of the biological target and its affinity for the biotinylated compound **16** were unknown, four different methods were explored to capture binding proteins, defined in Figure 4.2.

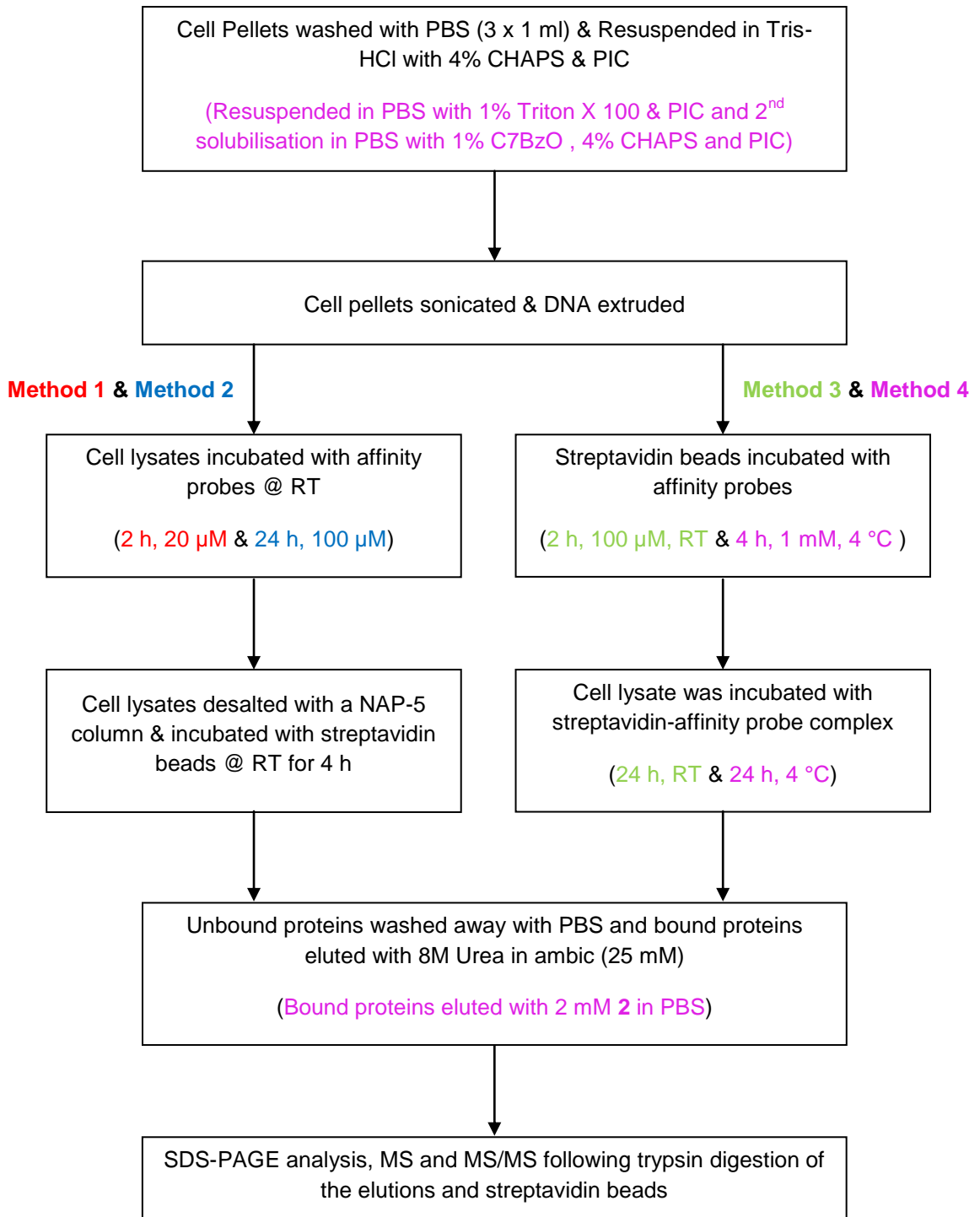


Figure 4.2. Affinity chromatography procedure. Method 1 (red), Method 2 (blue), Method 3 (green) and Method 4 (purple).

4.3.1 METHOD 1

Cells lysis can be achieved by mechanical or chemical techniques. The mechanical method sonication was chosen as it is less likely to result in denaturation of cellular proteins. The cells were suspended in Tris-HCl buffer containing 4 % 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS) detergent (to ensure solubilisation of membrane proteins) and complete protease inhibitor cocktail (PIC, to prevent proteolysis) prior to sonication. The DNA present in the samples was extruded using a fine needle.

The protein solution was split in two and one of the affinity probes (**16** or **17**, 20 μ M) was added to each batch and incubated for 2 h. Samples were then desalted using a NAP-5 column to remove any affinity probe not bound to a protein (therefore purifying the probe-protein complex) and to change the buffer to PBS. The protein-probe samples were incubated with streptavidin beads for 4 h at room temperature. Unbound proteins were removed, by washing with PBS and bound proteins eluted with 8 M urea in ammonium bicarbonate (25 mM).

The urea elutants and streptavidin beads from each experiment were analysed by SDS-PAGE (Figure 4.3). Visualisation of the proteins using an instant blue stain revealed only the presence of a band at \sim 13 kDa in lanes 5 and 9, corresponding to a streptavidin monomer from the streptavidin beads. The urea elutions did not contain any proteins. Silver staining of the gel confirmed the

presence of streptavidin, however, it still did not show any further bands which would suggest the presence of a bound target protein.

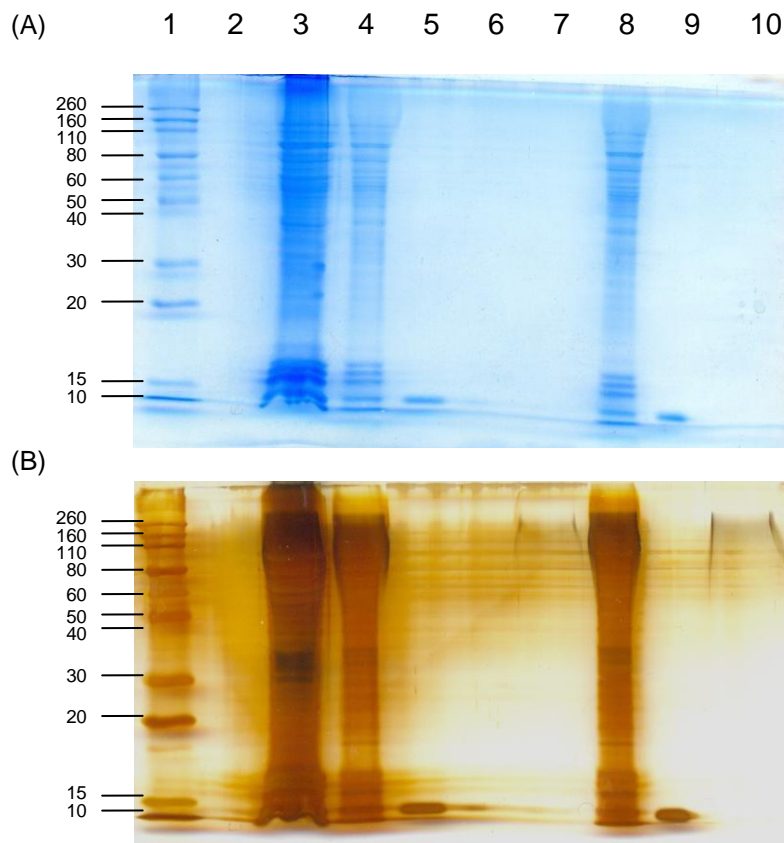


Figure 4.3. SDS-PAGE Analysis. (A) Stained with Instant Blue stain; (B) Silver Stained.

Lane 1, Molecular weight marker; **Lane 2** and **Lane 6**, SDS buffer; **Lane 3**, total protein content prior to incubation with affinity probes; **Lane 4**, 16 unbound proteins; **Lane 5**, 16 streptavidin beads; **Lane 7**, 16 urea elution; **Lane 8**, 17 unbound proteins; **Lane 9**, 17 streptavidin beads; **Lane 10**, 17 urea elution.

The same samples were trypsin digested and the peptides subjected to manual MALDI mass spectrometric analysis. MS analysis of the urea extracts showed only low mass peaks for experiments using both affinity probes, suggesting

contamination with detergents. Mascot protein searches did not identify any proteins from the peptide mass fingerprint (PMF) (Figure 4.4).

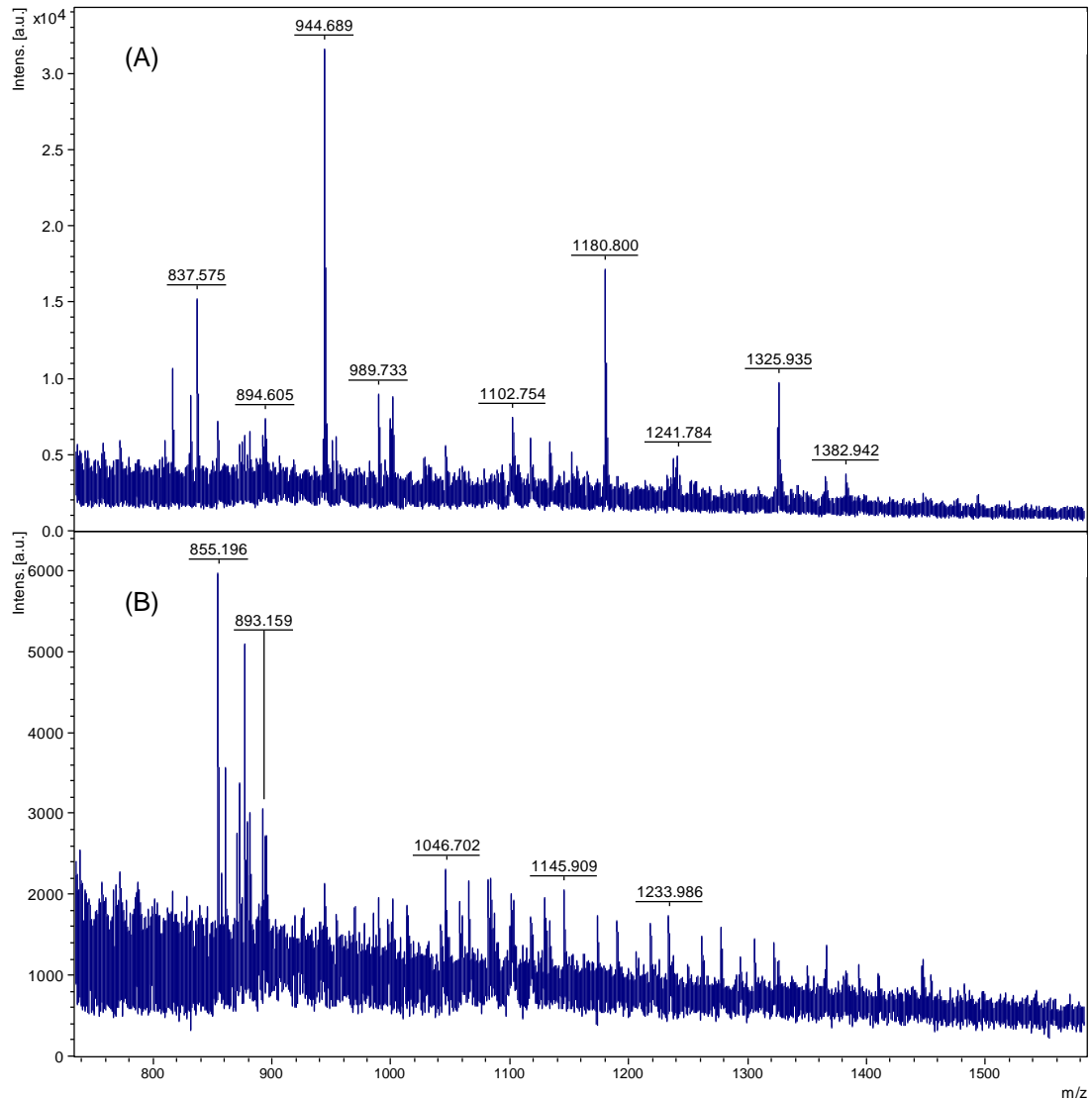


Figure 4.4. Mass spectra for urea elutions. (A) 16; (B) 17.

MS analysis of the streptavidin beads following incubation with the affinity probes revealed similar PMFs with only very few unique peaks, suggesting that proteins had failed to bind to the streptavidin-affinity probe complex (Figure 4.5).

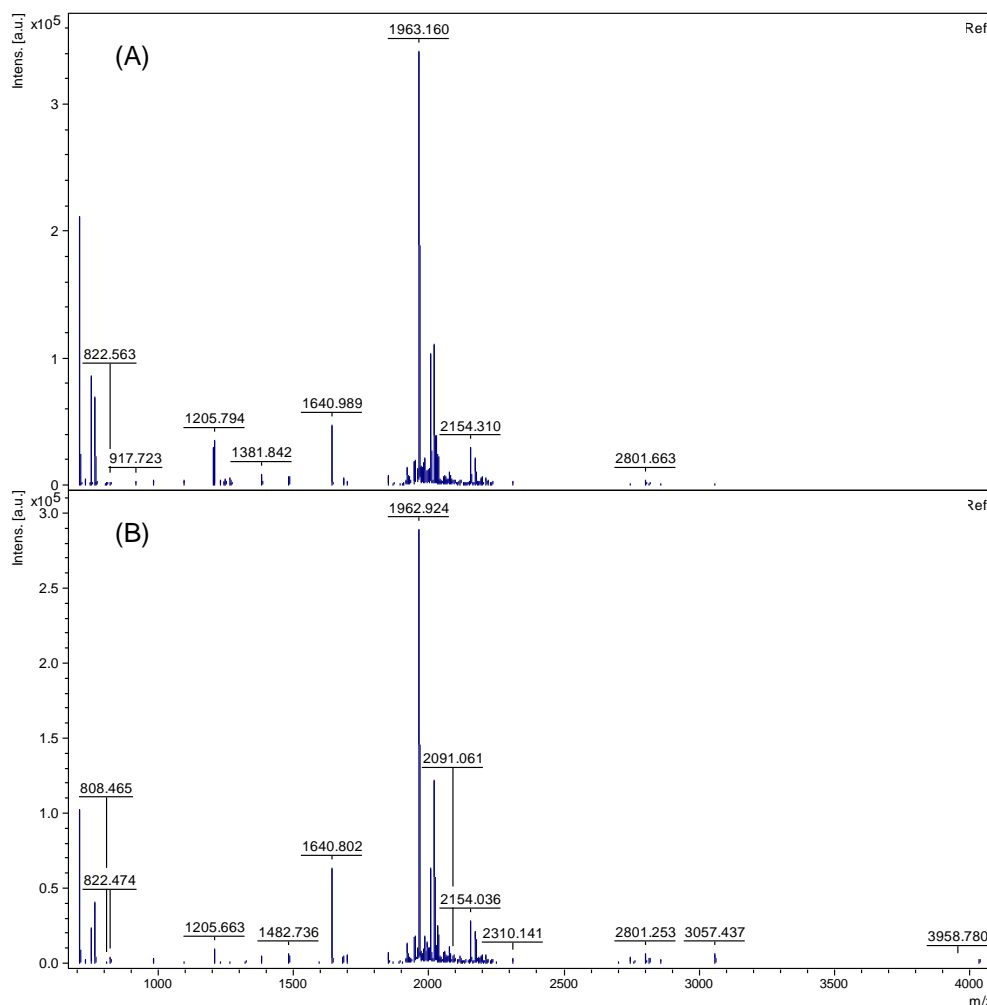


Figure 4.5. MS spectra for streptavidin beads. (A) **16**; (B) **17**.

Mascot searching identified one protein on the beads incubated with probe **17** - porphobilinogen deaminase from the species *Chlorobium limicola* (Mascot score 81, number of peptides 13). The peptides have a relatively high sequence coverage (53 %) and the matched peptides are shown in Table 4.2. However, the result can be considered a false positive as most of the peptides identified have one or more missed cleavages by trypsin digestion, several modifications and variable ppm differences in molecular weight. In addition, contamination with proteins of *Chlorobium* to either the streptavidin beads or the cell line used for protein extraction seems unlikely.

Table 4.2. Protein hit from the PMF of streptavidin beads incubated with affinity probe
17. (A) Sequence coverage identified by peptides highlighted in red. (B) Matched peptides.

(A)

1 MKK**QLIIGTR SSPLALWQAE FTKAELSR**HF PELDITLKLKLV KTTGDEVLLDS
 51 PLSK**IGDMGL FTKDIEKHLI AKEIDLAVHS LKDVPTSTPE GLIITSFTE**R
 101 EDTRDVIISK GGAKLADLPL NAKVATSSLR RMSQLK**SLRP DFEICDIRGN**
 151 **LNTRFKKFDE GEFDAMMLAY AGVFRLNFS**D **RISEILPHEI MLPVGGAL**
 201 **GIETRVDD**EQ **TREIVRILNH SNT**EY**CCK**AE RALLRHLQGG CQIPIGAYAS
 251 FKNGLTKLLA FVGSVDGTVG INNEITRSG L TSPDQAE EAG IALAEELLKQ
 301 GADK**ILSEIR** KTR

(B)

Amino Acids	ppm	Miss	Sequence
4 - 28	-91	2	K.QLIIGTRSSPLALWQAEFTKAELSR.H
11 - 28	-8	1	R.SSPLALWQAEFTKAELSR.H
55 - 67	-10	1	K.IGDMGLFTKDIEK.H Oxidation (M)
64 - 82	-63	2	K.DIEKHLIAKEIDLAVHSLK.D
83 - 100	-39	0	K.DVPTSTPEGLIITSFTE.R
137 - 154	-37	1	K.SLRPDFEICDIRGNLNTR.F Carbamidomethyl (C)
157 - 175	15	1	K.KFDEGEFDAMMLAYAGVFR.L
157 - 175	40	1	K.KFDEGEFDAMMLAYAGVFR.L 2 Oxidation (M)
158 - 181	-6	1	K.FDEGEFDAMMLAYAGVFRLNFSDR.I
176 - 181	99	0	R.LNFSDR.I
176 - 212	-37	2	R.LNFSDRISEILPHEIMLPVGGALGIETRVDDEQTR.E Oxidation (M)
213 - 228	-8	1	R.EIVRILNHS SNT EY CCK .A Carbamidomethyl (C)
305 - 310	-28	0	K.ILSEIR.K

Search Parameters: All species; Database – Swissprot; Enzyme – Trypsin; Mass Tolerance – 100 ppm; Variable modifications – oxidation (M) & carboxyamidomethyl (C); Partials – 2.

4.3.2 METHOD 2

Method 1 did not result in the identification of any proteins that bound to either affinity probe. This may have been due to the low amount of protein present in the cell lysates. Method 2 followed the same procedure with an increased number of cells used (3.0×10^7 , instead of 1.5×10^7) and higher concentration of affinity probes in the protein solution (100 μ M instead of 20 μ M).

With this method, SDS-PAGE analysis failed to visualise any proteins in the urea elutions or on the streptavidin beads after incubation using either affinity probe **16** and **17**. MS analysis of the urea elutions did not lead to the identification of proteins present in the samples from the PMFs. The most intense peptide peaks were subject to MS/MS analysis. The MS/MS spectra of two of the peptide peaks did not generate any hits with Mascot searching and the samples were too weak to allow good MS/MS spectra to be produced from less intense peaks. One peptide ($m/z = 944.4450$) was identified as Histone H2A type 1A (Mascot score 28) (Figure 4.6). Although this was a positive result, a protein identification from one peptide is not sufficient to confirm this as a specific binding protein of the affinity probe **17**.

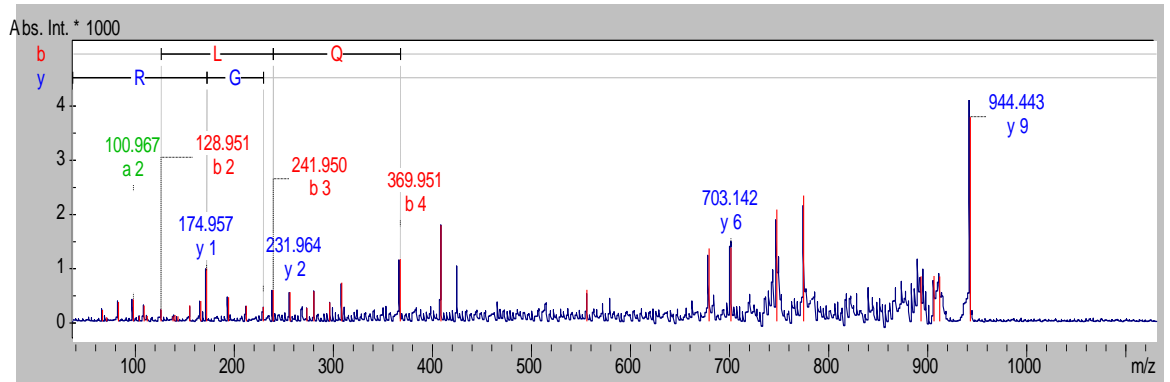
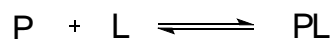


Figure 4.6. MS/MS spectrum of peptide $m/z = 944.4450$ from the MS of the urea elution with affinity probe **16**, highlighting the matched peptide fragments.

Search Parameters: All species; Database – Swissprot; Enzyme – Trypsin; Peptide Tolerance – 100 ppm; MS/MS Tolerance – 0.7 Da; Variable modifications – oxidation (M) & carboxyamidomethyl (C); Partial – 2.

4.3.3 METHOD 3

Method 3 followed a different approach to affinity capture. It was considered that desalting with a NAP-5 column, to remove excess affinity probe, before capture of the probe-protein complex with the streptavidin beads, may disrupt the binding equilibrium of the target with the affinity probe. This would result in a reduced amount of protein (P) binding to affinity probe (L).



Therefore, incubation of the affinity probes with the streptavidin beads was carried out prior to exposure to the cellular lysates. This would immobilise the affinity probes onto the column preventing loss by desalting and hence

increasing the amount of protein captured. The conditions of each stage remained the same as those used in Method 2.

Six parallel affinity experiments were carried out incorporating a sensitive (A2780) and an insensitive cell line (HT29), an active (**16**) and an inactive (**17**) affinity probe and a probe-free control (Table 4.3).

Table 4.3. Experiments carried out following Method 3

Experiment	Affinity Probe	Cell Line
1	16	A2780
2	17	A2780
3	No Compound	A2780
4	16	HT29
5	17	HT29
6	No Compound	HT29

SDS-PAGE analysis of the urea elutions from each experiment did not reveal the presence of proteins using instant blue staining (Figure 4.7A). Silver staining of the gels showed the presence of protein bands in lanes containing samples

from incubation of both affinity probes with the insensitive cell line HT29 (experiments 4 and 5). This contradicts the chemosensitivity data, therefore the bands may represent specific binding of proteins that do not have a direct involvement with the mechanism of action of the hit compounds. The absence of bands in other samples was most likely due to the low quantity of proteins captured.

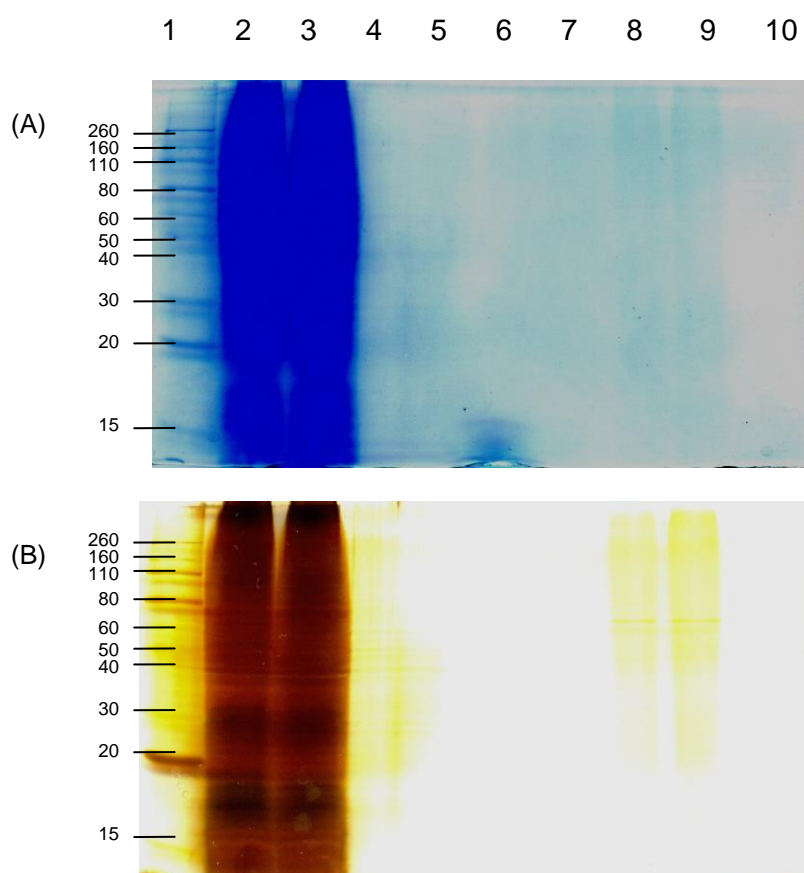


Figure 4.7. SDS-PAGE Analysis of Urea Extracts. (A) Stained with Instant Blue stain; (B) Silver Stained.

Lane 1, Molecular weight marker; **Lane 2**, total protein content from A2780 cell line prior to incubation with affinity probes; **Lane 3**, total protein content from HT29 cell line prior to incubation with affinity probes; **Lane 4**, SDS buffer; **Lane 5**, experiment 1; **Lane 6**, experiment 2; **Lane 7**, experiment 3; **Lane 8**, experiment 4; **Lane 9**, experiment 5; **Lane 10**, experiment 6.

The urea elutions from each experiment were subject to MS analysis. Comparing the spectra from experiments 1 and 3 eliminated peptides from non-specific binding proteins revealing peptides from proteins binding to **16**. However, as the samples were weak, identification of only one peptide ($m/z = 944.51$) was possible – Histone H2A type 1-B/E (Mascot score = 48). SDS-PAGE of the streptavidin bead samples from each experiment only showed bands due to streptavidin (Figure 4.8).

Comparisons of MS spectra for each sample allowed elimination of non-specific binding proteins in the same way as for the urea elutions. The peptide peaks present in these samples were more intense than those observed in the urea elutions suggesting that urea may not be eluting bound proteins from the affinity probe-streptavidin complex. The unique peptides were subject to MS/MS and the associated proteins identified using Mascot (Table 4.4).

Proteins from the Histone H2A family have been identified in three different samples (streptavidin beads, Method 3 and urea elutions, Methods 2 and 3). However, the identification has arisen from the same peptide ($m/z = 944$). The evidence is therefore not strong enough to confirm a hit. The other proteins identified from the streptavidin bead samples also have only one identified peptide and the Mascot scores of each were lower than the threshold of significance ($p < 0.05$), preventing confirmation of a hit.

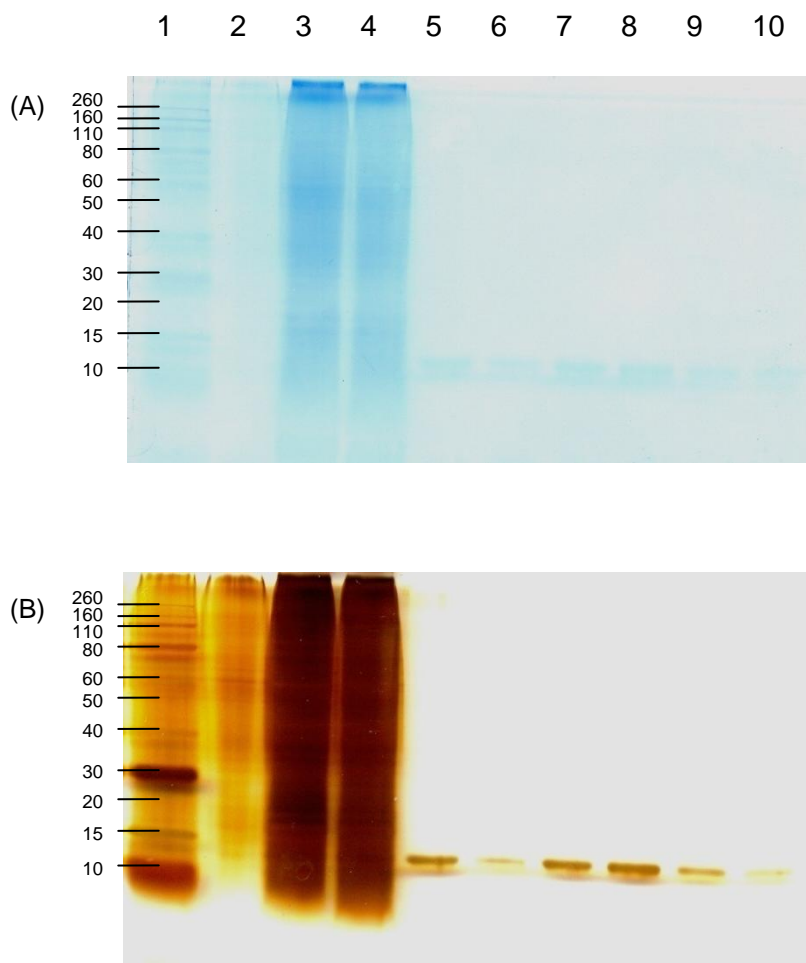


Figure 4.8. SDS-PAGE Analysis of Streptavidin beads. (A) Stained with Instant Blue stain; (B) Silver Stained.

Lane 1, Molecular weight marker; **Lane 2**, SDS buffer; **Lane 3**, total protein content from A2780 cell line prior to incubation with affinity probes; **Lane 4**, total protein content from HT29 cell line prior to incubation with affinity probes; **Lane 5**, experiment 1; **Lane 6**, experiment 2; **Lane 7**, experiment 3; **Lane 8**, experiment 4; **Lane 9**, experiment 5; **Lane 10**, experiment 6.

De novo sequencing was carried out from the MS/MS of two of the observed peptides, 1640.7730 and 3075.5880 as a hit was not confirmed using Mascot searching. This involved manually determining mass differences between fragments equal to one amino acid in weight (Figure 4.9). The sequences found were searched using the ExPASy proteomics server BLAST search engine and both were identified as streptavidin.

Table 4.4. Unique peptides on the streptavidin beads following affinity capture. **Search Parameters:** All species; Database – Swissprot; Enzyme – Trypsin; Mass Tolerance – 100 ppm; Variable modifications – ^aoxidation (M) & ^bcarboxyamidomethyl (C); Partial – 2.

Mass	Protein Identified	Mr _(exp)	Mr _(calc)	ppm	Miss	Score	Sequence
773.37	No protein identified						
831.40	No protein identified						
848.33	No protein identified						
859.36	No protein identified						
874.39	No protein identified						
899.40	No protein identified						
944.59	Histone H2A type 1-A	943.58	943.52	61.2	0	23	R.AGLQFPVGR.I
987.43	No protein identified						
1001.45	No protein identified						
1032.47	No protein identified						
1075.49	No protein identified						
1175.44	No protein identified						
1180.48	No protein identified						
1223.50	No protein identified						
1368.61	No protein identified						
1640.77	Ras association domain-containing protein 4	1639.77	1639.73	19.2	2	13	K.SDASCMSQRRPKCR.A ^a
1656.62	No protein identified						
1662.61	No protein identified						
1683.75	Protein orai-3 (Transmembrane protein 142C)	1682.74	1682.77	-13.95	0	24	R.SSASAAPSQAEPAECPQR.Q ^b
1697.64	No protein identified						
1699.75	No protein identified						
1715.77	No protein identified						
1833.92	Pleckstrin homology domain-containing G6	1832.91	1832.98	-37.13	2	14	R.RLQYYVFFARGSGQAR.G
1969.10	No protein identified						
2062.99	Misshapan-like kinase 1	2061.98	2062.01	-11.75	2	5	K.GQSPPSKDGSGDYQSRGLVK.A
2112.82	No protein identified						
2268.14	Kinesin-1 heavy chain	2267.13	2267.18	-19.81	0	20	K.SLSALGNVISALAEGSTVVPYR.D
2352.90	No protein identified						
2614.97	No protein identified						
3057.45	Peroxidasin homolog	3056.44	3056.62	-60.85	1	22	R.RCLLALVFCAWGTLAVVAQKPGAGCP SR.C ^b

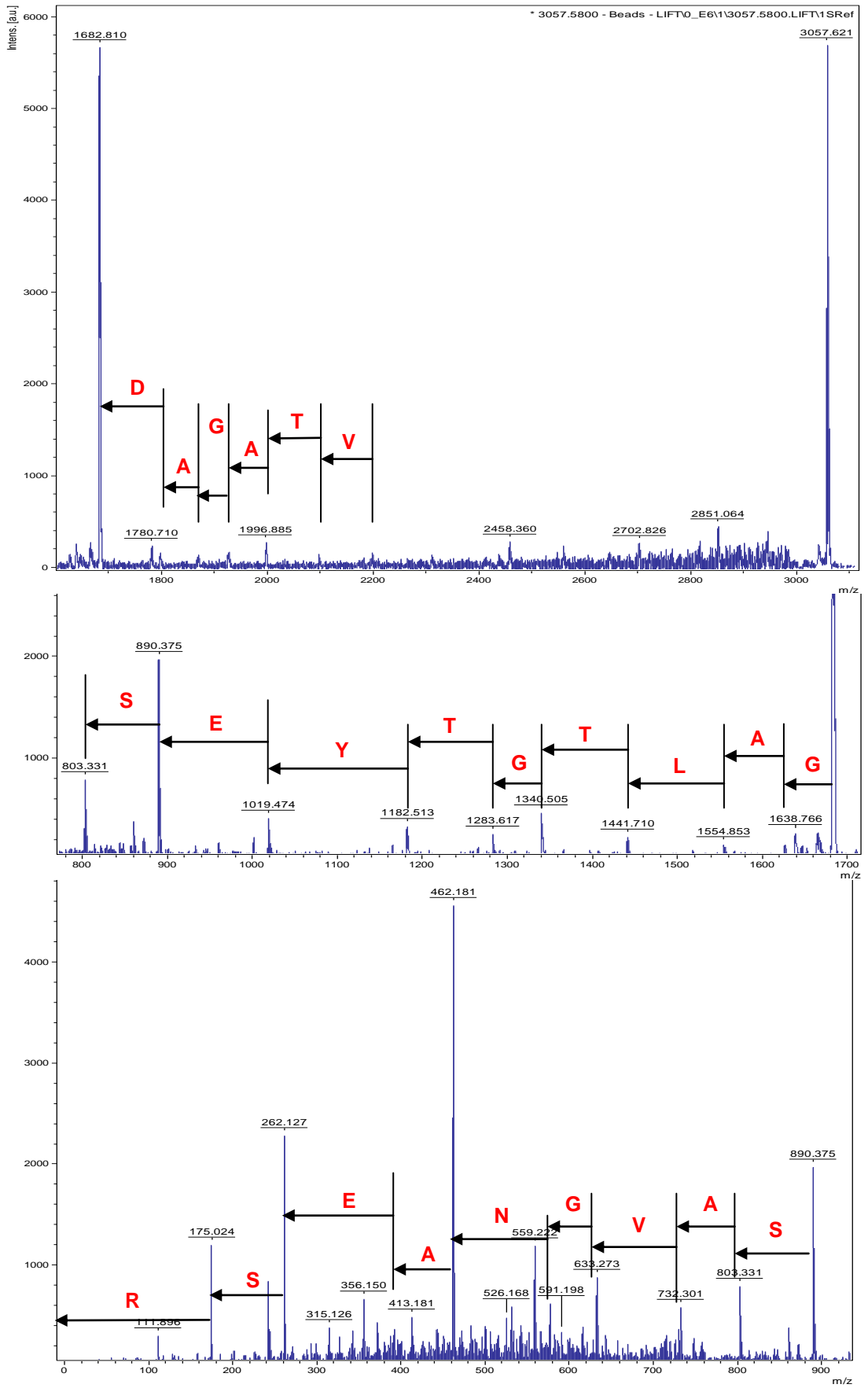


Figure 4.9. *De novo* sequencing of peptide m/z 3057.635.

SDS-PAGE and MS analysis for each method showed that the streptavidin protein is dominating the bead samples, with other proteins being less concentrated. A C18 ziptip was used to concentrate peptides which may belong to other proteins. Samples were eluted from the ziptip using increasing concentrations of acetonitrile in water. Although peptides were concentrated, MS and MS/MS analysis of each fraction did not identify any proteins in Mascot searching (Figure 4.10).

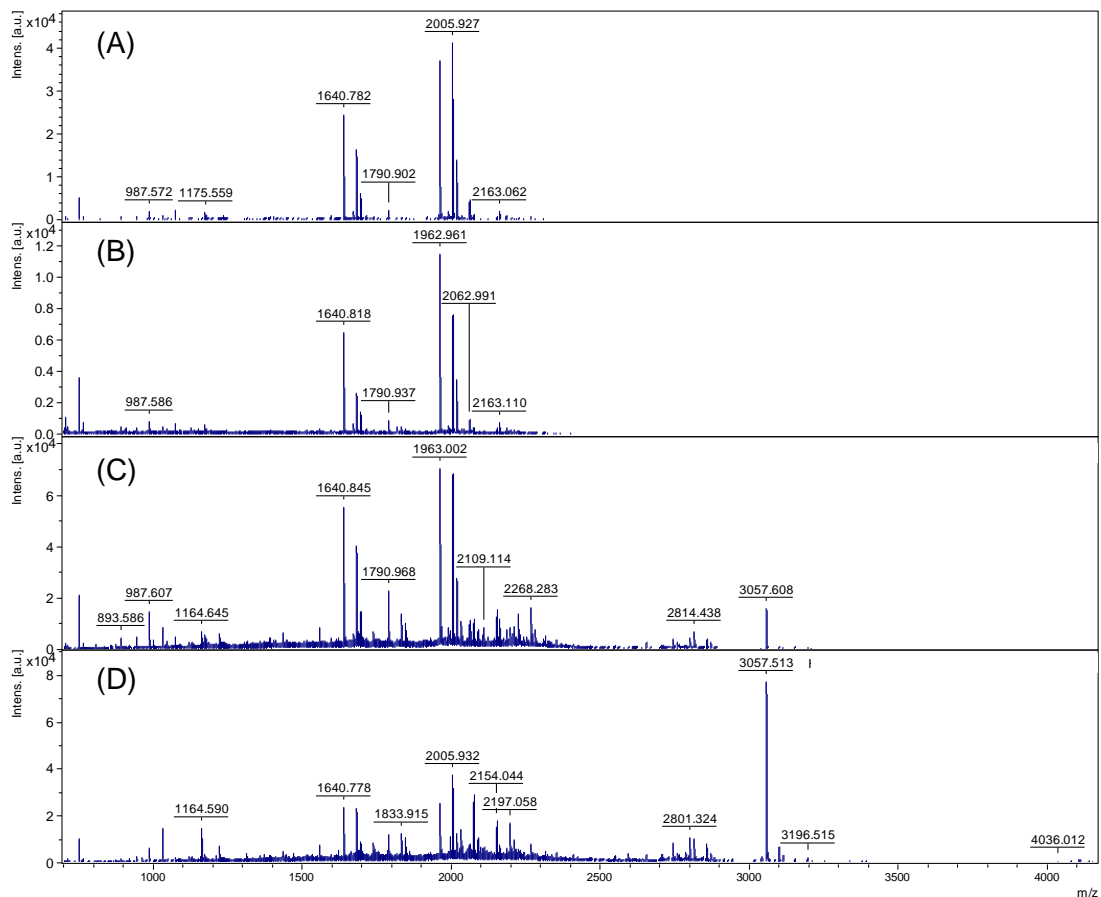


Figure 4.10. Mass spectra for peptides eluted with increasing concentrations of MeCN in H₂O. (A) 20 % MeCN; (B) 30 % MeCN; (C) 40 % MeCN; (D) 80 % MeCN.

Method 3 has resulted in more concentrated protein samples following affinity chromatography. Therefore, the method appears to be a more efficient approach to capturing binding proteins compared with Methods 1 and 2.

4.3.4 METHOD 4

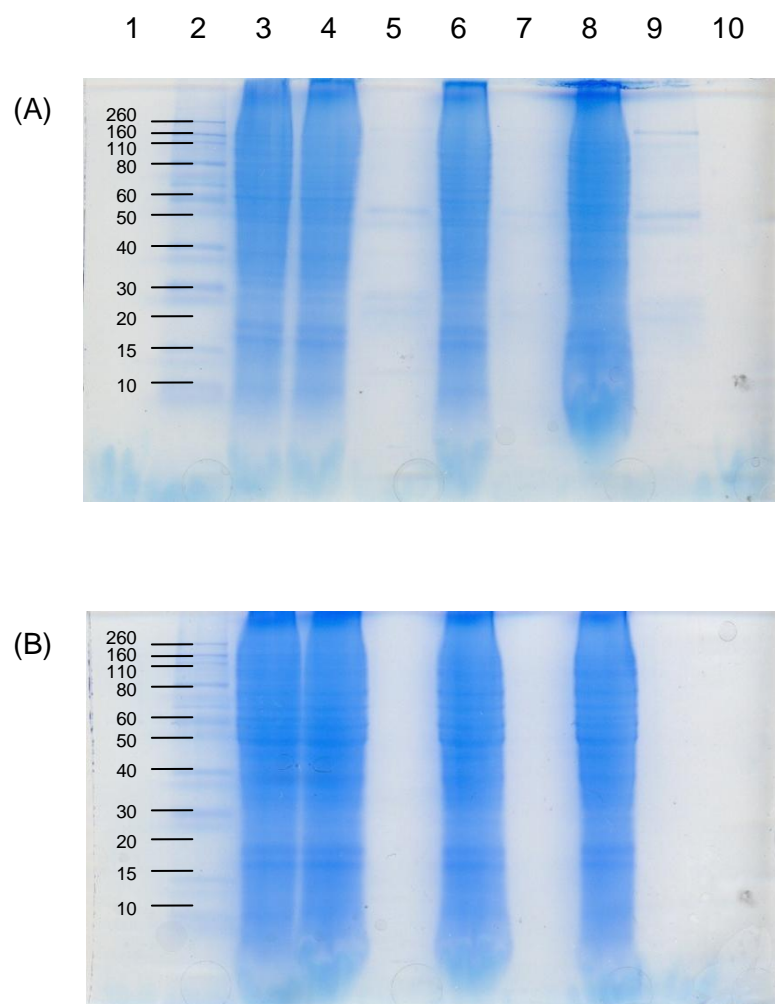
Method 4 followed the same path as Method 3, but was carried out on a larger scale. The volume of streptavidin beads (100 μ l) used remained constant, however, the number of cells (6.1×10^8) and the concentration of the affinity probes (1 mM) was increased. Incubation periods were also increased (24 h) and carried out at 4 °C instead of room temperature.

It was considered that the target protein may not be soluble in the Tris-HCl buffer containing 4 % CHAPS or that Tris was interfering with probe-protein interaction. Therefore, this method incorporated two solubilisations of the cell lysate prior to affinity capture. The first solubilisation used a PBS buffer containing 1 % Triton X and PIC and the second used PBS containing 1 % C7BzO, 4 % CHAPS and PIC. The experiments carried out are summarised in Table 4.5.

Table 4.5. Experiments performed using Method 4.

Experiment	Affinity Probe	Cell Line	Detergent
1	16	A2780	Triton X
2	17	A2780	Triton X
3	No compound	A2780	Triton X
4	16	HT29	Triton X
5	17	HT29	Triton X
6	No compound	HT29	Triton X
7	16	A2780	C7BzO + CHAPS
8	17	A2780	C7BzO + CHAPS
9	No compound	A2780	C7BzO + CHAPS
10	16	HT29	C7BzO + CHAPS
11	17	HT29	C7BzO + CHAPS
12	No compound	HT29	C7BzO + CHAPS

Bound proteins were eluted by competitive elution with hit compound **2** (2 mM, 24 h) instead of 8 M urea as urea elutions in the previous three methods did not lead to high concentrations of proteins. SDS-PAGE of the competition elutions (Figure 4.11) revealed proteins in experiments 1-3 (Figure 4.11A), representing incubation with A2780 cell lysates solubilised with Triton X. However, as these protein bands appear in control experiment 3, they were considered to have non-specific binding affinity. The gels from the remaining experiments failed to show the elution of any proteins.



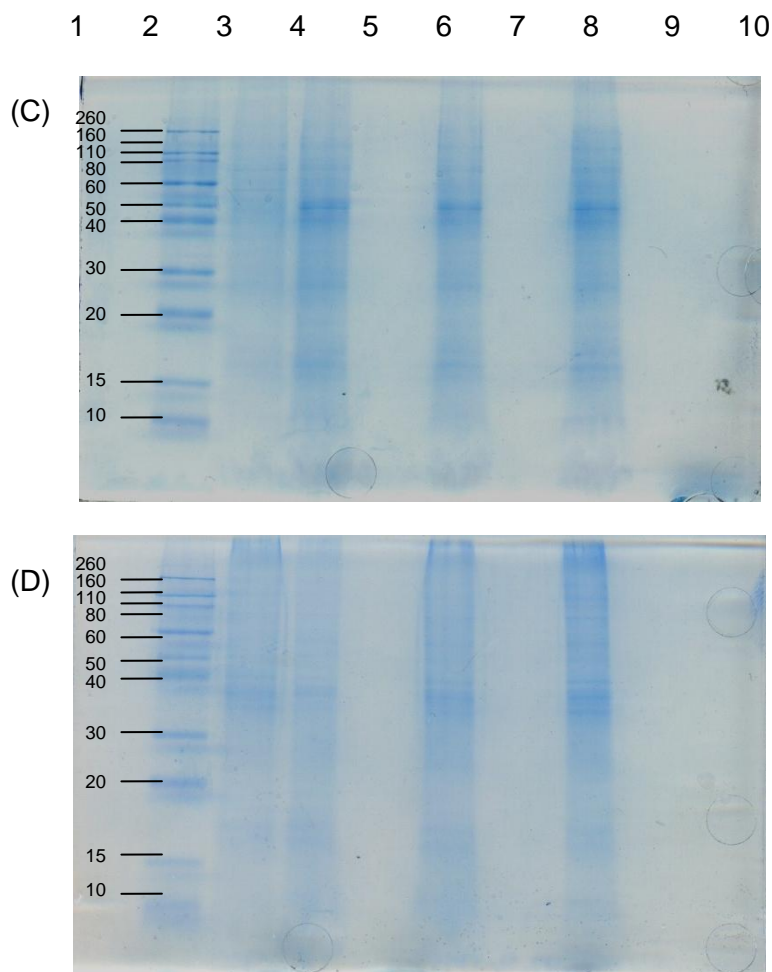


Figure 4.11. SDS-PAGE Analysis stained with Instant Blue stain of Competition washes with **2** from (A) The 1st solubilisation stage of cell lysates from the A2780 cell line; (B) The 1st solubilisation stage of cell lysates from the HT29 cell line; (C) The 2nd solubilisation stage of cell lysates from the A2780 cell line; (D) The 2nd solubilisation stage of cell lysates from the HT29 cell line.

Lane 1, SDS buffer; **Lane 2**, Molecular weight marker; **Lane 3**, total protein content prior to incubation with affinity probes; **Lane 4**, unbound proteins after incubation with **16**; **Lane 5**, bound proteins after incubation with **16**; **Lane 6**, unbound proteins after incubation with **17**; **Lane 7**, bound proteins after incubation with **17**; **Lane 8**, experiment 4; **Lane 9**, unbound proteins after incubation with no compound; **Lane 10**, bound proteins after incubation with no compound.

MS of the competition washes with hit compound **2** identified proteins based on their PMF (Table 4.6).

Table 4.6. Protein hits from the PMF of the hit compound **2** competition washes for each experiment.

Experiment	Protein	Score	Sequence Coverage
1	None identified		
2	Retinoid-inducible serine carboxypeptidase	38	15 %
3	Retinoid-inducible serine carboxypeptidase	44	22 %
4	None identified		
5	Histone H2A type 1-J	28	27 %
6	Epidermal growth factor-like protein 6	58	10 %
7	Alpha-enolase	45	9 %
8	RING finger protein 222	62	15 %
9	None identified		
10	None identified		
11	Epidermal growth factor-like protein 6	58	10 %
12	None identified		

Search Parameters: *Homo sapiens*; Database – Swissprot; Enzyme – Trypsin; Mass Tolerance – 100 ppm; Variable modifications – oxidation (M) & carboxyamidomethyl (C); Partial – 2.

Retinoid-inducible serine carboxypeptidase was identified in both experiment 2 and control experiment 3, therefore this result can be discarded as a non-specific binding protein. The same can be applied to the epidermal growth factor-like protein 6 seen in experiments 6 and 11. Experiment 5 identified another protein from the Histone H2A family. This result could now be considered as significant as it was identified from a PMF rather than MS/MS of the peptide found at $m/z = 944$. The sequence coverage for each of the identified proteins was relatively low and there was no protein identified from experiment 1, which would be expected if the target protein was captured.

The competitive elutions did not reveal any significant specific binding proteins, so it was presumed the target was still bound to the affinity column. This is possibly a result of the concentration of the hit compound **2** eluting solution being too low. Direct SDS-PAGE analysis of the streptavidin beads was therefore carried out (Figure 4.12).

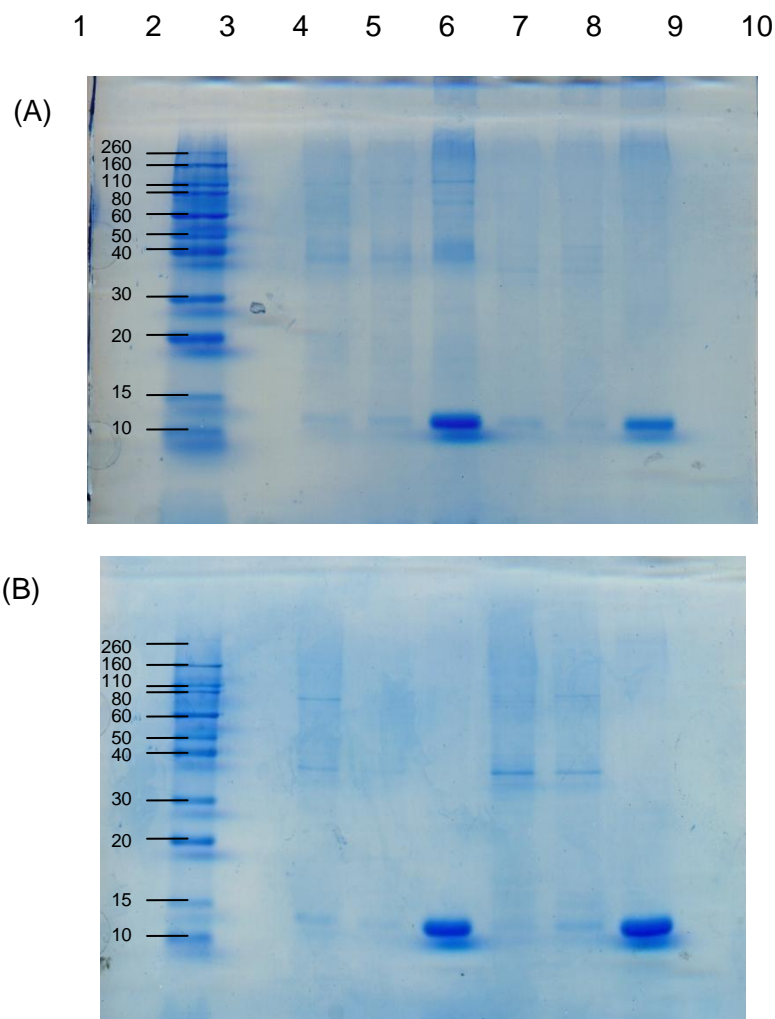


Figure 4.12. SDS-PAGE Analysis stained with Instant Blue stain of streptavidin beads from (A) The 1st solubilisation stage of cell lysates from both cell lines; (B) The 2nd solubilisation stage of cell lysates from both cell lines.

Lane 1, 3 & Lane 10, SDS buffer; **Lane 2**, Molecular weight marker; **Lane 4**, A2780 cell lysate incubated with **16**; **Lane 5**, A2780 cell lysate incubated with **17**; **Lane 6**, A2780 cell lysate incubated with no compound; **Lane 7**, HT29 cell lysate incubated with **16**; **Lane 8**, HT29 cell lysate incubated with **17**; **Lane 9**, HT29 cell lysate incubated with no compound.

Instant blue staining of the gels revealed of several bands. Instead of direct trypsin digestion of the streptavidin bead samples, in-gel trypsin digestion of the protein bands allowed separation from the streptavidin peptides which had

dominated the MS analysis during the previous three methods. The gels were silver stained to enable better visualisation of the proteins allowing for easier selection of bands to excise (Figure 4.13 and Figure 4.14). Once the bands had been identified, the SDS-PAGE was repeated and the bands indicated in Figure 4.13 and 4.14 were excised for digestion with trypsin.

The gels in Figure 4.13 and 4.14 showed a marked concentration of streptavidin in the control experiments compared with experiments which incorporated incubation with an affinity probe. This may be because on binding of biotin to streptavidin, a surface loop of the streptavidin protein folds over the binding site.¹²² The complex is more stable, preventing denaturation and dissociation from the agarose beads during SDS-PAGE sample preparation. This visual marker could act as an indicator that the affinity probes have bound to the streptavidin beads in future experiments.

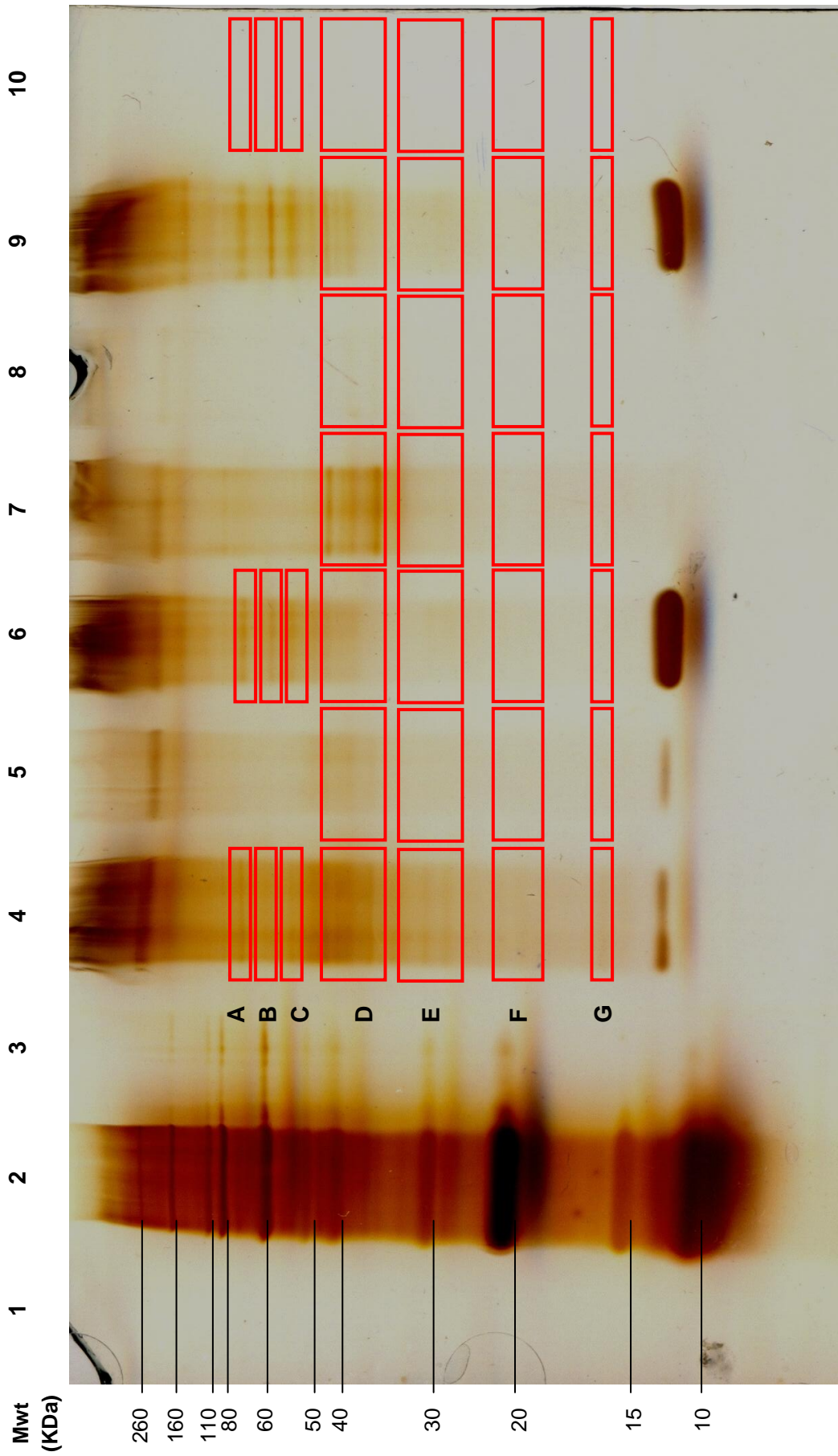


Figure 4.13. Silver stain of Figure 4.12A showing the bands excised for in-gel trypsin digestion.

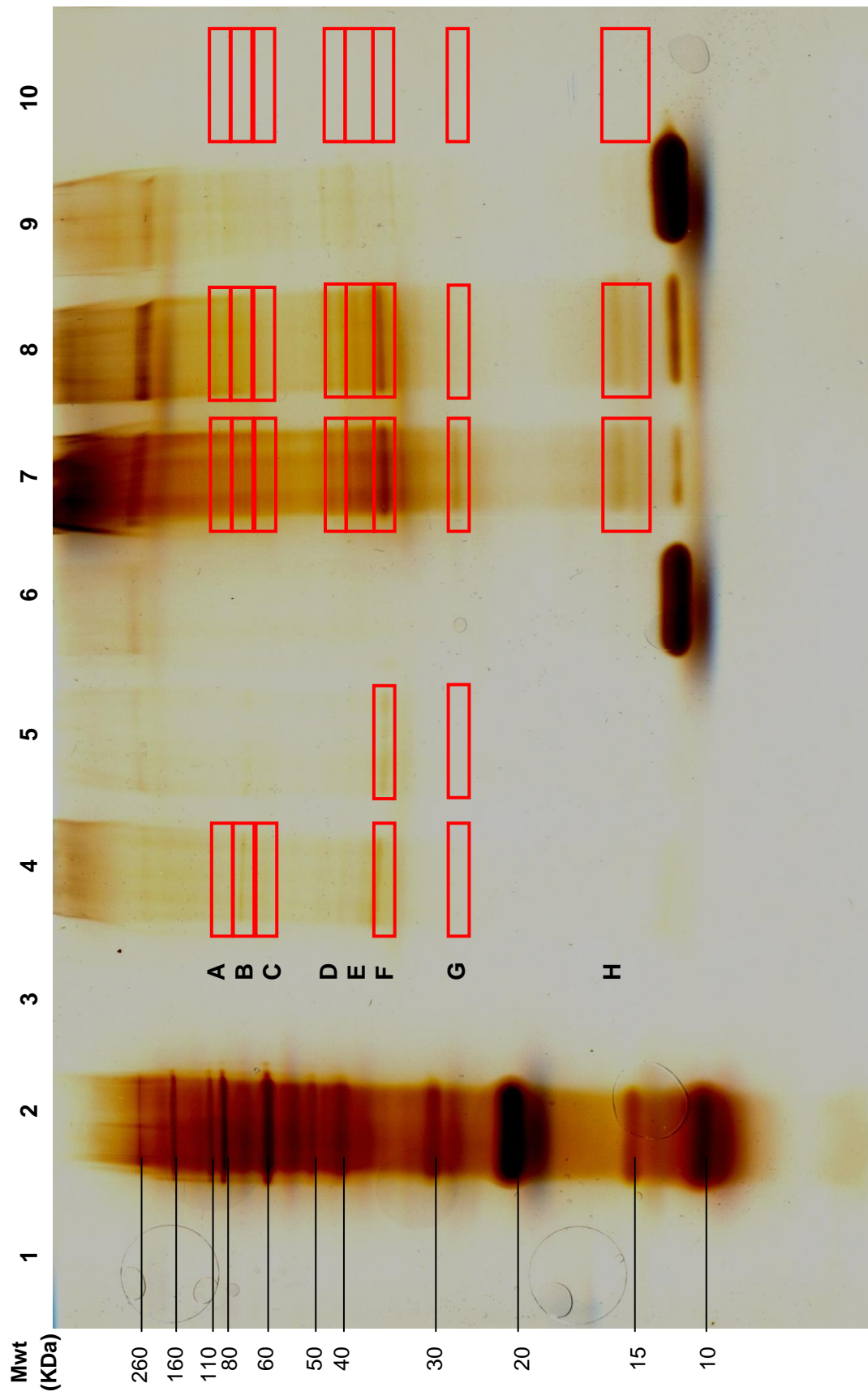


Figure 4.14. Silver stain of Figure 4.12B showing the bands excised for in-gel trypsin digestion.

The excised bands were trypsin digested and subject to MS analysis. The peptides shown in each spectrum were compared and unique peptides (those which were not present in lane 10) identified. The unique peptide lists were searched using the PMF feature of the Mascot database and identified proteins are summarised in Tables 4.7 and 4.8.

Comparisons of the PMF showed the highest numbers of unique peptides were present in band D of the gel from the first solubilisation (Figure 4.11A) and bands E and F of the gel from the second solubilisation (Figure 4.11B). The higher peptide numbers resulted in confirmed hits during database searching. Heterogeneous nuclear ribonuclear protein A2/B1 (hnRNP A2/B1) was identified in all of the bands outlined above. This result strongly suggests that the protein is binding to the affinity probes, **16** and **17**. In addition the sequence molecular weights of the proteins correspond to the molecular weights of the bands.

Cytoskeletal keratin type 1 was identified in band A lane 7 of the gel representing the second solubilisation (Figure 4.11B). Keratins are common contaminants in proteomic analysis and are probably not binding specifically to the affinity probe. The other proteins identified from each gel were below the threshold of protein confirmation ($p < 0.05$) due to their low Mascot scores and the theoretical molecular weights of the proteins identified did not correspond to the molecular weights of the bands in the SDS-PAGE gel. Therefore these identified proteins were ignored.

Table 4.7. Highest ranking protein identifications from unique bands in the PMF of bands excised from gel shown in Figure 4.11A. Blank rows indicate too few unique peptides to generate a hit. Red proteins scored a hit above the threshold. In brackets: number of peptides, sequence coverage and Mascot score. In blue, the theoretical molecular weight of the identified protein.

	4 (16, A2780)	5 (17, A2780)	7 (16, HT29)	8 (17, HT29)
A	None identified	None identified	None identified	None identified
B	None identified	None identified	None identified	None identified
C	None identified	None identified	None identified	None identified
D	Heterogeneous nuclear ribonucleoproteins A2/B1 (3, 16%, 32) B1 = 36,006 Da, A1 = 37,430 Da	Heterogeneous nuclear ribonucleoproteins A2/B1 (3, 16%, 32) B1 = 36,006 Da, A1 = 37,430 Da	Heterogeneous nuclear ribonucleoproteins A2/B1 (5, 28%, 56), B1 = 36,006 Da, A1 = 37,430 Da	None identified
E	None identified	None identified	Abhydrolase domain-containing protein 12B (4, 16%, 43) 40,776 Da	Aspartyl-tRNA synthetase, mitochondrial (4, 6%, 37) 73,563 Da
F	N-acetyl-beta-glucosaminyl-glycoprotein 4-beta-N-acetylgalactosaminyltransferase (6, 7%, 38) 114,975 Da	None identified	None identified	None identified
G	Putative 3-phosphoinositide-dependent protein kinase (5, 16%, 38) 44,765 Da	AAH06133 (3, 15 %, 14) 113,678 Da	None identified	None identified

Search Parameters: *Homo sapiens*; Database – Swissprot; Enzyme – Trypsin; Mass Tolerance – 100 ppm; Variable modifications – none; Partials – 2.

Table 4.8. Highest ranking protein identifications from unique bands in the PMF of bands excised from gel shown in Figure 4.11B. Blank rows indicate too few unique peptides to generate a hit. Red proteins scored a hit above the threshold. In brackets: number of peptides, sequence coverage and Mascot score. In blue, the theoretical molecular weight of the identified protein.

	4 (16, A2780)	7 (16, HT29)	8 (17, HT29)
A	Growth/differentiation factor 2 (5, 14%, 36) 47,320 Da	Keratin, type I cytoskeletal 16 (12, 26%, 89), 51,268 Da	Butyrophilin-like protein 9 (4, 10%, 36) 59,716 Da
B	Toll-like receptor 9 (8, 13%, 31), 115,860 Da	Protein APCDD1 (5, 13%, 41) 58,797 Da	Junctophilin-3 (10, 14%, 43) 81,469 Da
C	Probable N-acetyltransferase 8 (3, 22%, 57), 25,619 Da		Protein S100-A8 (2, 44%, 28) 10,835 Da
D	None identified	Cysteine and glycine-rich protein 3 (4, 30%, 39) 20,969 Da	Zinc finger protein 174 (5, 13%, 36) 46,455 Da
E	None identified	Heterogeneous nuclear ribonucleoproteins A2/B1 (8, 39%, 79) B1 = 36,006 Da, A1 = 37,430 Da	Heterogeneous nuclear ribonucleoproteins A2/B1 (5, 22%, 48) B1 = 36,006 Da, A1 = 37,430 Da
F	Heterogeneous nuclear ribonucleoproteins A2/B1 (7, 32%, 86) B1 = 36,006 Da, A1 = 37,430 Da	Heterogeneous nuclear ribonucleoproteins A2/B1 (12, 45%, 103) B1 = 36,006 Da, A1 = 37,430 Da	Heterogeneous nuclear ribonucleoproteins A2/B1 (13, 47%, 113) B1 = 36,006 Da, A1 = 37,430 Da
G	None identified	Protein LZIC (3, 20%, 22) 21,495 Da	PDZ domain-containing protein 7 (12, 14, 34) 55,677 Da
H	None identified	None identified	Protein PAT1 homolog 1 (5, 5%, 39) 86,850 Da

Search Parameters: *Homo sapiens*; Database – Swissprot; Enzyme – Trypsin; Mass Tolerance – 100 ppm; Variable modifications – none; Partials – 2.

The strongest unique peptide signals, for each band were subject to MS/MS.

Table 4.9 and Table 4.10 show the peptides with a Mascot score above 25.

Table 4.9. Protein identification by Mascot searching of MS/MS spectra of the most abundant peptides from the MS spectra of each band in the gel shown in Figure 4.13. Showing protein identifications with a Mascot score > 25.

Band	Mass	Protein Identified	Accession Number	Mr (expt)	Mr (Calc)	ppm	Miss	Score	Sequence	Protein Mwt
A	1125.57	Vimentin	VIME_HUMAN	1124.57	1124.60	-29	1	27	R.FANYIDKVR.F	53,619
B	1305.76	ATP synthase subunit epsilon-like protein, mitochondrial	AT5EL_HUMAN	1304.75	1304.75	2	2	25	K.VVRDALKTEFK.A	5,803
B	1567.62	Serum Albumin	ALBU_HUMAN	1566.61	1566.71	-82	0	42	K.DAFLGFLYEYSR.R	69,248
D	1377.51	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1376.50	1376.62	-87	0	26	R.GGGNFGPGGNSFR.G	37,407
D	1539.80	Vimentin	VIME_HUMAN	1538.79	1538.90	-71	1	86	K.ILLAELEQLKGGK.S	53,619
F	1377.56	Thymosin beta-15A	TB15A_HUMAN	1376.56	1376.65	-67	0	26	-MSDKPDLSEVEK.F	5,226
E	1539.80	E3 ubiquitin-protein ligase	TR133_HUMAN	1538.79	1538.94	-95	2	38	K.GAIENLLAKLLEKK.N	122,444
E	1918.87	Heterogeneous nuclear ribonucleoprotein H3	HNRH3_HUMAN	1917.86	1917.96	-52	0	32	R.ATENDIANFFSPLNPIR.V	36,909

Table 4.10. Protein identification by Mascot searching of MS/MS spectra of the most abundant peptides from the MS spectra of each band in the gel shown in Figure 4.14. Showing protein identifications with a Mascot score > 25.

Band	Mass	Protein Identified	Accession Number	Mr (expt)	Mr (Calc)	ppm	Miss	Score	Sequence	Protein Mwt
A	1283.68	Metallothionein-1E	MT1E_HUMAN	1282.67	1282.55	98	1	25	K.CAQQGCVCKGASEK.C	6,009
A	1407.68	Keratin, type II cytoskeletal 6A	K2C6A_HUMAN	1406.67	1406.70	-23	0	40	K.ADTLTDEINFLR.A	60,008
A	2151.06	Keratin, type I cytoskeletal	K1C16_HUMAN	2150.05	2150.09	-18	1	56	R.TDLEMQIEGLKEELAYLR.K	51,236
B	1640.09	Serum albumin	ALBU_HUMAN	1639.08	1638.93	90	1	56	K.KVPQVSTPTLVEVSR.N	69,321
D	1996.88	Heterogeneous nuclear ribonucleoprotein F	HNRPF_HUMAN	1995.87	1995.97	-49	0	25	K.ATENDIYNFFSPLNPVR.V	45,643
E	1377.65	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1376.64	1376.62	-11	0	36	R.GGGNFGPGGSGNFR.G	37,497
E	1410.73	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1409.72	1409.68	31	0	27	K.YHTINGHNAEVR.K	37,497
E	1798.94	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1797.94	1797.91	12	0	25	K.LFIGGLSFETTEESLR.N	37,497
F	1377.62	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1376.61	1376.62	-11	0	69	R.GGGNFGPGGSGNFR.G	37,497
F	1410.59	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1409.58	1409.68	-73	0	39	K.YHTINGHNAEVR.K	37,497
F	1798.81	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1797.81	1797.91	-60	0	47	K.LFIGGLSFETTEESLR.N	37,497
F	1918.86	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1917.85	1917.96	-57	0	45	R.ATENDIANFFSPLNPVR.V	37,497
F	1926.90	Heterogeneous nuclear ribonucleoproteins A2/B1	ROA2_HUMAN	1926.89	1926.01	-61	1	25	R.KLFIGGLSFETTEESLR.N	37,497

4.3.5 HETEROGENEOUS NUCLEAR RIBONUCLEAR PROTEIN A2/B1

Heterogeneous nuclear ribonuclear proteins (hnRNPs) are a family of proteins which share common structural motifs. They are some of the most abundant nuclear proteins with a vast array of functions.¹⁶⁴ These multi-tasking proteins play a central role in RNA metabolism, however they have also been reported to have roles in telomere biogenesis, cell signalling and regulation of gene expression at transcription and translational levels, although they lack any direct enzymatic activity.¹⁶⁵

hnRNP A2 and its spliced isoform B1 are major components of the 40S particles that package RNA. They have roles in addition to packaging hnRNA including alternative RNA splicing, mRNA export from the nucleus and cytoplasmic trafficking, stability and translation.¹⁶⁶ hnRNP A2/B1 are increased at many types of cancer – lung, breast, pancreatic, stomach – hnRNP B1 is increased in oral and oesophagus cancers along with leukaemia and lymphoma.¹⁶⁵ These increases can be linked to their role in telomere regulation and also in cell proliferation.

Levels of hnRNP A2/B1 have been determined at different stages of the cell cycle showing that they peak during the G1 and S phase and decline during the G2 and M phase. This has a direct effect on cell cycle regulators, decreasing expression of BRAC1 and increasing p21, as a result it has been suggested that the proteins may be classed as oncogenic development proteins. hnRNP A1/B2 bind to the single-stranded telomeric repeat sequence. SiRNAi

experiments to reduce the protein expression in HeLa cells led to a change in the distribution in the lengths of the telomere G-tails. However, the functional significance of this binding on regulation and maintenance of telomere structure has not been reported.¹⁶⁶

As hnRNP A2/B1 form large complexes with many proteins, it is uncertain whether these are the unique protein targets for the hit compounds **1**, **2** and **3**, or that they are masking the binding of another protein. Figure 4.15 shows the proteins associated with hnRNP A2/B1 in the ribonucleosome using the STRING database.¹⁶⁷

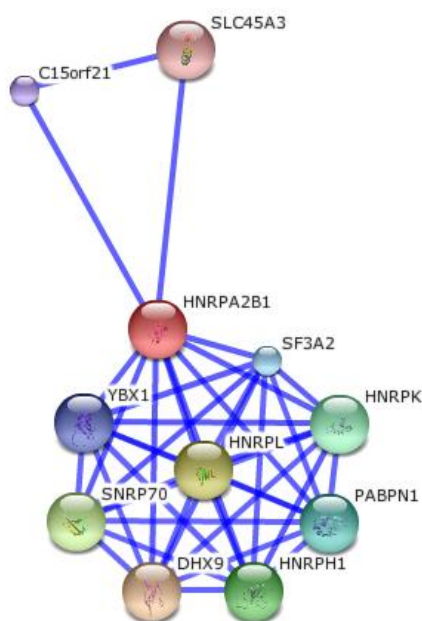


Figure 4.15. STRING network display of hnRNP A2/B1, showing proteins which have a known association.

4.3.6 CONCLUSIONS

The proteomic approach to target identification was successful in identifying proteins using affinity probes **16** and **17**. Methods 1-3 failed to capture any binding proteins, possibly due to the scale of the procedure. Histone H2A proteins were identified from MS/MS spectra from different sources, however, the peptide ($m/z = 944$) was the same in each identification.

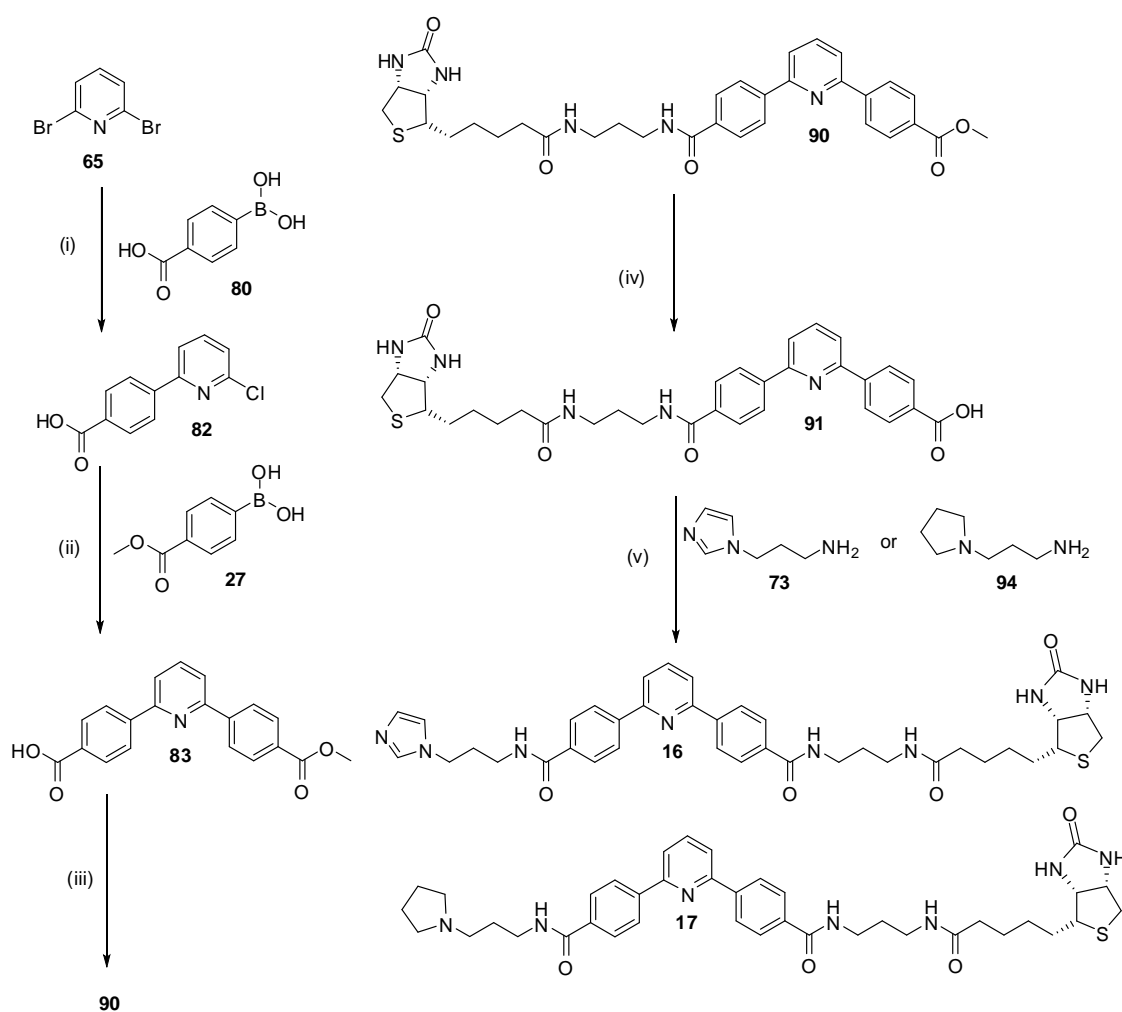
Method 4 captured many proteins, the majority of which were successfully identified as heterogeneous nuclear ribonuclear protein A2/B1. This may be attributed to the change of detergent used to solubilise the proteins from the cell lysate and the larger scale of the method. Most of the peptides were present in samples from affinity experiments using a detergent of C7BzO and CHAPS instead of CHAPS alone, which was used in previous methods.

CHAPTER 5

CONCLUSIONS AND FURTHER WORK

5.1 CONCLUSIONS

Non-symmetrical biarylheterocycles containing a biotin group were designed, synthesised and used as affinity probes in chemical proteomics. Many synthetic routes were explored and the final synthesis of affinity probes **16** and **17** is shown in Scheme 5.1.



Scheme 5.1. Synthesis of affinity probes **16** and **17**.

Reagents and Conditions: (i) Na_2CO_3 (0.4 M (aq)):MeCN (1:1), $\text{Pd}(\text{PPh}_3)_4$, 90 °C, N_2 , 24 h (93 %); (ii) K_2CO_3 , $\text{Pd}(\text{PPh}_3)_4$, PhMe:MeOH (9:1), Δ , 48 h (86 %); (iii) **88**, PyBOP, Et_3N , DMF, RT, 18 h (90 %); (iv) NaOH (10 eq), 50 °C, 2 h (46 %); (v) PyBOP, Et_3N , DMF, RT, 18 h (70-89 %).

A method of affinity capture has been established and successful in capturing a number proteins, identified from Mascot searches of PMFs and MS/MS ion searches. Most proteins had Mascot scores below the threshold ($p < 0.05$) and no correlation with the calculated molecular weights of the identified proteins, however, one significant protein was identified. This protein was identified as heterogeneous nuclear ribonucleoprotein A2/B1 (hnRNP A2/B1) which bound to both affinity probes **16** and **17**. The identifications are from Mascot searching of PMFs and MS/MS ion searches, leading to strong evidence that hnRNP A2/B1 is selectively binding to the affinity probes and not binding non-specifically to the streptavidin beads.

The potency of affinity probe **16** was reduced by bearing a single imidazole group, suggesting that both imidazole rings are required for binding. The half compound **4** shows the same selectivity as the hit compounds **1-3** but with reduced potency, implying that the target could contain a symmetrical binding site. Although there is not a complete 3D structure of the hnRNP A2/B1 protein available, the primary sequence information confirms that there are two RNA recognition motifs.¹⁶⁸ Each of these could bind to half of the hit compound or be binding two different RNAs which in turn are complexed to the hnRNP A2/B1. In addition, the amount of protein captured using affinity probes cannot be directly related to chemosensitivity data.

The distinction in activity observed between the two different cell lines, A2780 and HT29, with compound **16** was also lost during affinity capture and the protein was identified in both of the cell lines. It is possible that the target protein

only has a direct relation to the mechanism of cancerous growth in the A2780 cell line, as the chemosensitivity assay showed a contrast for the affinity probe **16** between the A2780 ($IC_{50} = 6.25 \mu\text{M}$) and HT29 ($IC_{50} > 100 \mu\text{M}$) cell lines. hnRNP A2/B1 is present in normal cells but found at increased levels in cancerous cells, suggesting that the presence of the protein may not be directly related to activity.

Although there is strong evidence to relate hnRNP A2/B1, or a protein complexed to hnRNP A2/B1, with binding to the affinity probes, the loss of potency in comparison to the hit compounds and the loss of contrast between the 'active' and the control affinity probe, has lead to uncertainty whether hnRNP A2/B1 is the target protein responsible for the potent activity of the hit compounds **1-3**. Further work is required to confirm and validate the target protein. Also, development of affinity probes with biotin attached to alternative sites of the hit compounds must be investigated.

5.2 TARGET VALIDATION

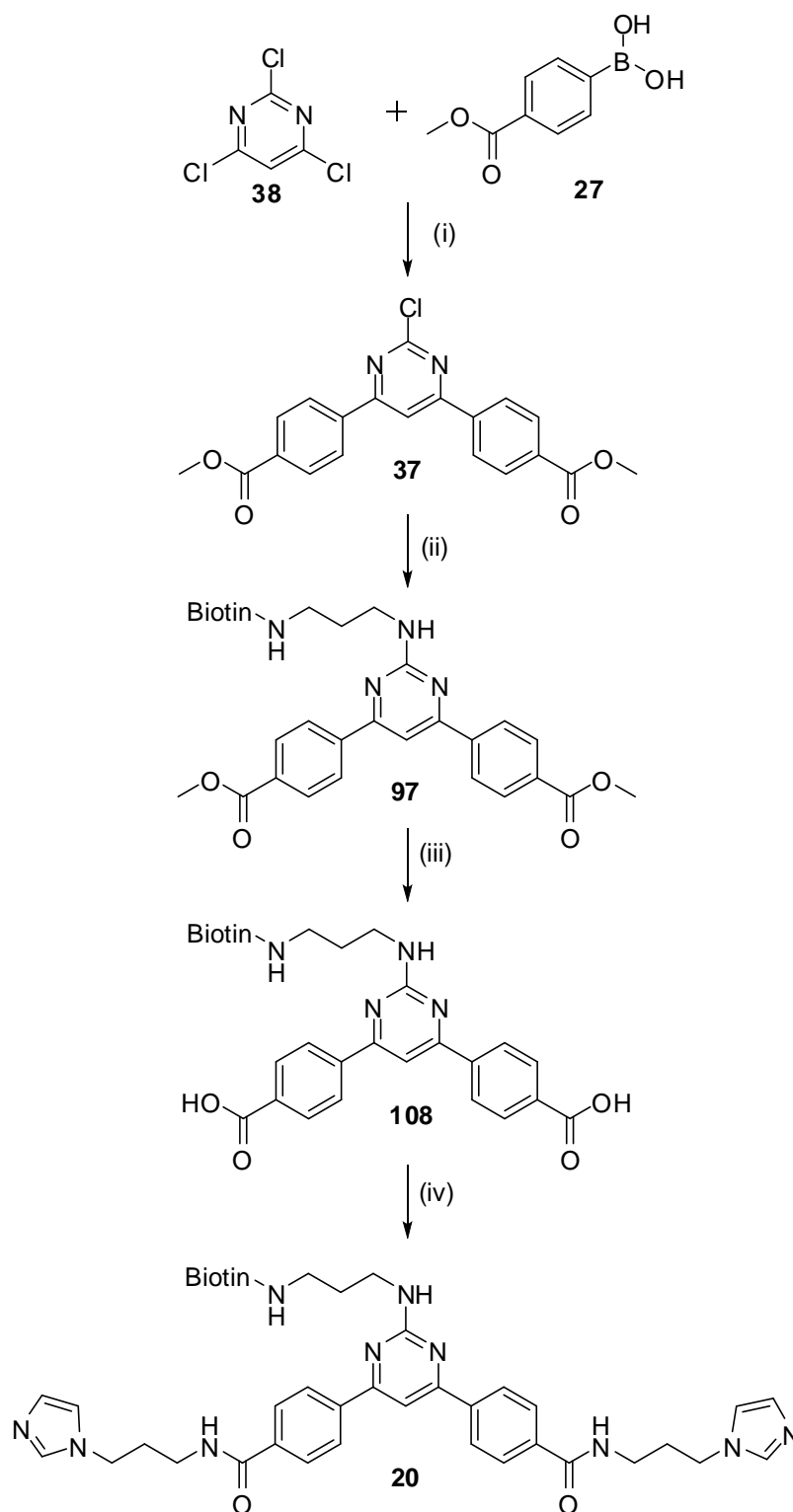
The NCI COMPARE data showed the H522 NSC lung cancer cell was the most sensitive to the hit compounds. Therefore, affinity chromatography using the affinity probes should be carried out using this cell line to confirm hnRNP A2/B1 as the biological target. In addition, western blotting can be used to determine the levels of hnRNP A2/B1 in the two tested cell lines, A2780 and HT29.

As another method of validation, affinity chromatography should be carried out in the presence of hit compound **2**. If the binding proteins are hnRNP A2/B1, the incubation of the total cell lysate with compound **2** prior to affinity capture with the affinity probes complexed to streptavidin, would result in none of the protein becoming captured.

5.3 SYNTHESIS OF ALTERNATIVE AFFINITY PROBES

Due to the loss of potency of the affinity probes, synthesis of compounds with an alternative attachment site must be explored. Progress has been made in the synthesis of pyrimidine compounds with the linker at the pyrimidine-2 position. However, the Ullmann copper-catalysed reactions only proceeded using a solvent of DMSO which led to purification problems for carboxylic acids **36**. Synthesis of compounds with the linker at the pyrimidine-2 position could be formed using palladium catalysed Buchwald-Hartwig amination reactions which proceed in organic solvents which can be easily removed.¹⁶⁹

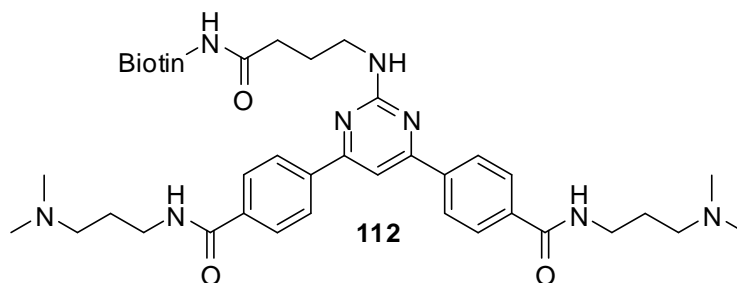
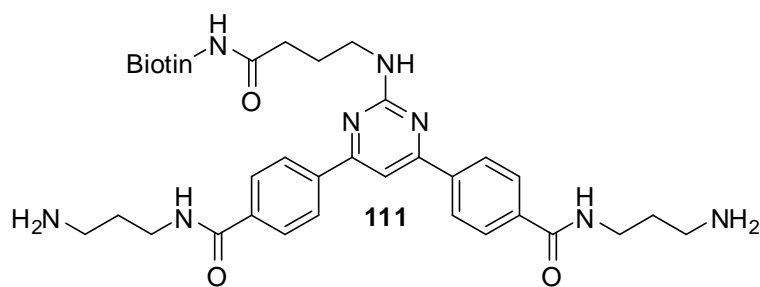
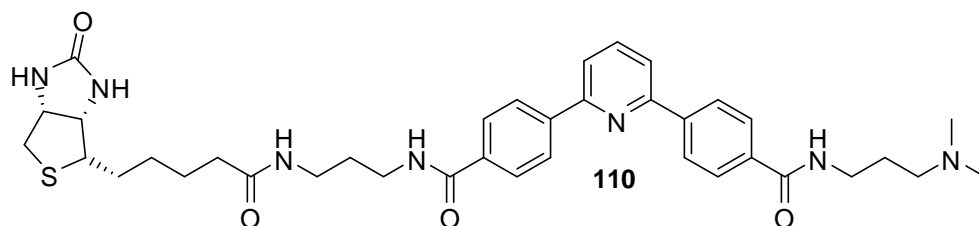
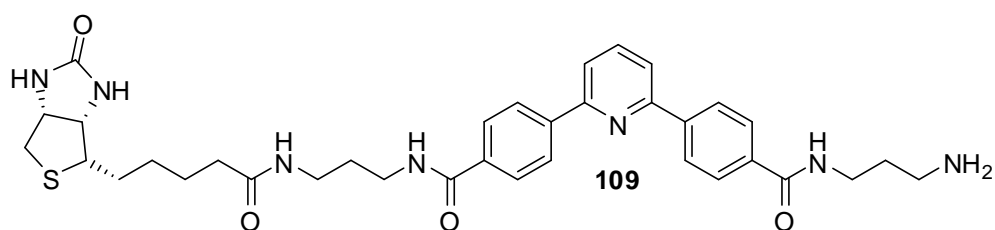
Also, hydrolysis of esters containing the Boc protected linker led to premature removal of the protecting group. Attaching the biotinylated linker, **88**, in place of the Boc-protected linker, **71**, to afford compound **20**, would overcome this problem. (Scheme 5.2).



Scheme 5.2. Possible synthesis of biotinylated affinity probe **20**.

Reagents and Conditions: (i) Na_2CO_3 , $\text{Pd}(\text{OAc})_2$, PPh_3 , glyme, Δ , 18-24 h; (ii) **77**, $\text{Pd}(\text{OAc})_2$, XPhos , $t\text{-BuONa}$, PhMe , 90°C , 4 h; (iii) NaOH (2M), Δ , 24 h, then HCl (4 M); (iv) **62**, PyBOP , Et_3N , DCM , RT , 18 h.

The contrast between the control compound **17** and the 'active' affinity probe **16** was lost. Therefore, a less potent control compound is required. The SAR carried out by Phillips has shown that replacing the imidazole ring with an aliphatic or free amine results in no activity. Therefore, the control affinity probes should be **109** and **110** for non-symmetrical compound **16**, and **111** and **112** for compounds with the biotin linker at the pyrimidine-2 position, **20**.



EXPERIMENTAL

EXPERIMENTAL DETAILS FOR CHAPTER 3: SYNTHESIS

REAGENTS

Boronic acids were purchased from Frontier Scientific, Europe Ltd., Carnforth, Lancashire, UK, other reagents were purchased from Sigma-Aldrich, Gillingham, UK or Lancaster, Morecambe, UK. Solvents were purchased from Riedel-de Haën.

INSTRUMENTATION

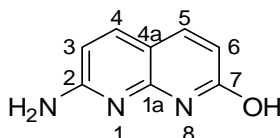
TLC was performed on highly purified silica gel plates with UV indicator (silica gel 60 F₂₅₄), manufactured by Merck and visualised under UV light (254 or 366 nm). Flash chromatography was performed using Silica gel (particle size 40-63 µm), purchased from Merck and followed the methodology described by Still *et al.*¹⁷⁰ Melting points were determined using an Electrothermal IA9200 digital melting point apparatus. Infrared data were obtained as KBr discs on a Perkin Elmer (Paragon 1000) FT-IR spectrophotometer or as an ATR on a Perkin Elmer (Spectrum 100) FT-IR spectrophotometer. Nuclear magnetic resonance spectra were acquired on a Bruker DMX400 spectrometer (observing ¹H at 400.13 MHz and ¹³C at 100.62 MHz) or a Jeol ECA600 spectrometer (observing ¹H at 600.17 MHz and ¹³C at 150.91 MHz). ¹³C assignments were made with the aid of the DEPT135 experiment. Mass spectra were obtained from the University of Bradford Analytical Centre, UK.

Elemental analysis was performed by the Advanced Chemical and Materials Analysis Unit at the University of Newcastle upon Tyne, UK.

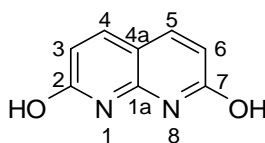
EXPERIMENTAL DETAILS FOR SECTION 3.2

Synthesis of 2,7-dichloronaphthyridine

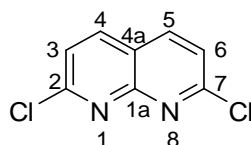
2-Amino-7-hydroxy-1,8-naphthyridine¹⁵²



2,6-Diaminopyridine (3.3 g, 30.2 mmol) and malic acid (4.5, 33.3 g) were ground to a fine powder and cooled in an ice bath. Concentrated sulphuric acid (15 ml) was added dropwise over 10 min and the mixture heated at 110 °C for 5 h. The reaction mixture was poured over ice and basified with ammonia solution. The precipitate was filtered to yield pure product (4.86 g, 100 %). m.p. decomp > 300 °C (lit. m.p > 350 °C); ¹H NMR (DMSO) δ 11.82 (br s, 1H, OH), 7.66 (d, *J* = 8.8 Hz, 2H, 4-H, 5-H), 7.00 (s, 2H, NH₂), 6.35 (d, *J* = 8.8 Hz, 1 H, 6-H), 6.12 (d, *J* = 8.8 Hz, 1H, 3-H); ¹³C NMR (DMSO) δ 164.2 (C-7), 161.0 (C-2), 150.6 (C-1a), 140.2 (C-4 or C-5), 137.8 (C-4 or C-5), 115.3 (C-6), 105.6 (C-3), C-4a not observed.

2,7-Dihydroxy-1,8-naphthyridine¹⁵²

2-Amino-7-hydroxy-1,8-naphthyridine (4.70 g, 29.2 mmol) was added to concentrated sulphuric acid (40 ml). Sodium nitrite (2.4 g) was added and the mixture was poured over ice. The solution was neutralised with sodium carbonate and the acidified with glacial acetic acid to pH 3. The resultant precipitate was collected by filtration (3.01 g, 64 %). m.p. decomp > 300 °C (lit. m.p 321-323 °C); ¹H NMR (DMSO) δ 11.59 (br s, 2H, OH), 7.82 (d, *J* = 8.8 Hz, 2H, 4-H, 5-H), 6.35 (d, *J* = 8.8 Hz, 2H, 3-H, 6-H); ¹³C NMR (DMSO) 163.8 (C-2, C-7), 139.9 (C-4, C-5), 120.6 (C-3, C-6), C-1a and C-4a not observed.

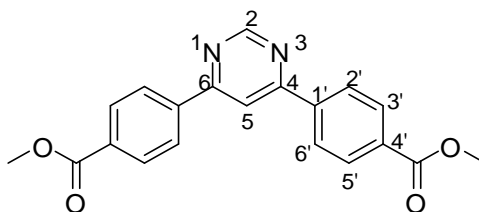
2,7-Dichloro-1,8-naphthyridine, 66¹⁵²

2,7-Dihydroxy-1,8-naphthyridine (2.5 g, 15.4 mmol), phosphorus pentachloride (6.45 g, 30.8 mmol) and phosphoryl chloride (4.72, 30.8 mmol) were heated under reflux for 2 h. Ice was added to the solution and basified to pH 8 with sodium carbonate. The resultant precipitate was collected by filtration and recrystallised from acetone to yield pure **66** (2.49 g, 81 %). m.p. decomp

256.6 – 257.3 °C (lit. sublimation point 258 °C); ^1H NMR (DMSO) δ 8.63 (d, $J = 8.6$ Hz, 2H, 4-H, 5-H), 7.80 (d, $J = 8.6$ Hz, 2H, 3-H, 6-H); ^{13}C NMR (DMSO) δ 154.2 (C-2, C-7), 141.7 (C-4, C-5), 124.6 (C-3, C-6), C-1a and C-4a not observed.

General Method 1: The Suzuki coupling Reaction

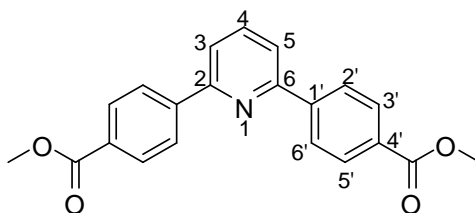
4,6- Bis-(4-methoxycarbonylphenyl) pyrimidine, **25**⁶²



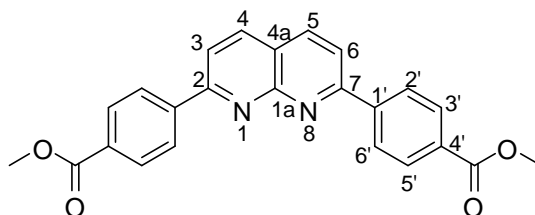
4,6-Dichloropyrimidine, **26**, (3.00 g, 20.14 mmol), 4-methoxycarbonylphenyl boronic acid, **27**, (6.12 g, 40.27 mmol) and finely powdered potassium carbonate (8.35 g, 60.41 mmol) were stirred at 60 °C in a solution of toluene:methanol (9:1, 450 ml) for 1 h with thorough degassing with nitrogen. $\text{Pd}(\text{PPh}_3)_4$ (0.10 g) was added to the reaction and it was heated at 120 °C under reflux for 48 h. The solvents were evaporated and the residue partitioned between chloroform (3 x 200 ml) and water (250 ml). The organic layers were combined and dried over MgSO_4 . The solids were removed by filtration through a celite pad. The solvent was evaporated and the residue recrystallised from ethyl acetate to give a crystalline solid **25** (4.68 g, 79%); m.p. 248.2 – 248.6 °C (Lit. 223.5 – 225.0 °C); ^1H NMR (CDCl_3) δ 9.37 (s, 1H, 2-H), 8.21 (m, 8H, 2'-H,

3'-H, 5'-H, 6'-H), 8.14 (s, 1H, 5-H), 3.96 (s, 6H, OMe); ^{13}C NMR (CDCl_3) δ 166.6 (C=O), 164.0 (C-4, C-6), 159.4 (C-2), 140.9 (C-4'), 132.4 (C-1'), 130.4 (C-3', C-5'), 127.3 (C-2', C-6'), 123.7 (C-5), 52.5 (OMe).

2,6-Bis-(4-methoxycarbonylphenyl) pyridine, **67**⁶²



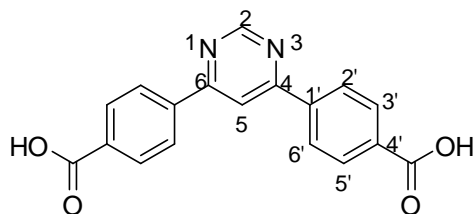
Prepared according to general method 1 from 2,6-dibromopyridine, **65**, (1.00 g, 4.22 mmol) and 4-methoxycarbonylphenyl boronic acid, **27**, (1.52 g, 8.44 mmol). Recrystallisation from ethyl acetate afforded crystalline solid **67** (1.12 g, 76 %); m.p. 212.9 – 213.5 °C (Lit. 194.5 – 195.5°C); ^1H NMR (CDCl_3) δ 8.20 ($\frac{1}{2}$ AB, $J = 8.4$ Hz, 4H, 2'-H, 6'-H), 8.15 ($\frac{1}{2}$ AB, $J = 8.4$ Hz, 4H, 3'-H, 5'-H), 7.86 (t, $J = 8.4$ Hz, 1H, 4-H), 7.77 (d, $J = 8.4$ Hz, 2H, 3-H, 5-H), 3.95 (s, 6H, OMe); ^{13}C NMR (CDCl_3) δ 167.0 (C=O), 160.0 (C-2, C-6), 143.4 (C-4'), 137.9 (C-4), 130.6 (C-1'), 130.1 (C-3', C-5'), 127.0 (C-2', C-6'), 120.0 (C-3, C-5), 52.3 (OMe).

2,7-Bis-(4-methoxycarbonylphenyl) naphthyridine, 68⁶²

Prepared according to general method 1 from 2,7-dichloronaphthyridine, **66**, (1.00 g, 5.02 mmol) and 4-methoxycarbonylphenyl boronic acid, **27**, (1.68 g, 11.1 mmol). Recrystallisation from ethyl acetate afforded a crystalline solid **68** (1.39 g, 70 %); m.p. 267.2–267.8 °C (Lit. 268.0 – 268.5 °C); ¹H NMR (CDCl₃) δ 8.35 (½ AB, *J* = 9.5 Hz, 4H, 2'-H, 6'-H), 8.31 (½ AB, *J* = 9.5 Hz, 2H, 4-H, 5-H), 8.19 (½ AB, *J* = 9.5 Hz, 4H, 4-H), 8.03 (½ AB, *J* = 9.5 Hz, 4H, 3-H, 5-H), 3.95 (s, 6H, OMe); ¹³C NMR (CDCl₃) δ 167.1 (C=O), 160.0 (C-2, C-7), 143.3 (C-4'), 138.0 (C-4, C-5), 131.6 (C-1'), 130.1 (C-3', C-5'), 128.1 (C-2', C-6'), 120.1 (C-3, C-6), 52.3 (OMe). C1a and C4a not observed.

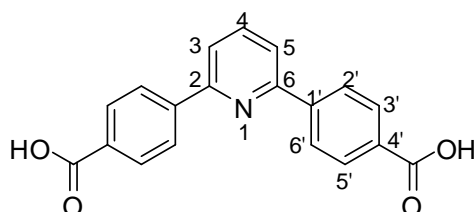
General Method 2 – Hydrolysis of esters

4,6-Bis-(4-carboxyphenyl) pyrimidine, **24**⁶²



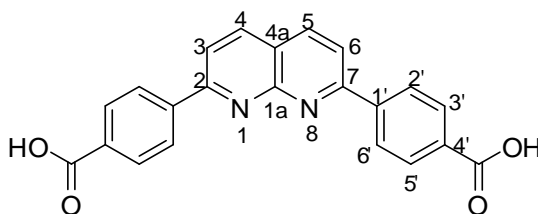
4,6-Bis-(4-methoxycarbonylphenyl) pyrimidine, **25**, (0.50 g, 1.44 mmol) was added to a solution of sodium hydroxide (2 M, 50 ml) which was heated under reflux for 24 h. The reaction mixture was cooled and acidified to pH 5 with HCl (4 M). The precipitate was collected by filtration and washed with water and ether to yield a white powder **24** (0.45 g, 100 %); m.p. decomp. >300 °C (Lit. Decomp. >300 °C); ¹H NMR (DMSO) δ 9.45 (s, 1H, 2-H), 8.74 (s, 1H, 5-H), 8.46 (½ AB, *J* = 8.0 Hz, 4H, 2'-H, 6'-H), 8.09 (½ AB, *J* = 8.2 Hz, 4H, 3'-H, 5'-H); ¹³C NMR (DMSO) δ 167.4 (C=O), 163.5 (C-4, C-6), 159.6 (C-2), 140.5 (C-4'), 133.5 (C-1'), 130.3 (C-3', C-5'), 128.1 (C-2', C-6'), 114.3 (C-5).

2,6-Bis-(4-carboxyphenyl) pyridine, **69**⁶²



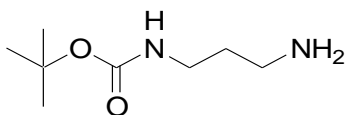
Prepare according to general method 2 from 2,6-bis-(4-methoxycarbonyl) pyridine, **67**, (1.0 g, 2.87 mmol) yielding a white solid **69** (0.91 g, 99 %) m.p. decomp > 300 °C (lit. m.p. 321 – 323 °C); ¹H NMR (DMSO) δ 8.38 (½ AB, *J* = 8.4 Hz, 4H, 2'-H, 6'-H), 7.92 (m, 5H, 3'-H, 5'-H, 4-H), 7.88 (d, *J* = 8.4 Hz, 2H, 3-H, 5-H); ¹³C NMR (DMSO) δ 167.8 (C=O), 155.3 (C-2, C-6), 143.0 (C-4'), 139.4 (C-4), 131.9 (C-1'), 130.5 (C-3', C-5'), 127.4 (C-2', C-6'), 121.0 (C-3, C-5).

2,7-Bis-(4-carboxyphenyl) naphthyridine, **70**⁶²



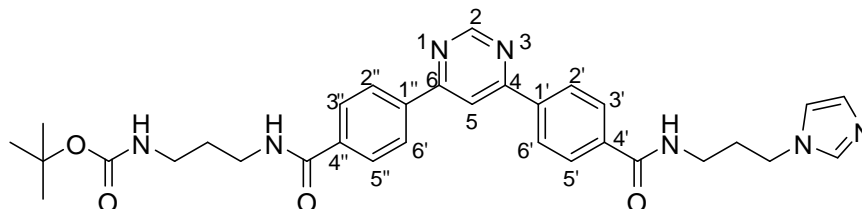
Prepare according to general method 2 from 2,7-bis-(4-methoxycarbonyl) naphthyridine, **68**, (0.5 g, 1.25 mmol) yielding a yellow solid **70** (0.44 g, 95 %) m.p. decomp > 300 °C (lit. m.p. 321-323 °C); ¹H NMR (DMSO) δ 8.67 (½ AB, *J* = 8.6 Hz, 2H, 4-H, 5-H), 8.51 (½ AB, *J* = 8.3 Hz, 4H, 2'-H, 6'-H), 8.39 (½ AB, *J* = 8.6 Hz, 2H, 3-H, 6-H), 8.16 (½ AB, *J* = 8.3 Hz, 4H, 3'-H, 5'-H); ¹³C NMR (DMSO) δ 167.8 (C=O), 159.3 (C-2, C-7), 139.6 (C-4, C-5), 138.7 (C-4'), 131.5 (C-1'), 130.0 (C-3', C-5'), 128.8 (C-2', C-6'), 120.3 (C-3, C-6), C1a and C4a not observed.

EXPERIMENTAL DETAILS FOR SECTION 3.3

***N*-Boc-1,3-diaminopropane, 71¹⁵⁵**

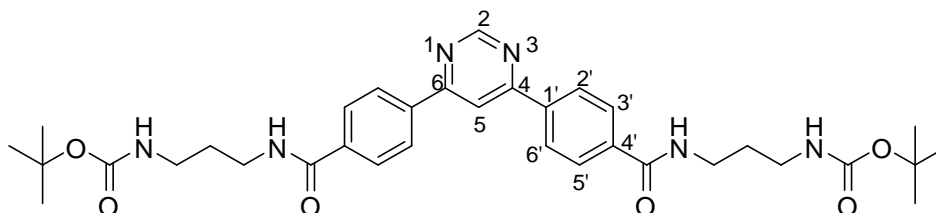
Di-*tert*-butyl dicarbonate, **72**, (10.00 g, 45.8 mmol) in THF (50 ml) was added dropwise to a solution of 1,3-diaminopropane, **5**, (11.20 g, 152.6 mmol) in THF (150 ml) at 0 °C. Once added, the solution was stirred at 0 °C for 30 min, and then left to stir at room temperature for 12 h. The solvents were evaporated and the residue dissolved in ethyl acetate (200 ml). The ethyl acetate layer was washed with brine (20 % w/v, 3 x 100 ml). The organic layer was collected, dried over MgSO₄, filtered through a celite pad and evaporated to leave an oil. The oil was purified by distillation (160-170 °C at 0.1 mmHg) to give a colourless oil **71** (7.81 g, 98 %) Lit. b.p. not given; ¹H NMR (CDCl₃) δ 5.06 (br s, 1H, CONH), 3.14 (q, 2H, *J* = 5.8 Hz, CH₂NHCO), 2.70 (t, 2H, *J* = 5.8 Hz, CH₂NH₂), 1.57 (quint., 2H, *J* = 5.8 Hz, H₂NCH₂CH₂), 1.40 (br s, 2H, NH₂), 1.36 (s, 9H, Bu^t); ¹³C NMR (CDCl₃) δ 156.3 (C=O), 77.4 (CMe₃), 39.3 (CH₂NH₂), 38.3 (CH₂NH), 32.9 (H₂NCH₂CH₂), 28.5 (CCH₃); NMR peaks correspond to literature.

Attempted synthesis of 4-(4-(3-[imidazol-1-yl]propyl)carboxamido phenyl)-6-(4-[3-*N*-Boc-aminopropyl] carboxyamidophenyl) pyrimidine, **22**



4,6-Bis-(4-carboxyphenyl)pyrimidine, **24**, (0.10 g, 0.31 mmol), PyBOP (0.16 g, 0.31 mmol) and Et₃N (0.03 g, 0.31 mmol) was stirred in a solution of dry DCM (10 ml) at RT for 1 h. *N*-Boc-1,3-diaminopropane, **71**, (0.05 g, 0.31 mmol) was added to the mixture which was left to stir at RT for 15 h. PyBOP (0.16 g, 0.31 mmol) and Et₃N (0.03 g, 0.31 mmol) was added and left at RT for 1 h. 1-(3-Propyl)imidazole, **73**, was added and the reaction stirred at RT for 18 h. The resultant precipitate was collected by filtration and washed with ether. The solid was recrystallised from methanol to give a mixture of products **22**, **75** and **76** as a white powder (0.11 g). ¹H NMR (DMSO) δ 9.28 (s, 1H, 2-H), 8.73-8.66 (m, 3H, 5-H, NCHN, NH), 8.54 (t, *J* = 5.5 Hz, 1H, NH) 8.42 (m, 4H, 2'-H, 6'-H, 2''-H, 6''-H), 7.95 (m, 4H, 3'-H, 5'-H, 3''-H, 5''-H), 7.62 and 7.41 (2 x s, 1H, NCHCHN), 6.76 (t, *J* = 5.5 Hz, 1H, NH), 4.18 (t, *J* = 6.5 Hz, 2H, CH₂-imidazole), 3.23-3.19 (m, 4H, CH₂CH₂CH₂-imidazole, CH₂CH₂CH₂NHBoc), 2.90 (q, *J* = 6.5 Hz, 2H, CH₂NHBoc), 1.99 (quint., *J* = 6.5 Hz, 2H, CH₂CH₂-imidazole), 1.55 (quint., *J* = 6.5 Hz, 2H, CH₂CH₂NHBoc), 1.28 (s, 9H, CMe₃); MS (ESI⁺): *m/z* (M+H)⁺ 633.5 (5 %), 584.4 (5 %), 535.4 (5 %).

Attempted synthesis of 4-(4-carboxyphenyl)-6-(4-[3-*N*-Boc-aminopropyl]carboxyamidophenyl) pyrimidine, **74, leading to 4,6-Bis-(4-[3-*N*-Boc-amino propyl]carboxyamidophenyl) pyrimidine, **75****



4,6-Bis-(4-carboxyphenyl)pyrimidine, **24**, (1.00 g, 3.12 mmol), PyBOP (1.62 g, 3.12 mmol) and Et₃N (0.32 g, 3.12 mmol) were stirred at room temperature in dry DCM (50 ml). *N*-Boc-1,3-diaminopropane, **71**, (0.54 g, 3.12 mmol) was added to the mixture which was left to stir at room temperature for 15 h. The precipitate was collected by filtration and washed with ether. The solid was triturated with hot methanol to remove excess starting material. The filtrate was concentrated and purified by flash chromatography on silica gel eluting with chloroform: methanol (98:2 to 50:50) to yield white solid **75** (0.32 g, 16 %); m.p. 220.9 – 221.7 °C; ¹H NMR (DMSO) δ 9.36 (s, 1H, 2-H), 8.76 (s, 1H, 5-H), 8.63 (t, *J* = 5.7 Hz, 2H, NHCOPh), 8.47 (½ AB, *J* = 8.4 Hz, 4H, 2'-H, 6'-H), 8.02 (½ AB, *J* = 8.4 Hz, 4H, 3'-H, 5'-H), 6.85 (t, 1H, *J* = 5.7 Hz, NHBoc), 3.29 (q, *J* = 6.5 Hz, 4H, CH₂NHCOPh), 2.98 (q, *J* = 6.5 Hz, 4H, CH₂NHBoc), 1.65 (m, *J* = 6.5 Hz, 4H, NHCH₂CH₂), 1.37 (s, 9H, CCH₃); ¹³C NMR (DMSO) δ 166.4 (C=OPh), 163.4 (C-4, C-6), 159.3 (C-2), 156.2 (COOBu^t), 139.1 (C-4'), 137.3 (C-1'), 128.3 (C-3', C-5'), 127.8 (C-2', C-6'), 113.8 (C-5), 78.0 (CMe₃), 38.2 and 37.3 (NHCH₂CH₂CH₂NH), 30.1 (CH₂CH₂NH), 28.8 (CH₃); IR (KBr): 3350s

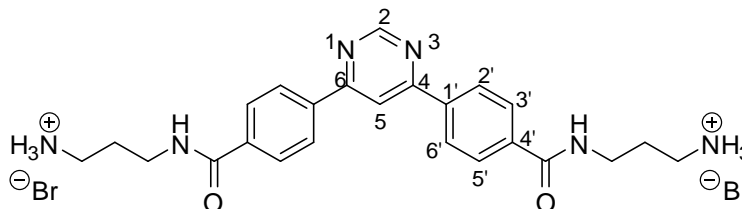
Experimental

(N-H), 2950m (Ar C-H), 2350m (C-H), 1700s (C=O), 1600s (Ar C-C), 1550s (N-H), 1550s (C-H₂), 1450m (C=C), 1375s (C-N), 1150m (C-O); MS (ESI⁺): m/z 633 (98 %) (M+H)⁺, 613 (4 %) (M-H₃O)⁺, 595 (13 %), 577 (10 %) (M-Bu^t)⁺, 566 (4 %), 533 (8 %) (M-Boc)⁺, 518 (4 %) (M-NBoc)⁺, 486 (8 %), 477 (5 %) (M-CH₂CH₂CH₂NHBoc)⁺, 381 (4 %), 347 (2 %), 339 (18 %), 309 (5 %), 300 (27 %), 278 (95 %), 250 (32 %), 239 (12 %) 217 (100 %) (M-2(Boc)-NH)⁺, 209 (44 %), 200 (40 %) (NH(CH₂)₃NHBoc)⁺, 181 (24 %), 163 (13 %), 149 (34 %), 131 (18 %), 117 (36 %) (BocNH₂)⁺.

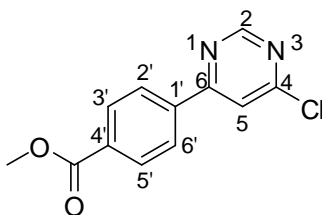
Elemental analysis found C, 64.19; H, 6.89; N, 13.32 %.

C₃₄H₄₄N₆O₆ requires: C, 64.54; H, 7.01; N, 13.32 %.

4,6-Bis-(4-[3-aminopropyl]carboxyamidophenyl)-6-(4-hydroxycarbonyl phenyl) pyrimidine hydrobromide salt



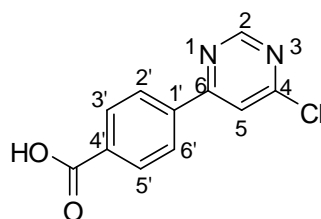
4-(4-[3-Boc-aminopropyl]carboxyamidophenyl)-6-(4-hydroxycarbonylphenyl) pyrimidine, **75**, (0.10 g, 0.16 mmol) was taken up in conc. HBr and evaporated to dryness to yield the hydrobromide salt (0.07 g, 85 %); m.p. decomp. > 250°C; ^1H NMR (DMSO) δ 9.36 (s, 1H, 2-H), 8.84 (br s, 2H, NH), 8.78 (s, 1H, 5-H), 8.49 ($\frac{1}{2}$ AB, $J = 7.2$ Hz, 4H, 2'-H, 6'-H), 8.06 ($\frac{1}{2}$ AB, $J = 7.2$ Hz, 4H, 3'-H, 5'-H), 7.75 (br s, 6H, $^+\text{NH}_3$), 3.36 (br, 4H, CH_2NHCOPh), 2.87 (br, 4H, CH_2^+NH_3), 1.88 (br, 4H NHCH_2CH_2); ^{13}C NMR (DMSO) δ 166.5 (C=O), 163.6 (C-4, C-6), 159.6 (C-2), 139.2 (C-1') 136.9 (C-4'), 128.4 (C-3', C-5'), 127.9 (C-2', C-6'), 113.9 (C-5), 37.8 and 37.3 (CH_2NH , CH_2^+NH_3), 28.0 ($\text{CH}_2\text{CH}_2\text{NH}$); IR (KBr): 3400s (N-H), 3000 br (N-H), 2900m (Ar C-H), 2350m (C-H), 1550s (Ar C-C), 1525s (C=O), 1550s (N-H), 1475s (Ar C-C), 1450m (C-H₂), 1425m (C=C), 1250s (C-N), 1150m (C-O); MS (ESI+): m/z 433 (22%) ($\text{M}+\text{H}$) $^+$, 416 ($\text{M}-\text{NH}_3$) $^+$, 399 (4 %) ($\text{M}-2\text{NH}_3$) $^{2+}$, 359 (45 %) ($\text{M}-\text{H}_2\text{NCH}_2\text{CH}_2\text{CH}_2\text{NH}$) $^{2+}$, 342 (3 %) ($\text{M}-[\text{H}_2\text{NCH}_2\text{CH}_2\text{CH}_2\text{NH}]-\text{NH}_3$) $^{2+}$, 274 (7 %), 256 (5 %), 229 (3 %) ($\text{M}-2[\text{H}_2\text{NCH}_2\text{CH}_2\text{CH}_2\text{NHCO}]$) $^{2+}$, 203 (5 %), 175 (5 %), 135 (4 %), 122 (6 %), 111 (14 %), 98 (15 %), 87 (18 %), 75 (25 %) ($\text{H}_2\text{NCH}_2\text{CH}_2\text{CH}_2\text{NH}$) $^{2+}$, 56 (100 %). Elemental analysis found C, 64.19; H, 6.89; N, 13.32 % $\text{C}_{34}\text{H}_{44}\text{N}_6\text{O}_2 \cdot 3.75\text{HBr} \cdot 0.2\text{H}_2\text{O}$ requires: C, 38.98; H, 4.38; N, 11.36 %.

4-chloro-6-(4-methoxycarbonylphenyl) pyrimidine, 32

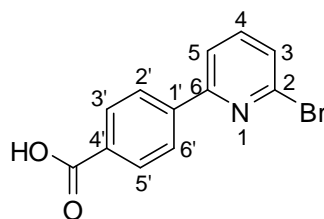
4,6-dichloropyrimidine, **26**, (0.50 g, 3.36 mmol), 4-methoxyphenylboronic acid, **27**, (0.60 g, 3.36 mmol) and aqueous (1 ml) sodium carbonate (1.10 g, 10.07 mmol) were stirred in a solution of glyme (50 ml) at room temperature with thorough degassing with N₂ for 1 h. Pd(OAc)₂ (0.04 g, 2.50 mmol%) and PPh₃ (0.09 g, 5.00 mmol%) were added to the reaction mixture and it was heated at 100 °C under N₂ reflux for 18 h. The solvents were removed by evaporation and the residue separated between water (200 ml) and chloroform (150 ml, 2 x 50 ml). The organic layers were combined and dried over MgSO₄. The solids were removed by filtration through a celite pad and the solvents evaporated. Recrystallisation from IPA afforded a pink solid **32** (0.68 g, 79 %); m.p. 165.0 – 166.3 °C; ¹H NMR (DMSO) δ 9.22 (s, 1H, 2-H), 8.50 (s, 1H, 5-H), 8.46 (½ AB, *J* = 8.4 Hz, 2H, 2'-H, 6'-H), 8.18 (½ AB, *J* = 8.4 Hz, 2H, 3'-H, 5'-H), 3.96 (s, 3H, OMe); ¹³C NMR (DMSO) δ 166.2 (C=O), 164.3 (C-6), 162.1 (C-4), 159.6 (C-2), 139.5 (C-4'), 132.7 (C-1'), 130.3 (C-3', C-5'), 128.4 (C-2', C-6'), 118.8 (C-5), 53.0 (CH₃); IR (KBr): 2350s, 1735s (C=O), 1575s (Ar C-C), 1460s (Ar C-C), 1450 (C-H₃), 1120 (C-O), 800s (C-Cl) cm⁻¹. MS (ESI⁺): *m/z* 251.1 (25 %) (M+H)⁺(³⁷Cl), 249.1 (100 %) (M+H)⁺(³⁵Cl), 239.2 (5 %), 205.1 (5 %);

Elemental analysis found C, 55.69; H, 3.72; N, 10.54 %

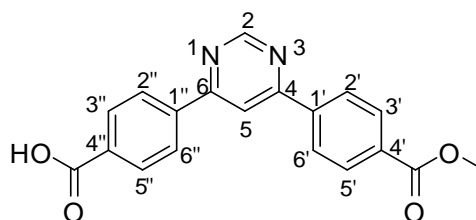
C₁₂H₂₄N₂O₂Cl•0.5H₂O requires: C, 55.94; H, 3.91; N, 10.87 %.

4-Chloro-6-(4-carboxyphenyl) pyrimidine, **31**¹⁵⁹

4-Carboxyphenylboronic acid, **80**, (2.78 g, 16.75 mmol) was stirred in a solution sodium carbonate (0.4 M, 150 ml) with thorough degassing with Ar for 30 min. This solution was added dropwise to a solution of acetonitrile (150 ml) containing 4,6-dichloropyrimidine (5.00 g, 33.56 mmol) and Pd(PPh₃)₄ (0.01 g) also thoroughly degassed with Ar. The combined reaction mixture was heated under Ar at reflux at 90 °C for 18 h. The mixture was cooled and washed with DCM (200 ml). The solution was acidified with conc. HCl and the precipitate was collected by filtration, washing with water and ether. The solid was loaded onto a plug of silica and washed through with 10% methanol in chloroform solution. The solvents were evaporated to yield white solid **31** (2.50 g, 64 %); m.p. 248.0 – 248.9 °C (lit m.p. not given); ¹H NMR (DMSO) δ 9.14 (s, 1H, 2-H), 8.42 (s, 1H, 5-H), 8.36 (½ AB, *J* = 8.3 Hz, 2H, 2'-H, 6'-H), 8.09 (½ AB, *J* = 8.3 Hz, 2H, 3'-H, 5'-H); ¹³C NMR (DMSO) δ 167.3 (C=O), 164.4 (C-6), 162.1 (C-4), 159.6 (C-2), 140.6 (C-4'), 139.2 (C-1'), 130.4 (C-2', C-6'), 128.2 (C-3', C-5'), 118.6 (C-5). NMR peaks correspond to literature.

2-Bromo-6-(4-carboxyphenyl) pyridine, **82**¹⁷¹

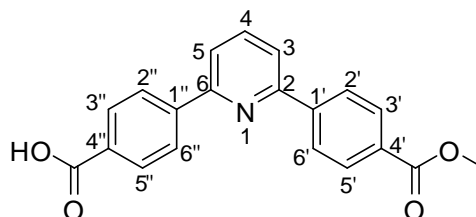
4-Carboxyphenylboronic acid, **80**, (1.00 g, 6.03 mmol) was stirred in a solution sodium carbonate (0.4 M, 50 ml) with thorough degassing with Ar for 30 min. This solution was added dropwise to a solution of acetonitrile (75 ml) containing 2,6-dibromopyridine, **65**, (2.86 g, 12.05 mmol) and Pd(PPh₃)₄ (0.01 g) also thoroughly degassed with Ar. The combined reaction mixture was heated under reflux at 90 °C for 18 h, then cooled and washed with DCM (150 ml). The solution was acidified with conc. HCl and the precipitate was collected by filtration, washing with water and ether. The solid was loaded onto a plug of silica and washed through with a 10 % methanol in chloroform solution. The solvents were evaporated to leave a white solid **82** (1.56 g, 93 %) m.p. 278.3 – 278.9 °C (Lit m.p. not given); ¹H NMR (DMSO) δ 13.11 (br s, 1H, OH), 8.18 (½ AB, *J* = 8.6 Hz, 2H, 2'-H, 6'-H), 8.12 (d, *J* = 7.6 Hz, 1H, 3-H), 8.06 (½ AB, *J* = 8.6 Hz, 2H, 3'-H, 5'-H), 7.89 (t, *J* = 7.6 Hz, 1H, 4-H), 7.67 (d, *J* = 7.6 Hz, 1H, 5-H); ¹³C NMR (DMSO) δ 167.4 (C=O), 156.7 (C-6), 141.9 (C-4'), 141.2 (C-4), 141.1 (C-2), 132.2 (C-1'), 130.3 (C-3', C-5'), 128.0 (C-3), 127.2 (C-2', C-6'), 120.8 (C-5). NMR peaks correspond to literature.

4-(4-methoxycarbonylphenyl)-6-(4-carboxyphenyl) pyrimidine, 30

4-Chloro-6-(4-methoxycarbonylphenyl)pyrimidine, **31**, (0.15 g, 0.64 mmol), 4-methoxycarbonyl phenyl boronic acid, **27**, (0.19 g, 1.28 mmol) and potassium carbonate (0.27 g, 1.92 mmol) were stirred with thorough degassing with Ar in a solution of toluene:methanol (8:2, 50 ml). After 2 h Pd(PPh₃)₄ (0.01 g) was added and the reaction mixture was heated under Ar at reflux at 120 °C for 48 h. The solvents were evaporated and the residue partitioned between chloroform (50 ml) and water (50 ml). The water layer was collected and acidified with conc. HCl. The precipitate was collected by filtration to give a white solid **30** (0.28 g, 55 %); m.p. decomp > 250 °C; ¹H NMR (DMSO) δ 9.38 (s, 1H, 2-H), 8.78 (s, 1H, 5-H), 8.48-8.52 (m, 4H, 2'-H, 6'-H, 2''-H, 6''-H), 8.15 (m, 4H, 3'-H, 5'-H, 3''-H, 5''-H), 3.90 (s, 3H, OMe); ¹³C NMR (DMSO) δ 167.4 (COOH), 166.4 (COOMe), 163.6 and 163.4 (C-4, C-6), 159.6 (C-2), 140.9 and 140.5 (C-4', C-4''), 133.5 and 132.3 (C-1', C-1''), 130.4 and 130.2 (C-3', C-5', C-3'', C-5''), 128.2 and 128.1 (C-2', C-6', C-2'', C-6''), 114.4 (C-5), 52.9 (OMe); IR (KBR): 3000m (Ar C-H), 2750s (O-H), 1725 (C=O(OMe)), 1700s (C=O(OH)), 1600 (Ar C-C), 1475s (Ar C-C), 1450s (C-H₃), 1375m (C-H₃), 1115s (C-O) cm⁻¹. MS (ESI-): m/z 333.0 (100%) (M-H)⁻, 289.2 (10%) (M-CO₂Me)⁻, 212.2 (5%) (M-PhCO₂Me)⁻, 145.1 (10%), 81.3 (5%), 59.2 (57%), (MeO₂)⁻. Elemental analysis found C, 56.88; H, 3.57; N, 6.88.

$C_{19}H_{14}N_2O_4 \cdot \frac{1}{2}HCl \cdot \frac{1}{2}CHCl_3$ requires: C, 56.81; H, 3.67; N, 6.80 %.

2-(4-Methoxycarbonylphenyl)-6-(4-carboxyphenyl) pyrimidine, **83**



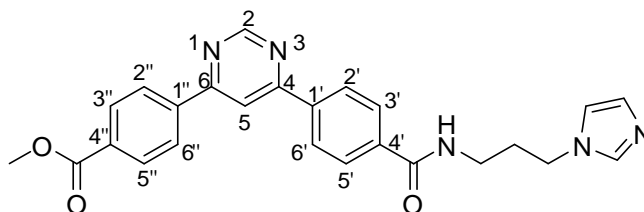
2-Bromo-6-(4-methoxycarbonylphenyl)pyridine, **82**, (1.00 g, 3.60 mmol), 4-methoxycarbonyl phenyl boronic acid, **27**, (0.60 g, 3.96 mmol) and potassium carbonate (0.99 g, 7.19 mmol) were stirred with thorough degassing with Ar in a solution of toluene:methanol (8:2, 150 ml). After 1 h $Pd(PPh_3)_4$ (0.01 g) was added and the reaction mixture was heated under reflux at 120 °C for 72 h. The solvents were evaporated and the residue partitioned between chloroform (50 ml) and water (50 ml). The water layer was collected and acidified with conc. HCl. The precipitate was collected by filtration to give a white solid **83** (1.12 g, 74 %); m.p. decomp >300 °C; 1H NMR (DMSO) δ 8.38 ($\frac{1}{2}$ AB, $J = 8.4$ Hz, 2H, 2''-H, 6''-H), 8.12 ($\frac{1}{2}$ AB, $J = 8.4$ Hz, 2H, 2'-H, 6'-H), 8.10 ($\frac{1}{2}$ AB, $J = 8.4$ Hz, 2H, 3'-H, 5'-H), 8.0 (m, 3H, 3-H, 4-H, 5-H), 7.96 ($\frac{1}{2}$ AB, $J = 8.4$ Hz, 2H, 3''-H, 5''-H), 3.91 (s, 3H, OMe); ^{13}C NMR (DMSO) δ 168.4 (C=O), 168.0 (C=O), 155.6 and 155.4 (C-2, C-6), 143.4 and 143.3 (C-1', C-1''), 139.0 (C-4), 130.8 and 130.6 (C-4', C-4''), 130.3 and 129.9 (C-3', C-5', C-3'', C-5''), 127.5 and 126.0 (C-2', C-6', C-2'', C-6''), 120.3 and 120.0 (C-3, C-5), 52.8 (OMe); IR (ATR): 3295 (O-H), 3000s (Ar C-H), 1712s (C=O(OMe)), 1699s (C=O(OH)), 1607 (Ar

C-C), 1475s (Ar C-C), 1441m (C-H₃), 1375m (C-H₃), 1016 (C-O) cm⁻¹. MS (ESI⁺): m/z 334.1 (100 %) (M+H)⁺, 279.0 (2 %), 269.0 (5 %), 232.0 (15 %) (M-CO₂-CO₂Me)⁺, 214.2 (10 %), 187.0 (10 %), 121.1 (5 %), 105.2 (15 %) (PhCO₂)⁺, 102.4 (3 %), 55.5 (5 %).

Elemental analysis found C, 56.47; H, 4.44; N, 3.31 %

C₂₀H₁₅NO₄·2HCl·1H₂O requires: C, 56.62; H, 4.51; N, 3.30 %.

4-(4-[3-Imidazol-1-ylpropylcarboxamido]phenyl)-6-(4-methoxycarbonylphenyl)pyrimidine, 28



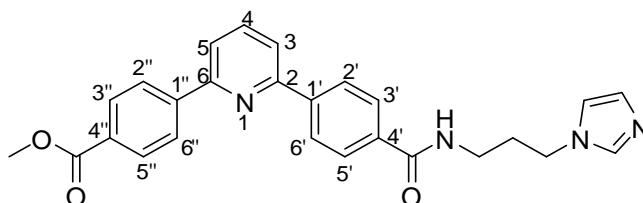
4-(4-Methoxycarbonylphenyl)-6-(4-carboxyphenyl)pyrimidine, **30**, (0.5 g, 1.50 mmol), PyBOP (0.87 g, 1.65 mmol) and Et₃N (0.17 g, 1.65 mmol) were stirred in DCM (50 ml) at RT for 1 h. 1-(3-Aminopropyl)imidazole, **73**, (0.21 g, 1.65 mmol) was added to the reaction mixture and it was left at RT for 18 h. The precipitate was collected by filtration to leave a white solid **28** (0.45 g, 48 %); m.p. 182.4 – 183.5 °C; ¹H NMR (DMSO) δ 9.38 (s, 1H, 2-H), 8.78 (s, 1H, 5-H), 8.71 (t, J = 5.7 Hz, 1H, NH), 8.52 (½ AB, J = 8.4 Hz, 2H, 2'-H, 6'-H), 8.48 (½ AB, J = 8.4 Hz, 2H, 2''-H, 6''-H), 8.14 (½ AB, J = 8.4 Hz, 2H, 3'-H, 5'-H), 8.04 (½ AB, J = 8.4 Hz, 2H, 3''-H, 5''-H), 7.99 (s, 1H NCHN), 7.36 and 7.07 (2 x s, 1H, NCHCHN), 4.09 (t, J = 6.6 Hz, 2H, CONHCH₂CH₂CH₂), 3.89 (s, 3H, OMe), 3.28 (q, J = 6.6 Hz, 2H, CONHCH₂CH₂CH₂), 2.01 (quint., J = 6.6 Hz, 2H, CONHCH₂CH₂CH₂); ¹³C NMR (DMSO) δ 166.4 (C=O), 166.3 (C=O), 163.7 and 163.3 (C-4, C-6),

159.6 (C-2), 141.0 (C-4''), 139.0 (C-4'), 137.5 and 137.2 (C-1', C-1''), 132.2 (NCHN), 130.2 (C-3'', C-5''), 128.4 (C-3', C-5'), 128.3 (C-2', C-6'), 127.8 (C-2'', C-6''), 127.1 and 120.5 (NCH₂CH₂N), 114.2 (C-5), 52.6 (OMe), 44.9 (CONHCH₂CH₂CH₂), 37.1 (CONHCH₂), 31.0 (CONHCH₂CH₂); IR (KBr): 3100m (Ar C-H), 1725s (C=O(OH)), 1650s (C=O(NH)), 1625s (N-H), 1500m (C-H₂), 1450m (C-H₃), 1300s (C-N), 1100m (C-O) cm⁻¹. MS (ESI+): m/z 442.1 (100 %) (M+H)⁺, 374.1 (25 %) (M-C₃H₄N₂)⁺, 257.0 (2 %), 86.0 (2 %), 69.1 (2 %) (C₃H₅N₂)⁺.

Elemental analysis found C, 61.72; H, 4.56; N, 14.29 %.

C₂₅H₂₃N₅O₃·²/₃CH₂Cl₂ requires: C, 61.86; H, 4.92; N, 14.05 %.

2-(4-[3-Imidazol-1-ylpropylcarboxamido]phenyl)-6-(4-methoxycarbonylphenyl) pyridine, **84**



2-(4-Methoxycarbonylphenyl)-6-(4-carboxyphenyl)pyridine, **83**, (0.5 g, 1.38 mmol), PyBOP (0.79 g, 1.52 mmol) and Et₃N (0.28 g, 2.77 mmol) were stirred in a solution of DCM (50 ml) at RT for 1 h. 1-(3-Aminopropyl)imidazole, **73**, (0.19 g, 1.52 mmol) was added and the mixture stirred for 18 h at RT. The precipitate was collected by filtration to leave a white solid **84** (0.61 g, 100 %); m.p. 183.2 – 183.6 °C; ¹H NMR (DMSO) δ 8.54 (t, *J* = 5.5 Hz, 1H, NH); 8.26 (½ AB, *J* = 8.4 Hz, 2H, 2'-H, 6'-H), 8.20 (½ AB, *J* = 8.4 Hz, 2H, 2''-H, 6''-H), 8.00 (½ AB, *J* = 8.4 Hz, 2H, 3''-H, 5''-H), 7.97-7.94 (m, 3H, 3-H, 4-H, 5-H), 7.89 (½ AB,

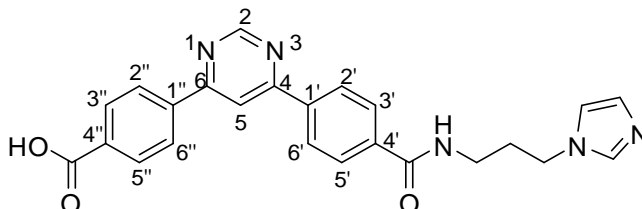
Experimental

$J = 8.4$ Hz, 2H, 3'-H, 5'-H), 7.56 (s, 1H NCHN), 7.11 and 6.78 (2 x s, 1H, NCHCHN), 3.93 (t, $J = 6.7$ Hz, 2H, CONHCH₂CH₂CH₂), 3.78 (s, 3H, OMe), 3.16 (q, $J = 6.7$ Hz, 2H, CONHCH₂), 1.87 (quint., $J = 6.7$ Hz, 2H, CONHCH₂CH₂); ¹³C NMR (DMSO) δ 166.6 (C=O), 166.3 (C=O), 155.6 and 151.0 (C-2, C-6), 143.5, (C-4''), 141.4 (C-4'), 139.3 (C-4), 137.9 (NCHN), 130.6 and 130.5 (C-1', C-1''), 130.3 (C-3'', C-5''), 129.0 (C-3', C-5') 128.4 (C-2', C-6'), 127.5 (C-2'', C-6''), 127.1 (NCHCHN or NCHCHN), 120.9 and 120.7 (C-3, C-5), 119.9 (NCHCHN or NCHCHN), 52.9 (OMe), 44.4 (CONHCH₂CH₂CH₂), 37.2 (CONHCH₂), 31.3 (CONHCH₂CH₂); IR (ATR): 3054m (Ar C-H), 1713s (C=O(OMe)), 1645s (C=O(NH)), 1609s (Ar C-C), 1558 (N-H), 1509m (C-H₂), 1438m (C-H₃), 1279s (C-N) cm⁻¹; MS (ESI+): m/z 441.1 (100 %) (M+H)⁺, 373.1 (20 %) (M-imidazole)⁺, 152.2 (5 %) (CO(CH₂)₃(CH)₃N₂)⁺, 150.2 (15 %), 86.4 (42 %), 58.5 (COOMe)⁺.

Elemental analysis found C, 70.62; H, 5.19; N, 12.95 %

C₂₆H₂₄N₄O₃ requires: C, 70.89; H, 5.49; N, 12.72 %.

4-(4-[3-Imidazol-1-ylpropylcarboxyamido]phenyl)-6-(4-methoxycarbonylphenyl)pyrimidine, 23



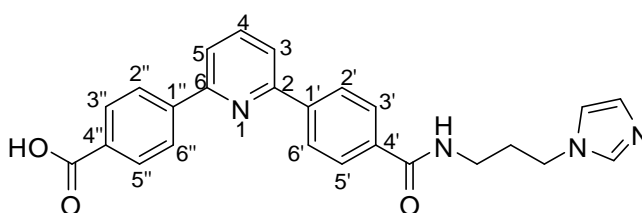
4-(4-[3-Imidazol-1-ylpropylcarboxyamido]phenyl)-6-(4-methoxycarbonylphenyl)pyrimidine, **28**, (0.2 g, 0.45 mmol) and NaOH (0.18 g, 4.5 mmol) were heated in water (50 ml) to 50 °C until the solid had dissolved. The solution was acidified with HCl to pH 3 (4 M) and the resultant precipitate filtered and washed with water and ether to leave a waxy solid **23** (0.18 g, 73 %); m.p. decomp. >300 °C; ¹H NMR (DMSO) δ 9.38 (d, *J* = 1.2 Hz, 1H, 2-H), 9.26 (s, 1H, NCHN), 8.95 (t, *J* = 5.5 Hz, 1H, NH), 8.79 (d, *J* = 1.2 Hz, 1H, 5-H), 8.50 and 8.48 (2 x ½ AB, *J* = 8.4 Hz, 2H, 2'-H, 6'-H and 2''-H, 6''-H), 8.11 and 8.09 (2 x ½ AB, *J* = 8.4 Hz, 2H, 3'-H, 5'-H and 3''-H, 5''-H), 7.89 and 7.68 (2 x s, 1H, NCHCHN), 4.29 (t, *J* = 6.6 Hz, 2H CONHCH₂CH₂CH₂), 3.31 (q, *J* = 6.6 Hz, 2H CONHCH₂), 2.11 (quint., *J* = 6.6 Hz, 2H CONHCH₂CH₂); ¹³C NMR (DMSO) δ 166.7 (C=O), 166.6 (C=O), 162.4 and 162.1 (C-4, C-6), 155.6 (C-2), 143.4 and 141.9 (C-4' or C-4''), 139.3 (C-1' or C-1''), 137.9 (NCHN), 130.3 (C-3'', C-5''), 129.0 (C-3', C-5'), 128.4 and 127.5 (C-2', C-6', C-2'', C-6''), 127.1 and 120.9 (NCH₂CH₂N or NCH₂CH₂N), 119.9 (C-5), 44.4 (CONHCH₂CH₂CH₂), 37.2 (CONHCH₂), 31.4 (CONHCH₂CH₂); IR (ATR): 3230 (O-H), 3056m (Ar C-H), 1697s (C=O(OH)), 1650s (C=O(NH)), 1465m (C-H₂), 1226s (C-N) cm⁻¹; MS (ESI+):

m/z 428.1 (100 %) (M+H)⁺, 214.2 (6 %) (M-PhCONH(CH₂)₃(CH)₃N₂)⁺, 179.1 (2 %), 158.1 (3 %), 151.1 (5 %), 125.2 (3 %), 69.5 (7 %);

Elemental analysis found C, 52.44; H, 3.96; N, 12.74 %

C₂₄H₂₁N₅O₃•3HCl•²/₃CO₂ requires: C, 52.34; H, 4.27; N, 12.38 %.

2-(4-[3-Imidazol-1-ylpropylcarboxyamido]phenyl)-6-(4-carboxyphenyl) pyridine, 85



2-(4-[3-Imidazol-1-ylpropylcarboxyamido]phenyl)-6-(4-methoxycarbonylphenyl) pyridine, **84**, (0.5 g, 1.14 mmol) and NaOH (0.45 g, 11.35 mmol) were heated in water (50 ml) to 50 °C for 2 h. The solution was acidified with HCl (4 M) to pH 3 and the resultant precipitate collected by filtration and washed with water and ether to leave a white solid (0.40 g, 63 %); m.p. decomp >300 °C; ¹H NMR (DMSO) δ 9.23 (s, 1H NCH_N), 8.80 (t, *J* = 5.3 Hz, 1H, NH); 8.36 (½ AB, *J* = 9.1 Hz, 2H, 2''-H, 6''-H), 8.20 (½ AB, *J* = 9.1 Hz, 2H, 2'-H, 6'-H), 8.11-8.03 (m, 7H, 3'-H, 5'-H, 3''-H, 5''-H, 3-H, 4-H, 5-H), 7.86 and 7.70 (2 x s, 1H, NCH_{CHN}), 7.70 (s, 1H, NCH_{CHN} or NCH_{CHN}), 4.32 (t, *J* = 6.7 Hz, 2H, CONHCH₂CH₂CH₂), 3.32 (q, *J* = 6.7 Hz, 2H, CONHCH₂), 2.14 (quint., *J* = 6.7 Hz, 2H, CONHCH₂CH₂); ¹³C NMR (DMSO) δ 167.7 (COOH), 166.7 (CONH), 155.5 and 155.3 (C-2, C-6), 143.0 (C-4''), 141.5 (C-4'), 139.3 (C-4), 136.1 (NCH_N), 135.3 (C-1'), 131.8 (C-1''), 130.4 (C-3'', C-5''), 128.4 (C-3', C-5'), 127.4 (C-2'', C-6''), 127.1 (C-2', C-6'), 122.6 (NCH_{CHN} or NCH_{CHN}),

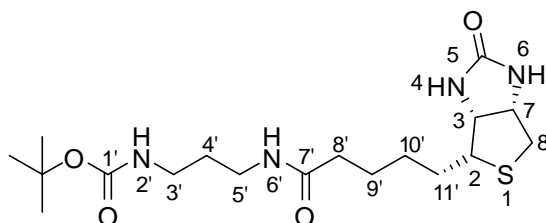
120.71 and 120.70 (C-3, C-5), 120.4 (NCHCHN or NCHCHN), 47.1 (CONHCH₂CH₂CH₂), 36.6 (CONHCH₂), 30.3 (CONHCH₂CH₂); IR (ATR): 3000m (Ar C-H), 1679s (C=O(OH)), 1634s (C=O(NH)), 1610s (Ar C-C), 1587 (N-H), 1449m (C-H₃), 1291 (C-N) cm⁻¹; MS (ESI+): m/z 427.1 (100%) (M+H)⁺, 381.2 (10%) (M-CO₂H)⁺, 353.2 (7%), 214.1 (5%).

Elemental analysis found C, 54.69; H, 3.93; N, 9.82 %

C₂₅H₂₂N₄O₃•2½HCl•1½H₂O requires: C, 54.54; H, 4.23; N, 9.6 %.

6.2.5 EXPERIMENTAL DETAILS FOR SECTION 3.4

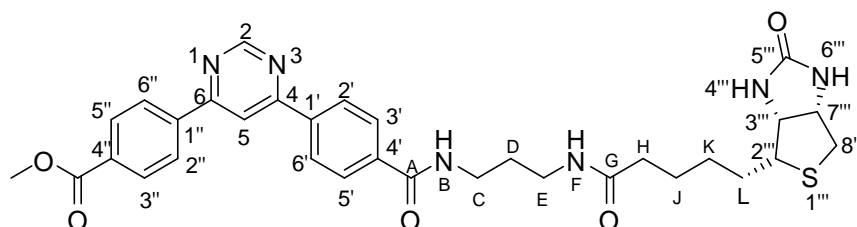
N-Boc-1-Amino-3-[biotin-1-ylamido]carboxyamidopropane, **88**



Biotin, **87**, (2.00 g 8.19 mmol), PyBOP (4.69 g, 9.00 mmol) and Et₃N (0.91 g, 9.00 mmol) were stirred at room temperature in DMF (10 ml) for 1 h. *N*-Boc-1,3-Diaminopropane, **71**, (1.58 g, 9.00 mmol) was added to the solution and left to stir for 15 h. The resultant oil was washed with ethyl acetate and the residue filtered to leave a white solid **88** (2.66 g, 92 %); m.p. 120.6-121.3 °C; ¹H NMR (DMSO, 400.13 MHz) δ 7.73 (t, *J* = 5.3 Hz, 1H, 6'-H), 6.74 (t, *J* = 5.3 Hz, 1H, 2'-H), 6.41 (s, 1H, 6-H), 6.35 (s, 1H, 4-H), 4.31 (s, 1H, 7-H), 4.14 (s, 1H, 3-H), 3.10 (m, 1H, 2-H), 3.01 (q, *J* = 6.1 Hz, 2H, 5'-H), 2.90 (q, *J* = 6.3 Hz, 2H, 3'-H), 2.82

(dd, 1H, $J = 5.1, 12.4$ Hz, 8-H_a), 2.58 (d, 1H, $J = 12.4$ Hz, 8-H_b), 2.05 (t, $J = 7.3$ Hz, 2H, 8'-H), 1.60 (m, 1H, 11'-H_a), 1.52-1.45 (m, 5H, 4'-H, 9'-H, 11'-H_b), 1.38 (s, 9H, ^tBu), 1.31 (m, 2H, 10'-H). ¹³C NMR (DMSO) δ 172.5 (C-7'), 163.2 (C-5), 156.1 (C-1'), 78.0 (C-CH₃), 61.5 (C-3), 59.7 (C-7), 55.9 (C-2), 38.1 (C-3'), 36.7 (C-5'), 35.8 (C-8'), 30.2 (C-4'), 28.8 (C-CH₃), 28.7 (C-10'), 28.6 (C-11'), 25.8 (C-9') C-8''' under DMSO solvent peak; IR (ATR): 3261m (C-H), 1765s (C=O), 1629 (C=O), 1525 (N-H), 1473m (C-H₂), 1450m (C-H₃) cm⁻¹; MS (ESI+): m/z 401.2 (100 %) (M+H)⁺, 345.1 (5%), 301.1 (20 %) (M-Boc)⁺, 214.2 (10%); Elemental analysis found C, 51.27; H, 7.88; N, 13.77; C₁₈H₃₂N₄O₄S·H₂O requires: C, 51.65; H, 8.19; N, 13.39 %.

4-[4-(3-[Biotin-1-ylamido]carboxyamidopropyl)carboxyamidophenyl]-4-[4-methoxycarbonylphenyl]pyrimidine, 90



4-(4-Methoxycarbonylphenyl)-6-(4-carboxyphenyl)pyrimidine, **30**, (0.25 g, 0.75 mmol), PyBOP (0.43 g, 0.82 mmol) and Et₃N (0.38 g, 3.74 mmol) were stirred in DMF (10 ml) at RT for 1 h. In parallel, *N*-Boc-1-amino-3-[biotin-1-ylamido] carboxyamidopropane, **88**, (0.29 g, 0.82 mmol) was taken up in HBr (1 ml) and the HBr was evaporated under reduced pressure. The residue was added to the PyBOP reaction mixture and it was left to stir for 18 h at RT. The DMF was evaporated under reduced pressure and the residue triturated with ethyl acetate

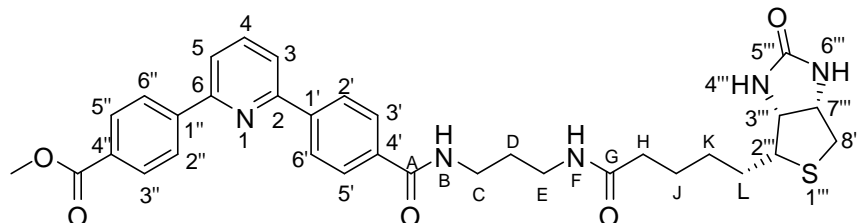
Experimental

to leave a white solid **90** (0.16 g, 46 %); m.p. 154.2-154.6 °C; ¹H NMR (DMSO) δ 9.26 (d, *J* = 1.2 Hz, 1H, 2-H), 8.68 (d, *J* = 1.2 Hz, 1H, 5-H), 8.55 (t, *J* = 5.5 Hz, 1H, NH_B), 8.43 (½ AB, *J* = 8.6 Hz, 2H, 2'-H, 6'-H), 8.37 (½ AB, *J* = 8.6 Hz, 2H, 2''-H, 6''-H), 8.03 (½ AB, *J* = 8.6 Hz, 2H, 3''-H, 5''-H), 7.92 (½ AB, *J* = 8.6 Hz, 2H, 3'-H, 5'-H), 7.75 (t, *J* = 5.7 Hz, 1H, NH_F), 6.32 (s, 1H, 6'''-H), 6.24 (s, 1H, 4'''-H), 4.17 (m, 1H, 7'''-H), 4.00 (m, 1H, 3'''-H), 3.79 (s, 3H, OMe), 3.22 (m, 1H, 2'''-H), 3.18 (q, *J* = 6.2 Hz, 2H, H_C), 3.00 (q, *J* = 6.2 Hz, 2H, H_E), 2.68 (dd, *J* = 5.0 Hz, 12.5 Hz, 1H, 8'''-H_a), 2.44 (d, *J* = 12.5 Hz, 1H, 8'''-H_b), 1.95 (t, *J* = 6.2 Hz, 2H, H_H), 1.55 (quint., *J* = 6.2 Hz, 2H, H_D) 1.51-1.38 (m, 6H, H_K, H_J, H_L); ¹³C NMR (DMSO) δ 172.7 (C_G), 166.4 and 166.0 (COOH, C_A), 163.8 (C-5'''), 163.30 (C-4, C-6), 159.6 (C-2), 141.0 (C-4'), 138.9 (C-4''), 137.4 and 137.3 (C-1', C-1''), 130.3 and 128.32 (C-3', C-5', C-3'', C-5''), 128.29 and 127.9 (C-2', C-6', C-2'', C-6''), 114.2 (C-5), 61.9 (C-3'''), 59.7 (C-7'''), 56.0 (C-2'''), 53.0 (OMe), 40.6 (C-8'''), 37.6 (C_C), 36.9 (C_E) 35.8 (C_H), 29.9 (C_D), 28.8 (C_K) 28.6 (C_L), 25.9 (C_J); IR (ATR): 3281m (Ar C-H), 1699s (C=O(OMe)), 1640 (C=O(NH)), 1588 (N-H), 1467m (C-H₂) cm⁻¹; MS (ESI+): *m/z* 617.1 (100 %) (M+H)⁺, 539.3 (5 %), 473.2 (10 %), 391.1 (10 %), 374.2 (10 %), 317.1 (8 %) (M-NH(CH₂)₃NH-Biotin)⁺, 284.1 (7 %), 227.1 (7 %) (Biotin)⁺, 187.2 (6 %), 102.4 (50 %), 83.3 (38 %), 58.5 (65 %) (COOMe)⁺.

Elemental analysis found C, 62.68; H, 5.67; N, 13.72 %

C₃₂H₃₆N₆O₅S· requires: C, 62.32; H, 5.88; N, 13.63%.

2-[4-(3-[Biotin-1-ylamido]carboxyamidopropyl)carboxyamidophenyl]-4-[4-methoxycarbonylphenyl] pyridine, **91**



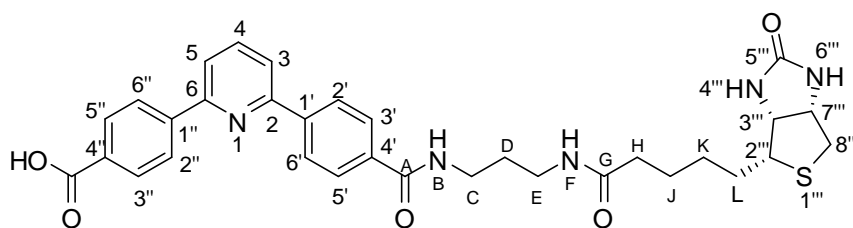
2-(4-Methoxycarbonylphenyl)-6-(4-carboxyphenyl)pyridine, **83**, (0.3 g, 0.90 mmol), PyBOP (0.52 g, 1.00 mmol) and Et₃N (0.45 g, 4.49 mmol) were stirred in DMF (10 ml) at RT for 1 h. In parallel, *N*-Boc-1-amino-3-[biotin-1-ylamido]carboxyamidopropane, **88**, (0.29 g, 0.82 mmol) was taken up in HBr (1 ml) and the HBr was evaporated under reduced pressure. The residue was added to the PyBOP reaction mixture and it was left at RT for 18 h. The DMF was evaporated under reduced pressure and the residue triturated with ethyl acetate to leave a white solid **91** (0.54 g, 83 %); m.p. 187.5-187.7 °C; ¹H NMR (DMSO) δ 8.57 (t, *J* = 5.6 Hz, 1H, NH_B), 8.38 (½ AB, *J* = 8.6 Hz, 2H, 2'-H, 6'-H), 8.32 (½ AB, *J* = 8.6 Hz, 2H, 2''-H, 6''-H), 8.08 (½ AB, *J* = 8.6 Hz, 2H, 3'-H, 5''-H), 8.09 (m, 3H, 3-H, 4-H, 5-H), 8.01 (½ AB, *J* = 8.6 Hz, 2H, 3'-H, 5'-H), 7.83 (t, *J* = 5.6 Hz, 1H, NH_F), 6.42 (s, 1H, 6'''-H), 6.35 (s, 1H, 4'''-H), 4.29 (m, 1H, 7'''-H), 4.13 (m, 1H, 3'''-H), 3.90 (s, 1H, OMe), 3.29 (q, *J* = 6.8 Hz, 2H, H_C), 3.12 (m, 3H, 2'''-H, H_E), 2.80 (dd, *J* = 5.1 Hz, 12.4 Hz, 1H, 8'''-H_a), 2.57 (d, *J* = 1245 Hz, 1H, 8'''-H_b), 2.09 (t, *J* = 7.3 Hz, 2H, H_H), 1.70 (quint., *J* = 6.8 Hz, 2H, H_D), 1.55-1.44 (m, 4H, H_J, H_L), 1.32 (m, 2H, H_K); ¹³C NMR (DMSO) δ 172.5 (C_G), 166.6 (COOH), 166.2 (C_A), 163.4 (C-5'''), 155.4 and 155.0 (C-2, C-6), 143.3 and 141.2

(C-4' or C4''), 139.2 (C-4), 135.6 and 130.5 (C-1', C-1''), 130.2 (C-3'', C-5''), 128.2 (C-3', C-5'), 127.4 (C-2'', C-6''), 127.0 (C-2', C-6'), 120.7 and 120.6 (C-3, C-5), 61.5 (C-3'''), 59.2 (C-7'''), 56.0 (C-2'''), 52.7 (OMe), 40.6 (C-8'''), 37.5 (C_C), 36.8 (C_E) 35.9 (C_H), 29.8 (C_D), 28.9 (C_K) 28.6 (C_L), 25.7 (C_J); IR (ATR): 2928m (Ar C-H), 1693s (C=O(OMe)), 1631 (C=O(NH)), 1587 (N-H), 1434m (C-H₂) 1270 (C-O) cm⁻¹; MS (ESI+): m/z 617.1 (100 %) (M+H)⁺, 539.3 (5 %), 473.2 (10 %), 391.1 (10 %), 374.2 (10 %), 317.1 (8 %) (M-NH(CH₂)₃NH-Biotin)⁺, 284.1 (7 %), 227.1 (7 %) (Biotin)⁺, 187.2 (6 %), 102.4 (50 %), 83.3 (38 %), 58.5 (65 %) (COOMe)⁺;

Elemental analysis found C, 61.87; H, 5.55; N, 61.87 %

C₃₃H₃₇N₅O₅S• $\frac{1}{3}$ HBr requires: C, 61.70; H, 5.86; N, 10.90 %.

2-[4-(3-[Biotin-1-ylamido]carboxyamidopropyl)carboxyamidophenyl]-4-[4-carboxyphenyl]pyridine, **92**



2-[4-(3-[Biotin-1-ylamido]carboxyamidopropyl)carboxyamidophenyl]-4-[4-methoxycarbonylphenyl]pyridine, **90**, (0.14 g, 0.22 mmol) and NaOH (0.18 g, 20 mmol) were stirred in a solution of water (20 ml) at 50 °C for 2 h. The basic solution was acidified with HCl (4 M) and the precipitate was collected by filtration. The resultant solid was washed with water and ether to leave a white

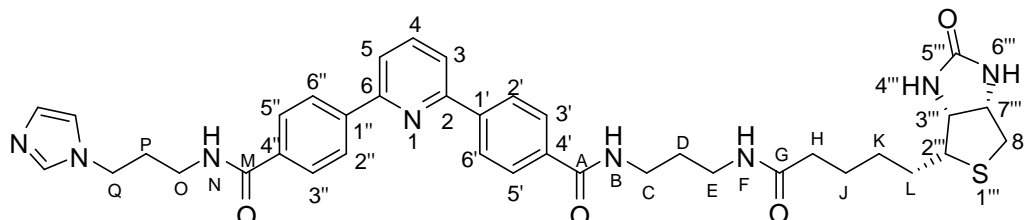
Experimental

waxy solid (0.12 g, 89 %). m.p.decomp >300 °C; ^1H NMR (DMSO) δ 8.77 (br t, 1H, NH_C), 8.33-8.30 (m, 4H, 2'-H, 6'-H, 2''-H, 6''-H), 8.08-8.05 (m, 8H, 3-H, 4-H, 5-H, 3'-H, 5'-H, 3''-H, 5''-H, NH_F), 6.44 (s, 1H, 6'''-H), 6.37 (s, 1H, 4'''-H), 4.27 (m, 1H, 7'''-H), 4.11 (m, 1H, 3'''-H), 3.29 (q, $J = 6.5$ Hz, 2H, H_C), 3.09 (m, 3H, 2'''-H, H_E), 2.79 (dd, $J = 5.0$ Hz, 12.5 Hz, 1H, 8'''- H_a), 2.55 (d, $J = 12.5$ Hz, 1H, 8'''- H_b), 2.07 (t, $J = 7.4$ Hz, 2H, H_H), 1.65 (quint., $J = 6.5$ Hz, 2H, H_D) 1.60 (m, 1H, H_{La}), 1.49 (m, 2H, H_J), 1.44 (m, 1H, H_{Lb}), 1.30 (m, 2H, H_K); ^{13}C NMR (DMSO) δ 172.6 ($\text{C}_\text{G}=\text{O}$), 167.9 (COOH), 166.2 ($\text{C}_\text{A}=\text{O}$), 163.2 (C-5'''), 155.5 (C-2 or C-6), 155.4 (C-2 or C-6), 141.4 (C-4' or C-4''), 141.3 (C-4' or C-4''), 139.3 (C-4), 135.7 (C-1' or C-1''), 135.5 (C-1' or C-1''), 130.4 (C-3', C-5' or C-3'', C-5''), 128.3 (C-3', C-5' or C-3'', C-5''), 127.3 (C-2', C-6' or C-2'', C-6''), 127.1 (C-2', C-6' or C-2'', C-6''), 120.6 (C-3, C-5), 61.6 (C-3'''), 59.8 (C-7'''), 56.0 (C-2'''), 37.5 (C_C), 36.7 (C_E), 35.8 (C_H), 29.7 (C_D), 28.9 (C_K), 28.6 (C_L), 25.9 (C_J), C-8''' under DMSO solvent peak; IR (ATR): 3282m (O-H), 2850m (Ar C-H), 1690s ($\text{C}=\text{O}(\text{OH})$), 1635 ($\text{C}=\text{O}(\text{NH})$), 1539 (N-H), 1448m (C-H₂) 1266 (C-O) cm^{-1} ; MS (ESI+): m/z 602.2 (100 %) ($\text{M}+\text{H}$)⁺, 464.6 (10 %), 442.7 (9 %), 376.1 (8 %) (M -Biotin)⁺, 320.0 (8 %), 301.1 (6 %) (M - $\text{NH}(\text{CH}_2)_3\text{NHBiotin}$)⁺, 227.1 (7 %) (Biotin)⁺, 179.1 (5 %), 151.1 (6 %), 141.1 (5 %), 105.2 (12 %), 81.4 (6 %), 53.5 (11 %);

Elemental analysis found C, 51.68; H, 4.92; N, 9.21 %

$\text{C}_{32}\text{H}_{35}\text{N}_5\text{O}_5\text{S}\cdot 3\text{HCl}\cdot 1\frac{1}{2}\text{CO}_2$ requires: C, 51.78; H, 4.93; N, 9.01 %.

2-[4-(3-[Biotin-1-ylamido]carboxyamidopropyl)carboxyamidophenyl]-4-[4-(3-[imidazol-1-yl]carboxyamidophenyl)pyridine, 16



2-[4-(3-[Imidazol-1-ylamido]carboxyamidopropyl)phenyl]-6-[4-carboxyphenyl] pyridine, **85**, (0.10 mg, 0.23 mmol), PyBOP (0.13 g, 0.26 mmol) and Et₃N (0.26 g, 2.6 mmol) were stirred in DMF (5 ml) for 1 h at room temperature. In parallel, *N*-Boc-1-amino-3-[biotin-1-ylamido]carboxyamidopropane, **88**, (0.10 g, 0.26 mmol) was taken up in HBr (1 ml) and the HBr was evaporated under reduced pressure. The residue was added to the PyBOP mixture and left to stir for 18 h. The DMF was evaporated under reduced pressure and the residue washed with ethyl acetate and acetonitrile. Recrystallisation from methanol yielded 0.11 g of product **16** (62 %). m.p. decomp >138.4-138.7 °C; ¹H NMR (DMSO) δ 8.69 (t, *J* = 5.15 Hz, 1H, NH_B or NH_N), 8.64 (s, 1H NCH_N), 8.58 (t, *J* = 5.15 Hz, 1H, NH_B or NH_N), 8.33 (m, 4H, 2'-H, 6'-H, 2''-H, 6''-H), 8.07-8.03 (m, 3H, 3-H, 4-H, 5-H), 8.00 (m, 4H, 3'-H, 5'-H, 3''-H, 5''-H), 7.86 (t, *J* = 5.15 Hz, 1H, NH_F), 7.63 (s, 1H, and 7.42 (2 x s, 1H, NCH_{CH}N), 6.43 (s, 1H, 6'''-H), 6.36 (s, 1H, 4'''-H), 4.28 (m, 1H, 7'''-H), 4.19 (t, *J* = 6.8 Hz, 2H, H_R), 4.11 (m, 1H, 3'''-H), 3.29 (m, 4H, H_C, H_P), 3.10 (m, 3H, 2'''-H, H_E), 2.79 (dd, *J* = 5.0 Hz, 12.4 Hz, 1H, 8'''-H_a), 2.55 (d, *J* = 12.4 Hz, 1H, 8'''-H_b), 2.09 (m, 4H, H_H, H_Q), 1.64 (quint., *J* = 6.8 Hz, 2H, H_D), 1.61 (m, 1H, H_{La}), 1.52 (m, 2H, H_J), 1.47 (m, 1H, H_{Lb}), 1.31 (m, 2H, H_K); ¹³C NMR (DMSO) δ 172.9 (C_G), 166.7 (C_A, C_M), 163.3 (C-5'''), 155.5 (C-2, C-6),

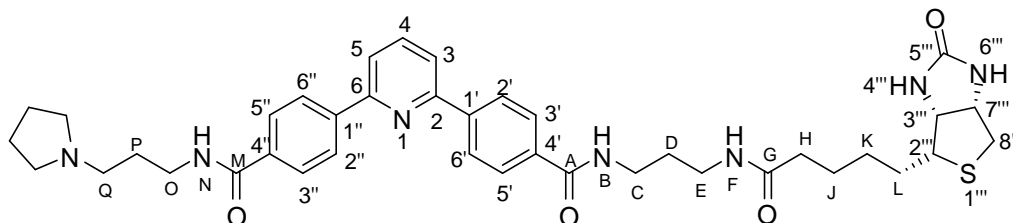
Experimental

141.5 and 141.4 (C-4', C4''), 139.2 (C-4), 135.6 (NCHN), 135.4 (C-1', C-1''), 128.4 and 128.3 (C-3', C-5', C-3'', C-5''), 127.1 (C-2', C-6', C-2'', C-6''), 121.5 and 120.5 (NCHCHN), 120.4 (C-3, C-5), 61.6 (C-3'''), 59.7 (C-7'''), 56.0 (C-2'''), 46.4 (C_R) 39.7 (C_P), 37.7 (C_C), 36.9 (C_E), 35.8 (C_H), 30.4 (C_Q), 29.8 (C_D), 28.8 (C_K), 28.6 (C_L), 25.9 (C_J), C-8''' under DMSO solvent peak; IR (ATR): 2950m (Ar C-H), 1691s (C=O(OH)), 1630 (C=O(NH)), 1536 (N-H), 1448m (C-H₂) 1250 (C-N) cm⁻¹; MS (ESI+) m/z 709.3 (100 %) (M+H)⁺, 641.2 (25 %) (M-imidazole)⁺, 584.1 (5 %) (M-NH(CH₂)₃imidazole)⁺, 483.1 (3 %), 415.1 (10 %), 409.1 (18 %) (M- NH(CH₂)₃NHBiotin)⁺, 358.1 (15 %), 341.1 (47 %), 313.0 (34 %), 283.1 (11 %), 228.2 (8 %) (Biotin)⁺, 219.2 (14 %), 187.1 (3 %), 177.2 (9 %), 151.2 (4 %) (CONH(CH₂)₃imidazole)⁺, 114.2 (22 %), 97.3 (8 %), 70.5 (18 %) (CH₅H₁₀)⁺, 56.4 (45 %) (CH₄H₈)⁺.

Elemental analysis found C, 59.44; H, 6.26; N, 14.11.

C₃₈H₄₄N₈O₄S•²/₃HBr•¹/₂H₂O requires C, 59.11; H, 5.96; N, 14.51.

2-[4-(3-[biotin-1-ylamido]carboxyamidopropyl)carboxyamidophenyl]-4-[4-(3-[pyrrolidin-1-yl]carboxyamidophenyl)pyridine, 17



2-[4-(3-[Biotin-1-ylamido]carboxyamidopropyl)phenyl]-6-[4-carboxyphenyl]pyridine, **92**, (0.10 mg, 0.17 mmol), PyBOP (0.95 g, 0.18 mmol) and Et₃N (0.50 g, 0.50 mmol) were stirred in DMF (5 ml) for 1 h at RT. 1-(3-Aminopropyl)pyrrolidine, **94**, (0.23 g, 0.18 mmol) was added and the solution stirred for 18 h at RT. The DMF was evaporated under reduced pressure and the residue washed with ethyl acetate. Recrystallisation from methanol yielded product **17** (0.089 g, 52 %). m.p. decomp >300 °C; ¹H NMR (DMSO) δ 8.69 and 8.64 (2 x t, *J* = 5.5 Hz, 1H, NH_B, NH_N), 8.31 (m, 4H, 2'-H, 6'-H, 2''-H, 6''-H), 8.07-7.99 (m, 7H, 3-H, 4-H, 5-H, 3'-H, 5'-H, 3''-H, 5''-H), 7.90 (q, *J* = 5.50 Hz, 1H, NH_F), 6.42 (s, 1H, 6'''-H), 6.35 (s, 1H, 4'''-H), 4.28 (m, 1H, 7'''-H), 4.11 (m, 1H, 3'''-H), 3.38 (q, *J* = 6.5 Hz, 2H, H_R), 3.34 (q, *J* = 6.5 Hz, 2H, H_D), 3.30 (m, 6H, H_C, HNCH₂(CH₂)₂CH₂), 3.10 (m, 3H, 2'''-H, H_E), 2.79 (dd, *J* = 5.2 Hz, 12.5 Hz, 1H, 8'''-H_a), 2.55 (d, *J* = 12.5 Hz, 1H, 8'''-H_b), 2.07 (m, 4H, H_H, H_Q), 1.93-1.89 (m, 4H, HNCH₂(CH₂)₂CH₂) 1.64 (q, *J* = 6.9 Hz, 2H, H_D) 1.59 (m, 1H, H_{La}), 1.50 (m, 2H, H_J), 1.44 (m, 1H, H_{Lb}), 1.32 (m, 2H, H_K); ¹³C NMR (DMSO) δ 172.6 (C_G), 166.5 and 166.2 (C_A, C_M), 163.4 (C-5'''), 155.5 (C-2, C-6), 141.4 (C-4', C-4''), 139.4 (C-4), 135.6 (C-1', C-1''), 128.4 and 128.2 (C-3', C-5', C-3'', C-5''), 127.1

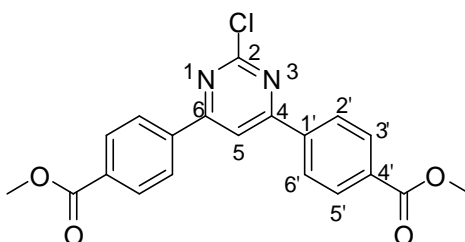
(C-2', C-6', C-2'', C-6''), 120.6 (C-3, C-5), 61.6 (C-3'''), 59.7 (C-7'''), 55.8 (C-2'''), 53.4 (NCH₂(CH₂)₂CH₂, C_R), 37.5 (C_C), 37.2 (C_P), 37.0 (C_E), 35.9 (C_H), 31.3 (C_Q) 29.7 (C_D), 28.8 (C_K), 28.6 (C_L), 25.9 (C_J), 23.0 (NCH₂(CH₂)₂CH₂), C-8''' under DMSO solvent peak; IR (ATR): 2929m (Ar C-H), 1691s (C=O(OH)), 1633 (C=O(NH)), 1536 (N-H), 1448m (C-H₂) 1305 (C-N) cm⁻¹; MS (ESI+) m/z 712.3 (100 %) (M+H)⁺, 679.2 (18 %), 641.2 (10 %) (M-pyrrolidine)⁺, 584.1 (5 %) (M-NH(CH₂)₃pyrrolidine)⁺, 415.1 (4 %), 412.1 (7 %) (M- NH(CH₂)₃NHBiotin)⁺, 358.1 (12 %), 341.1 (35 %), 313.0 (20 %), 286.0 (10 %), 283.1 (5 %), 114.2 (2 %), 84.4 (13 %), 56.4 (45 %) (CH₄H₈)⁺.

Elemental analysis found C, 45.88; H, 5.13; N, 9.71.

C₃₉H₄₉N₇O₄S·7²/₃ HCl·1¹/₃ CO₂O required C, 46.13; H, 5.44; N, 9.34 %

6.1.5 EXPERIMENTAL DETAILS FOR SECTION 3.5

2-Chloro-4,6-bis-(4-methoxycarbonylphenyl)pyrimidine, **37**



2,4,6-Trichloropyrimidine, **38**, (2.00 g, 10.9 mmol), 4-methoxycarbonylphenyl boronic acid, **27**, (3.92 g, 21.80 mmol) and aqueous sodium carbonate (3.47 g, 32.71 mmol) were stirred in glyme (150 ml) with thorough degassing with N₂ for 30 min. Pd(OAc)₂ (0.06 g, 2.50 mol%) and PPh₃ (0.14 g, 5.00 mol%) were

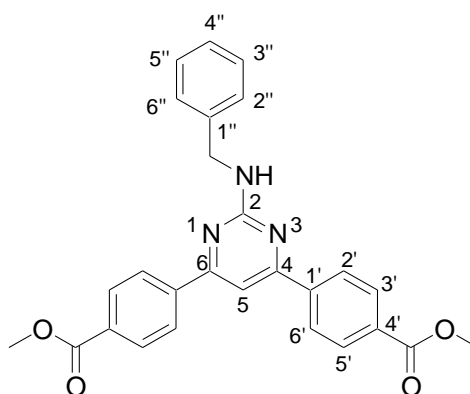
added and the reaction heated at 100 °C for 15 h. The reaction mixture was cooled and the precipitate collected by filtration. The solid was recrystallised from ethanol and washed with petroleum ether to leave a white powder **37** (3.62 g, 87%); m.p. 282.9 – 283.9°C; ¹H NMR (CDCl₃) δ 8.21 (½ AB, *J* = 8.6 Hz, 4H, 2'-H, 4'-H), 8.18 (½ AB, *J* = 8.6 Hz, 4H, 3'-H, 5'-H), 8.08 (s, 1H, 5-H), 3.96 (s, 6H, CH₃); ¹³C NMR (CDCl₃) δ 166.9 (C=O), 166.4 (C-4, C-6), 139.4 (C-4'), 133.0 (C-1'), 130.4 (C-3', C-5'), 127.6 (C-2', C-6'), 111.9 (C-5), 52.6 (OCH₃); IR (KBr): 2350m, 1750s (C=O), 1575s (Ar C-C), 1460m (Ar C-C), 1450s (C-H₃), 1375m (C-H₃), 1115 (C-O), 750s (C-Cl) cm⁻¹. MS (ESI+): *m/z* 383.2 (8 %) (M+H)⁺(³⁷Cl), 381.2 (30 %) (M+H)⁺(³⁵Cl), 352.2 (20 %) (M-OMe)⁺(³⁷Cl), 349.1 (100 %) (M-OMe)⁺(³⁵Cl), 301.0 (5 %), 279.1 (6 %), 249.2 (5 %), 196.2 (2 %), 183.1 (5 %), 151.1 (28 %), 119.2 (20 %), 105.2 (10 %), 94.4 (7 %);

Elemental analysis found C, 62.63; H, 3.72; N, 6.99.

C₂₀H₁₅N₂O₄Cl requires: C, 62.75; H, 3.95; N, 7.32 %.

General Method 3: Attachment of a linker to the 2-pyrimidine position

2-Benzylamino-4,6-bis-(4-methoxycarbonylphenyl)pyrimidine, 103

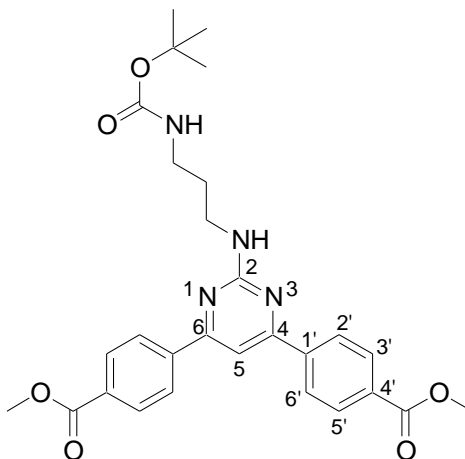


2-Chloro-4,6-bis-(4-methoxycarbonylphenyl)pyrimidine, **37**, (0.20 g, 0.52 mmol), benzylamine, **102**, (0.08 g, 0.78 mmol), potassium carbonate (0.14 g, 1.04 mmol), copper iodide (0.01 g, 0.05 mmol) and *L*-proline (0.01 g, 0.10 mmol) were stirred in DMSO (2 ml) at 80 °C for 5 h. Water (20 ml) was added and the mixture extracted with ethyl acetate (3 x 15 ml). The organic layers were combined and dried over MgSO₄. The solids were removed by filtration through a celite pad. The solvent was evaporated and the residue was recrystallised from ethanol to leave a fine solid **103** (0.15 g, 64 %); m.p. 194.1-194.8 °C; ¹H NMR (CDCl₃) δ 8.37 (½ AB, *J* = 8.3 Hz, 4H, 2'-H, 6'-H), 8.09 (½ AB, *J* = 8.3 Hz, 4H, 3'-H, 5'-H), 7.90 (s, 1H, 5-H), 7.43 (d, 2H, *J* = 7.4 Hz, 2''-H, 6''-H), 7.31 (t, *J* = 7.4 Hz, 2H, 3''-H, 5''-H), 7.20 (t, *J* = 7.4 Hz, 1H, 4''-H), 4.66 (d, *J* = 6.2 Hz, 2H, CH₂), 3.89 (s, 6H, CH₃), 3.33 (s, 1H, NH); ¹³C NMR (DMSO) δ 166.4 (C=O), 163.2 (C-4, C-6), 142.0 (C-2), 141.1 (C-4'), 131.8 (C-1'), 130.0 (C-3', C-5'), 128.8 (C-2'', C-6''), 127.91 and 127.90 (C-2', C-6', C-3''', C-5'''), 127.1 (C-4''), 103.4 (C-5), 52.9 (OCH₃), 44.9 (CH₂), C-1''' not observed; IR (KBr): 3400s (N-H), 3025w (Ar C-H), 2950w (C-H), 2350w, 1700s (C=O), 1600s (Ar C-C) 1550s (N-H), 1475m (Ar C-C), 1460m (C-H₃), 1450m (C-H₂), 1275s (C-N), 1100s (C-O) cm⁻¹; MS (ESI⁺): *m/z* 454 (100 %) (M+H)⁺, 381 (8 %) (M-MeCOOMe)⁺, 353 (5 %), 301 (4 %), 241 (18 %), 236 (7 %), 213 (5 %), 196 (13 %), 170 (2 %), 165 (15 %), 151 (6 %), 133 (18 %), 117 (7 %), 101 (23 %), 91 (29 %) (PhCH₂)⁺, 79 (28 %) (Ph+H)⁺, 74 (4 %) (MeCOOMe)⁺.

Elemental analysis found C, 69.66; H, 5.27; N, 9.03 %

C₂₇H₂₃N₃O₄•²/₃ H₂O requires: C, 69.71; H, 5.26; N, 9.03 %.

2-(*N*-Boc-1,3-diaminopropane)-4,6-bis-(4-methoxycarbonylphenyl)pyrimidine, 104



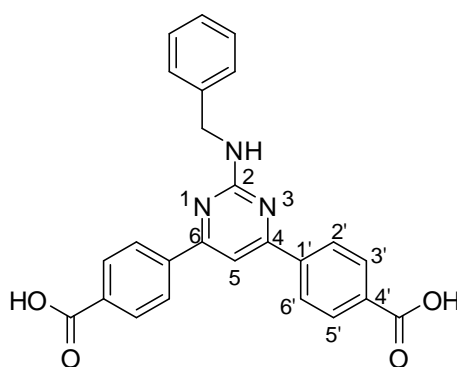
Prepared according to general method 3, from 2-chloro-4,6-bis-(4-methoxy carbonylphenyl)pyrimidine, **37**, (1.00 g, 2.60 mmol) and *N*-Boc-1,3-diaminopropane, **71**, (0.68 g, 3.90 mmol). Recrystallisation from ethanol yielded a white solid (0.65 g, 48 %); m.p. 163.5-164.5 °C; ^1H NMR (DMSO) δ 8.38 ($\frac{1}{2}$ AB, $J = 8.3$ Hz, 4H, 2'-H, 6'-H), 8.08 ($\frac{1}{2}$ AB, $J = 8.3$ Hz, 4H, 3'-H, 5'-H), 7.87 (s, 1H, 5-H), 7.44 (t, $J = 5.6$ Hz, 1H, NHCH_2), 6.85 (t, $J = 5.6$ Hz, 1H, NHCO), 3.90 (s, 6H, OCH_3), 3.44 (q, $J = 6.5$ Hz, 2H, CH_2NHAr), 3.03 (q, $J = 6.5$ Hz, 2H, CH_2NHCO), 1.72 (quint., $J = 6.5$ Hz, 2H, $\text{CH}_2\text{CH}_2\text{CH}_2$), 1.35 (s, 9H, ^tBu); ^{13}C NMR (DMSO) δ 166.4 (COOCH_3), 163.3 (COOBu^t), 156.2 (C-2), 142.0 (C-4'), 131.7 (C-1'), 130.0 (C-3', C-5'), 127.9 (C-2', C-6'), 103.0 (C-5), 78.0 (CMe_3), 52.9 (OCH_3), 39.1 (ArNHCH_2), 38.3 (CH_2NHBoc), 30.0 ($\text{NHCH}_2\text{CH}_2\text{CH}_2\text{NH}$), 28.8 ($\text{C(CH}_3)_3$); IR (KBr): 3400s (N-H), 3025w (Ar C-H), 2950m (C-H), 2350m, 1750s (C=O), 1580s (Ar C-C), 1550s (N-H), 1450m (C-H₃), 1450m (C=C), 1275s (C-N), 1100s (C-O) cm^{-1} ; MS (ESI+): m/z 521 (100 %) ($\text{M}+\text{H}$) $^{+}$, 241 (18 %), 197

(7 %), 165 (34 %), 151 (10 %), 133 (31 %), 119 (10 %), 101 (10 %) (Boc)⁺, 79 (6 %) (Ph+H)⁺, 61 (4 %) (CH₃COOH+H)⁺.

Elemental analysis found C, 64.66; H, 5.83; N, 10.59 %

C₂₈H₃₂N₄O₆ requires: C, 64.60; H, 6.20; N, 10.76 %.

2-Benzylamine-4,6-bis-(4-carboxyphenyl)pyrimidine, **105**



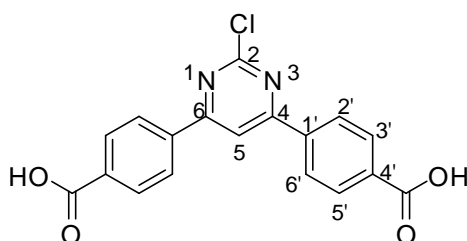
2-Benzylamino-4,6-bis-(4-methoxycarbonylphenyl)pyrimidine, **104**, (82 mg, 0.18 mmol) was heated at 100°C in a solution of sodium hydroxide (2 M, 10 ml) for 18 h. The solution was acidified with HCl (4 M) and the precipitate collected by filtration and washed with water and chloroform to leave a fine powder (76 mg, 75 %); m.p. decomp. > 250°C; ¹H NMR (DMSO) δ 8.11 (½ AB, *J* = 8.1 Hz, 4H, 2'-H, 6'-H), 7.93 (½ AB, *J* = 8.1 Hz, 4H, 3'-H, 5'-H), 7.82 (t, *J* = 6.4 Hz, 1H, NH), 7.01 (s, 1H, 5-H), 7.43 (d, *J* = 7.6 Hz, 2-H, 2''-H, 6''-H), 7.30 (t, *J* = 7.6 Hz, 2-H, 3''-H, 5''-H), 7.19 (t, *J* = 7.4 Hz, 1H, 4''-H), 4.64 (d, *J* = 6.36 Hz, 2H, CH₂); ¹³C NMR (DMSO) δ 169.6 (C=O), 163.2 (C-4, C-6), 143.3 (C-2), 141.5 (C-4'), 137.7 (C-1'), 129.7 (C-3', C-5'), 128.7 (C-2'', C-6''), 127.9 (C-3'', C-5''), 127.0 (C-2', C-6'), 126.5 (C-4''), 102.4 (C-5), 44.9 (CH₂); IR (KBr): 3400s (N-H), 3350br (O-H) 2975w (Ar C-H), 2350m (C-H), 1600s (Ar C-C), 1575s (C=O), 1550s

(N-H), 1500m (C-H₂), 1425m (C=C), 1400s (C-N), 1050m (C-O); MS (ESI+): m/z 448 (12%) (M+H+Na)⁺⁺, 426 (100 %) (M+H)⁺⁺, 381 (4 %) (M-CO₂)⁺⁺, 117 (2 %), 91 (5 %) (PhCH₂)⁺⁺, 85 (10 %).

Elemental analysis found C, 53.77; H, 4.18; N, 7.61 %

C₂₅H₁₉N₃O₄•1.5H₂O•3HCl requires: C, 54.81; H, 4.24; N, 7.59 %.

2-Chloro-4,6-bis-(4-carboxyphenyl)pyrimidine, **36**



2,4,6-Trichloropyrimidine, **38**, (0.50 g, 2.73 mmol), 4-carboxyphenylboronic acid pinacol ester, **81**, (1.35 g, 5.44 mmol) and aqueous sodium carbonate (0.87 g, 8.21 mmol) were stirred in glyme (50 ml) with thorough degassing with N₂ for 1 h. Pd(OAc)₂ (0.02 g, 2.50 mol%) and PPh₃ (0.04 g, 5.00 mol %) were added and the mixture was heated at 100 °C for 18 h. The reaction mixture was cooled and the precipitate collected by filtration. The solid was washed with petroleum ether and the solid triturated with hot methanol to leave a solid (0.69 g, 72%); m.p. decomp. > 250 °C; ¹H NMR (DMSO) δ 8.81 (s, 1H, 5-H), 8.47 (½ AB, J = 8.4 Hz, 4H, 2'-H, 6'-H), 8.11 (½ AB, J = 8.4 Hz, 4H, 3'-H, 5'-H); ¹³C NMR (DMSO) δ 167.3 (C=O), 166.9 (C-4, C-6), 161.4 (C-2), 139.1 (C-4'), 134.2 (C-1'), 130.5 (C-3', C-5'), 128.4 (C-2', C-6'), 113.7 (C-5); IR (KBr): 3400br (O-H), 3100s (Ar C-H) 1700s (C=O), 1600s (Ar C-C), 1460m (Ar C-C), 1235 (C-O), 750s (C-Cl) cm⁻¹. MS (ESI-): m/z 355.0 (22 %) (M-H)⁻ (³⁷Cl), , 353.0 (55 %) M-H)⁻ (³⁵Cl), 335.1 (3 %), 311.0 (2 %) (M-COOH)⁻(³⁷Cl), 309.0 (5 %)

(M-COOH)⁻ (³⁵Cl), 251.1 (7 %), 249.1 (18 %), 241.2 (20 %), 212.3 (18 %), 198.0 (12 %), 163.0 (5 %), 145.2 (100 %), 137.2 (16 %), 121.2 (8 %) (PhCOOH)⁻, 93.3 (10 %).

EXPERIMENTAL DETAILS FOR CHAPTER 4: BIOLOGICAL EVALUATION AND PROTEOMICS

MATERIALS

Acrylamide (30%), Dodeca silver stain kit, and trishydroxymethyl aminomethane were purchased from Bio-Rad. Peptide calibrant standard II (ACTH clip 1-17, ACTH clip 18-39, angiotensin I, angiotensin II, bombesin, bradykinin 1-7, renin substrate, somatostatin 28, substance P) was purchased from Bruker Daltonics (Bremen, Germany). Complete mini EDTA-free protease inhibitor cocktail tablets and trypsin were purchased from Roche Diagnostics (Mannheim, Germany). NAP-5 columns were purchased from GE Healthcare. All other materials, including streptavidin beads and HPLC solvents were purchased from Sigma-Aldrich, Gillingham, UK or Lancaster, Morecambe, UK. The Human ovarian carcinoma A2780 (#93112520) and human colon adenocarcinoma HT29 (#91072201) were obtained from the ECACC.

CELL MAINTENANCE

Cell lines were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with fetal bovine serum (10 %), L-glutamine (2 mM) and sodium pyruvate (1 mM). Cells lines were maintained at 37° C in a 5 % CO₂ humidified atmosphere.

Cells were maintained in exponential growth by subculturing every 3-4 days at approximately 75 % confluence. Adhered cells were washed with HBSS (2 x 10 ml) and trypsinisation was carried out using trypsin-EDTA solution (0.25 %, 3 ml). Upon the cells detachment from the flask, the trypsin solution was diluted with medium (10 ml). The cell suspension was centrifuged (200 g, 5 mins) and the supernatant discarded. The cell pellet was either re-suspended in fresh culture medium (10 ml) and transferred (1 ml) to a new flask containing fresh medium (19 ml), or frozen at -80 °C for proteomics.

MTT ASSAY

Cells at approximately 75 % confluence were washed with HBSS (2 x 10 ml), trypsinised and resuspended in fresh medium (10 ml). The cell concentration was determined by counting the cells with a haemocytometer and the cells were plated at 2000 cells per well in 200 µl of medium (the first lane was left blank) and incubated at 37 °C for 24 h.

The medium was removed and fresh medium (180 μ l) was added to the wells (fresh medium was also added to lane 1 (200 μ l)). DMSO solution (0.1 %, 20 μ l) was added to wells in lane 2. Serial dilutions of drugs ranging from 1 mM – 2 μ M, were prepared in culture medium and added (20 μ l) to cells in lanes 3-12, giving a drug concentration of 100 μ M – 0.2 μ M. Cells were left to incubate for 96 h at 37 °C.

MTT solution was made by diluting a stock MTT solution (0.5 mg/ml, 2 ml) with culture medium (18 ml). Medium was removed from all wells and replaced with MTT solution and the plate was left to incubate for 4 h at 37 °C. The MTT solution was removed and the resultant formazan crystals were dissolved in DMSO (150 μ l). The absorbance of each lane was measured at 540 nm.

Analysis of Chemosensitivity Data

The percentage of cell survival was calculated using the following equation from the mean absorbance for all eight wells in each lane:

$$\% \text{ Cell Survival} = \frac{A_{540} - A_{540}^{Blank}}{A_{540}^{Control} - A_{540}^{Blank}} \times 100$$

Equation 6.1. Where blank is lane 1 containing culture medium and control is lane 2 containing cells treated with DMSO (0.1 %).

Percentage cell survival was plotted against drug concentration and IC₅₀ values obtained from the concentration of drug when cell survival was equal to 50 % of the control.

NAP-5 COLUMN

A NAP-5 column was equilibrated with PBS or ambic (~10 ml). The sample was loaded on the column (500 µl, diluted with buffer if necessary). The sample was eluted using the same buffer used for equilibration (1 ml). Desalted samples were lyophilised and stored at -20 °C.

AFFINITY CHROMATOGRAPHY

Method 1

A2780 cell pellets (3) were thawed and washed with PBS (3 x 1 ml, 15 min). They were resuspended in Tris HCl (pH 7.5, 100 µl) containing CHAPs (4%) and complete protease inhibitor cocktail (14 µl). The cell suspensions were combined and sonicated using a SH70A sonicator (2 x 20 s). The suspensions were centrifuged (4 °C, 13000 g, 10 min). The DNA was extruded from cell pellets with a fine pipette needle (10 x) and centrifuged (4 °C, 13000 g, 10 min). The supernatant was collected and split in two. An affinity probe (**16** or **17**, 10 µl, 400 µM in Tris-HCl) was added to the supernatant (190 µl), to give a final affinity probe concentration of 20 µM, and incubated for 2 h at room

temperature on a carousel. Protein samples were desalted with a NAP-5 column using PBS.

Streptavidin beads (2 x 100 μ l) were washed with PBS (3 x 1 ml). The desalted protein solutions were added to the streptavidin beads and incubated at RT for 4 h on a carousel. The bead suspensions were centrifuged (13000 *g*, 1 min) and the supernatant removed and stored. The streptavidin beads were washed with PBS (3 x 1 ml, 15 mins) and Urea (15 mins, 300 μ l, 8M in ambic (25 mM)). The urea elutions were desalted with a NAP-5 column using ambic (25 mM).

Method 2

Carried out according to method 1 (Section 6.1.3.1) using A2780 cell pellets (6) and incubating protein suspensions (290 μ l) with affinity probes (**16** or **17**, 10 μ l, 3 mM in Tris-HCl), to give a final drug concentration of 100 μ M, for 24 h.

Method 3

Streptavidin beads (6 x 100 μ l) were washed with PBS (3 x 1 ml, 15 min), resuspended in PBS (290 μ l) and incubated with either **16** (10 μ l, 3 mM), **17** (10 μ l, 3 mM) or no compound for 2 h. In parallel, Cell pellets (9) from A2780 and HT29 cell lines were thawed and washed with PBS (3 x 1 ml, 15 min). They were resuspended in Tris HCl (pH 7.5, 100 μ l) containing CHAPs (4 %) and complete protease inhibitor cocktail (14 μ l). The cell suspensions were combined and sonicated using a SH70A sonicator (2 x 20 s). The suspensions

were centrifuged (4 °C, 13000 g, 10 min). The DNA was extruded from cell pellets with a fine pipette needle (10 x) and centrifuged (4 °C, 13000 g, 10 min). The supernatant was collected.

The streptavidin beads suspensions were centrifuged (13000 g, 1 min) and the supernatant removed. The protein suspensions (300 µl) were added to the streptavidin beads and left to incubate at room temperature for 24 h on a carousel. The bead suspensions were centrifuged (13000 g, 1 min) and the supernatant removed and stored. The streptavidin beads were washed with PBS (3 x 1 ml, 15 min) and Urea (15 min, 300 µl, 8 M in ambic (25 mM)). The urea elutions were desalted with a NAP-5 column using ambic (25 mM).

Method 4

Streptavidin beads (12 x 100 µl) were washed with PBS (3 x 1 ml, 15 min), resuspended in PBS (100 µl) and incubated with either **16** (100 µl, 1 mM), **17** (100 µl, 1 mM) or no compound for 24 h at 4 °C. In parallel, Cell pellets (9) from A2780 and HT29 cell lines were thawed and washed with PBS (3 x 1 ml, 15 min). They were resuspended in PBS (500 µl) containing Triton X (1 %), and complete protease inhibitor cocktail (14 µl). The cell suspensions were sonicated using a SH70A sonicator (2 x 20 s). The suspensions were centrifuged (4 °C, 13000 g, 10 min). The supernatant was collected and the cell pellets were suspended in PBS (500 µl) containing C7BzO (1 %), CHAPs (4%) and complete protease inhibitor cocktail (14 µl). The cell suspensions sonicated using a SH70A sonicator (2 x 20 s). The suspensions were centrifuged (4 °C,

13000 *g*, 10 min). The DNA was extruded from cell pellets with a fine pipette needle (10 x) and centrifuged (4 °C, 13000 *g*, 10 min). The supernatant was collected.

The streptavidin bead suspensions were centrifuged (13000 *g*, 1 min) and the supernatant removed. The protein suspensions (300 µl) were added to the streptavidin beads and left to incubate at room temperature for 24 h on a carousel. The bead suspensions were centrifuged (13000 *g*, 1 min) and the supernatant removed and stored. The streptavidin beads were washed with PBS (3 x 1 ml, 15 min) and Compound **2** (3 x 100 µl, 2 mM) in PBS. The competition elutions were desalted with a NAP-5 column using ambic (25 mM).

SDS-PAGE

Resolving gel [HPLC water (750 µl), acrylamide (30 %, 2 ml), Tris HCl (1.5 M, pH 8.8, 1.95 ml), SDS (10 %, 50 µl), APS (10%, 50 µl), TEMED (2 µl)] was transferred to mini-PROTEAN casting plates attached to a casting frame on a casting stand. Isopropanol (1 ml) was added and the gel was left to set for 45 min. The isopropanol was removed and the top of the gel washed with water (3 x 1 ml). Stacking gel [HPLC water (3.4 ml), acrylamide (30 %, 830 µl), Tris HCl (0.5 M, pH 6.8, 630 µl), SDS (10 %, 50 µl), APS (10%, 50 µl), TEMED (5 µl)] was added on top of the set resolving gel and a comb inserted to create the wells. The gel was left to set for 20 min. The gels were removed from the casting frame and inserted into a mini-PROTEAN electrophoresis system and

filled with SDS running buffer [Tris HCl (5 mM), Glycine (38.4 mM), SDS (0.02%)].

Samples were prepared by heating equal volumes of SDS buffer [water (3.55 ml water, Tris-HCl (0.5 M, 1.25 ml), Glycerol (2.5 ml), SDS (10 %, 2.0 ml), (0.5 %, 0.2 ml)] and protein sample to 70 °C for 20 min. The comb was removed and the samples (10-15 µl) loaded into the wells. The gels were run at 80 V for 15 min and then 150 V for 90 min. Gels were stained by immersion in Instant Blue stain for 30 min and then stored in water at 4 °C.

Silver staining

Pre-prepared solutions from the Bio-Rad Dodeca silver staining kit were used. Gels were destained for 24 h by immersing in a solution of water (40 ml), ethanol (32 ml) and acetic acid (8 ml). The gel was sensitised using a solution of Bio-Rad sensitising concentrate (8 ml), ethanol (24 ml), Bio-Rad background reducer concentrate (800 µl) and water (47.2 ml) for 30 min, washed with water (3 x 80 ml, 5 min), stained with a solution of Bio-Rad silver reagent concentrate (1.6 ml) in water (78.4 ml) for 20 min in the dark and rinsed with water (80 ml, 1 min). The gel was developed by immersing in a solution of Bio-Rad development buffer concentrate (8 ml), Bio-Rad image development concentrate (16 µl), Bio-Rad background reducing concentrate (4 µl) and water (72 ml) until the bands were visible (5 – 15 min) and stopped using a solution of acetic acid (4 ml) in water (76 ml) for 10 min. The gels were stored in water at 4 °C.

TRYPsin DIGESTION

Samples were dissolved in MeCN (5 %, 5 μ l or 50 μ l for beads). DTT (50 mM, 1 μ l or 5 μ l for beads) was added and the solutions heated to 70 °C for 20 min. Samples were alkylated with iodoacetamide (100 mM, 1 μ l or 5 μ l for beads) in the dark at room temperature for 20 min. Trypsin (10 μ g/ml in ambic, 1 μ l or 5 μ l for beads) was added and incubated at 37 °C for 18 h. Samples were desalted using C18 cartridges, lyophilised and stored at -20 °C.

IN-GEL TRYPsin DIGESTION

Protein bands were excised using a scalpel and the gel pieces were dehydrated by incubating in MeCN (200 μ l, 15 min). The solvent was removed and the gel pieces rehydrated by incubating in ambic (20 μ l, 15 min). The process was repeated three times and a final dehydration with MeCN (200 μ l, 15 min) was carried out. The solvent was removed and fragments left to dry (30 min, 37 °C). The gel was incubated in a solution of trypsin (20 μ l, 10 ng/ μ l in ambic (25 mmol), 18 h, 28 °C). The supernatants were desalted using C18 cartridges, lyophilised and stored at -20 °C.

C18 CARTRIGES

Isolute C18 cartridges (Kinesis Ltd., Cambridgeshire, England) were wetted with methanol (1 ml) and equilibrated with acetonitrile (2 %, 2 x 1 ml). The trypsin digested sample was loaded and made up to a total volume of 500 μ l with

acetonitrile (2 %). The column was washed with acetonitrile (2 %, 2 x 1 ml) and eluted with acetonitrile (80%, 1 ml). Samples were lyophilised and stored at -20 °C.

MASS SPECTROMETRY

Mass spectra were generated for desalted, trypsin-digested peptides using a MALDI-TOF/TOF UltraFlex II Instrument (Bruker Daltonics, Bremen, Germany). α -CHCA (1 μ l) was applied onto a MTP 384 'massive' target T plate (Bruker Daltonics, Bremen, Germany), followed by peptides (0.5 μ l) and then α -CHCA (0.5 μ l). Mass spectrometry was carried out in reflector positive mode using a mass range of 700-4000 Da and a total of 500 cumulative shots. MS/MS was carried out using LIFT mode and peak list and a total of 1500 cumulative shots. Peak lists for MS and MS/MS data were generated using FlexAnalysis version 3.0 software.

Proteins were identified by their peptide fragmentation fingerprints or MS/MS fragments. Peaklists were compiled into a batch using Bruker BioTools version 3.2 (BrukerDaltonics, Bremen, Germany) and searched using Mascot software version 2.2 (Matrix Science, USA) against a Swiss-Prot database.

REFERENCES

REFERENCES

1. World Health Organization.; Boyle, P.; International Agency for Research on Cancer., *World cancer report*. IARC Press: Lyon, **2009**.
2. Pecorino, L., *Molecular biology of cancer: mechanisms, targets, and therapeutics*. Oxford University Press Inc.: New York, **2005**.
3. Neidle, S., *Cancer drug design and discovery*. Academic: London, **2007**.
4. Sawyers, C. L., Opportunities and challenges in the development of kinase inhibitor therapy for cancer. *Gene Dev.* **2003**, 17, (24), 2998-3010.
5. Sikic, B. I., Anticancer drug discovery. *J. Natl. Cancer Inst.* **1991**, 83, (11), 738-740.
6. Pors, K.; Goldberg, F. W.; Leamon, C. P.; Rigby, A. C.; Snyder, S. A.; Falconer, R. A., The changing landscape of cancer drug discovery: a challenge to the medicinal chemist of tomorrow. *Drug Discov. Today* **2009**, 14, (21-22), 1045-1050.
7. Collins, I.; Workman, P., New approaches to molecular cancer therapeutics. *Nat. Chem. Biol.* **2006**, 2, (12), 689-700.
8. Hambley, T. W.; Hait, W. N., Is anticancer drug development heading in the right direction? *Cancer Res.* **2009**, 69, (4), 1259-1262.
9. Zubrod, C.; Schepartz, S.; Leiter, J.; Endicott, K.; Carrese, L.; Baker, C., The chemotherapy program of the National Cancer Institute: history, analysis and plans. *Cancer Chemother. Rep.* **1966**, 50, (7), 349-96.
10. Johnson, J. I.; Decker, S.; Zaharevitz, D.; Rubinstein, L. V.; Venditti, J. M.; Schepartz, S.; Kalyandrug, S.; Christian, M.; Arbuck, S.; Hollingshead, M.; Sausville, E. A., Relationships between drug activity in

- NCI preclinical *in vitro* and *in vivo* models and early clinical trials. *Br. J. Cancer* **2001**, 84, (10), 1424-1431.
11. Shoemaker, R. H.; Scudiero, D. A.; Melillo, G.; Currens, M. J.; Monks, A. P.; Rabow, A. A.; Covell, D. G.; Sausville, E. A., Application of high-throughput, molecular-targeted screening to anticancer drug discovery. *Curr. Top. Med. Chem.* **2002**, 2, (3), 229-46.
 12. Sausville, E. A.; Burger, A. M.; Becher, O. J.; Holland, E. C., Contributions of human tumor xenografts to anticancer drug development. *Cancer Res.* **2006**, 66, (7), 3351-3354.
 13. Alley, M. C.; Scudiero, D. A.; Monks, A.; Hursey, M. L.; Czerwinski, M. J.; Fine, D. L.; Abbott, B. J.; Mayo, J. G.; Shoemaker, R. H.; Boyd, M. R., Feasibility of drug screening with panels of human tumor cell lines using a microculture tetrazolium assay. *Cancer Res.* **1988**, 48, (3), 589-601.
 14. Decker, S.; Hollingshead, M.; Bonomi, C. A.; Carter, J. P.; Sausville, E. A., The hollow fibre model in cancer drug screening: the NCI experience. *Eur. J. Cancer* **2004**, 40, (6), 821-826.
 15. Shoemaker, R. H., The NCI60 human tumour cell line anticancer drug screen. *Nat. Rev. Cancer* **2006**, 6, (10), 813-823.
 16. Monks, A.; Scudiero, D.; Skehan, P.; Shoemaker, R.; Paull, K.; Vistica, D.; Hose, C.; Langley, J.; Cronise, P.; Vaigrowloff, A.; Graygoodrich, M.; Campbell, H.; Mayo, J.; Boyd, M., Feasibility of a high-flux anticancer drug screen using a diverse panel of cultured human tumor-cell lines. *J. Natl. Cancer Inst.* **1991**, 83, (11), 757-766.

References

17. Monks, A.; Scudiero, D. A.; Johnson, G. S.; Paull, K. D.; Sausville, E. A., The NCI anti-cancer drug screen: a smart screen to identify effectors of novel targets. *Anti-Cancer Drug Des.* **1997**, 12, (7), 533-541.
18. Boyd, M. R.; Paull, K. D., Some practical considerations and applications of the national cancer institute in vitro anticancer drug discovery screen. *Drug Dev. Res.* **1995**, 34, (2), 91-109.
19. Paull, K. D.; Shoemaker, R. H.; Hodes, L.; Monks, A.; Scudiero, D. A.; Rubinstein, L.; Plowman, J.; Boyd, M. R., Display and analysis of patterns of differential activity of drugs against human-tumor cell-lines - development of mean graph and compare algorithm. *J. Natl. Cancer Inst.* **1989**, 81, (14), 1088-1092.
20. Zaharevitz, D. W.; Holbeck, S. L.; Bowerman, C.; Svetlik, P. A., COMPARE: a web accessible tool for investigating mechanisms of cell growth inhibition. *J. Mol. Graph. Model.* **2002**, 20, (4), 297-303.
21. Weinstein, J. N.; Kohn, K. W.; Grever, M. R.; Viswanadhan, V. N.; Rubinstein, L. V.; Monks, A. P.; Scudiero, D. A.; Welch, L.; Koutsoukos, A. D.; Chiausa, A. J.; Paull, K. D., Neural computing in cancer drug development - predicting mechanism of action. *Science* **1992**, 258, (5081), 447-451.
22. Scherf, U.; Ross, D. T.; Waltham, M.; Smith, L. H.; Lee, J. K.; Tanabe, L.; Kohn, K. W.; Reinhold, W. C.; Myers, T. G.; Andrews, D. T.; Scudiero, D. A.; Eisen, M. B.; Sausville, E. A.; Pommier, Y.; Botstein, D.; Brown, P. O.; Weinstein, J. N., A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.* **2000**, 24, (3), 236-244.

References

23. Ross, D. T.; Scherf, U.; Eisen, M. B.; Perou, C. M.; Rees, C.; Spellman, P.; Iyer, V.; Jeffrey, S. S.; Van de Rijn, M.; Waltham, M.; Pergamenschikov, A.; Lee, J. C. E.; Lashkari, D.; Shalon, D.; Myers, T. G.; Weinstein, J. N.; Botstein, D.; Brown, P. O., Systematic variation in gene expression patterns in human cancer cell lines. *Nat. Genet.* **2000**, 24, (3), 227-235.
24. Rabow, A. A.; Shoemaker, R. H.; Sausville, E. A.; Covell, D. G., Mining the National Cancer Institute's tumor-screening database: Identification of compounds with similar cellular activities. *J. Med. Chem.* **2002**, 45, (4), 818-840.
25. Wallqvist, A.; Rabow, A. A.; Shoemaker, R. H.; Sausville, E. A.; Covell, D. G., Linking the growth inhibition response from the National Cancer Institute's anticancer screen to gene expression levels and other molecular target data. *Bioinformatics* **2003**, 19, (17), 2212-2224.
26. Giuliani, A.; Colosimo, A.; Benigni, R.; Zbilut, J. P., On the constructive role of noise in spatial systems. *Phys. Lett. A* **1998**, 247, (1-2), 47-52.
27. Huang, R.; Wallqvist, A.; Covell, D. G., Assessment of in vitro and in vivo activities in the National Cancer Institute's anticancer screen with respect to chemical structure, target Specificity, and mechanism of action. *J. Med. Chem.* **2006**, 49, (6), 1964-1979.
28. 3DMIND <http://spheroid.ncifcrf.gov/spheroid/>.
29. Bradshaw, T. D.; Stevens, M. F. G.; Westwell, A. D., The discovery of the potent and selective antitumour agent 2-(4-amino-3-methylphenyl)benzothiazole (DF 203) and related compounds. *Curr. Med. Chem* **2001**, 8, (2), 203-210.

References

30. Bradshaw, T. D.; Westwell, A. D., The development of the antitumour benzothiazole prodrug, Phortress, as a clinical candidate. *Curr. Med. Chem* **2004**, 11, (8), 1009-1021.
31. NCI <http://dtp.nci.nih.gov/timeline/flash/index.htm>.
32. Weinstein, I. B.; Joe, A. K., Mechanisms of disease: oncogene addiction - a rationale for molecular targeting in cancer therapy. *Nat. Clin. Prac. Oncol.* **2006**, 3, (8), 448-457.
33. Merlo, L. M. F.; Pepper, J. W.; Reid, B. J.; Maley, C. C., Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* **2006**, 6, (12), 924-935.
34. Hanahan, D.; Weinberg, R. A., The hallmarks of cancer. *Cell* **2000**, 100, (1), 57-70.
35. Hollstein, M.; Rice, K.; Greenblatt, M. S.; Soussi, T.; Fuchs, R.; Sorlie, T.; Hovig, E.; Smith-Sorensen, B.; Montesano, R.; Harris, C. C., Database of p53 gene somatic mutations in human tumors and cell lines. *Nucleic Acids Res.* **1994**, 22, (17), 3551-5.
36. Davies, H.; Bignell, G. R.; Cox, C.; Stephens, P.; Edkins, S.; Clegg, S.; Teague, J.; Woffendin, H.; Garnett, M. J.; Bottomley, W.; Davis, N.; Dicks, E.; Ewing, R.; Floyd, Y.; Gray, K.; Hall, S.; Hawes, R.; Hughes, J.; Kosmidou, V.; Menzies, A.; Mould, C.; Parker, A.; Stevens, C.; Watt, S.; Hooper, S.; Wilson, R.; Jayatilake, H.; Gusterson, B. A.; Cooper, C.; Shipley, J.; Hargrave, D.; Pritchard-Jones, K.; Maitland, N.; Chenevix-Trench, G.; Riggins, G. J.; Bigner, D. D.; Palmieri, G.; Cossu, A.; Flanagan, A.; Nicholson, A.; Ho, J. W. C.; Leung, S. Y.; Yuen, S. T.; Weber, B. L.; Seigler, H. F.; Darrow, T. L.; Paterson, H.; Marais, R.;

References

- Marshall, C. J.; Wooster, R.; Stratton, M. R.; Futreal, P. A., Mutations of the BRAF gene in human cancer. *Nature*. **2002**, 417, (6892), 949-954.
37. Santarius, T.; Shipley, J.; Brewer, D.; Stratton, M. R.; Cooper, C. S., A census of amplified and overexpressed human cancer genes. *Nat. Rev. Cancer* **2010**, 10, (1), 59-64.
38. Chatterjee-Kishore, M.; Miller, C. P., Exploring the sounds of silence: RNAi-mediated gene silencing for target identification and validation. *Drug Discov. Today* **2005**, 10, (22), 1559-65.
39. Sleno, L.; Emili, A., Proteomic methods for drug target discovery. *Curr. Opin. Chem. Biol.* **2008**, 12, (1), 46-54.
40. Haberman, A., Strategies to move beyond target validation: Targets and druggability for small and large molecule drugs. *Genet. Eng. Biotechn. News* **2005**, 25, (21), 36.
41. Yang, Y.; Adelstein, S. J.; Kassis, A. I., Target discovery from data mining approaches. *Drug Discov. Today* **2009**, 14, (3-4), 147-154.
42. Weinstein, I. B.; Joe, A.; Felsher, D., Oncogene addiction. *Cancer Res.* **2008**, 68, (9), 3077-3080.
43. Smith, C., Drug target identification: A question of biology. *Nature* **2004**, 428, (6979), 225-231.
44. Terstappen, G. C.; Schlupen, C.; Raggiaschi, R.; Gaviraghi, G., Target deconvolution strategies in drug discovery. *Nat. Rev. Drug Discov.* **2007**, 6, (11), 891-903.
45. Metcalf, B. W.; Dillon, S., *Target validation in drug discovery*. Academic Press: Amsterdam; Oxford, **2007**; p xii, 279 p.

References

46. Aherne, G. W.; McDonald, E.; Workman, P., Finding the needle in the haystack: why high-throughput screening is good for your health. *Breast Cancer Res.* **2002**, 4, (4), 148-54.
47. Carr, R. A. E.; Congreve, M.; Murray, C. W.; Rees, D. C., Fragment-based lead discovery: leads by design. *Drug Discov. Today* **2005**, 10, (14), 987-992.
48. Everts, S., Piece by piece. *Chem. Eng. News* **2008**, 86, (29), 15-23.
49. Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D., Application of the 3-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* **1994**, 37, (8), 1035-1054.
50. Grosdidier, S.; Totrov, M.; Fernandez-Recio, J., Computer applications for prediction of protein-protein interactions and rational drug design. *Adv. Appl. Bioinfo. Chem.* **2009**, 2, 101-123.
51. Sun, H.; Scott, D. O., Structure-based drug metabolism predictions for drug design. *Chem. Biol. Drug Des.* **2010**, 75, (1), 3-17.
52. Keseru, G. M.; Makara, G. M., Hit discovery and hit-to-lead approaches. *Drug Discov. Today* **2006**, 11, (15-16), 741-748.
53. Liszewski, K., Drug discovery: Successful lead optimization strategies *Genet. Eng. Biotechn. News* **2006**, 26, (14).
54. Jorgensen, W. L., Efficient drug lead discovery and optimization. *Accounts Chem. Res.* **2009**, 42, (6), 724-733.
55. Raymond, J. W.; Watson, I. A.; Mahoui, A., Rationalizing lead optimization by associating quantitative relevance with molecular structure modification. *J. Chemical Inf. Model.* **2009**, 49, (8), 1952-1962.

References

56. Keseru, G. M.; Makara, G. M., The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discov.* **2009**, 8, (3), 203-212.
57. Oprea, T.; Allu, T.; Fara, D.; Rad, R.; Ostopovici, L.; Bologna, C., Lead-like, drug-like or "pub-like": how different are they? *Journal of Computer-Aided Molecular Design* **2007**, 21, (1), 113-119.
58. Oprea, T. I.; Davis, A. M.; Teague, S. J.; Leeson, P. D., Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comp. Sci.* **2001**, 41, (5), 1308-1315.
59. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliver. Rev.* **1997**, 23, (1-3), 3-25.
60. Navia, M. A.; Chaturvedi, P. R., Design principles for orally bioavailable drugs. *Drug Discover. Today* **1996**, 1, (5), 179-189.
61. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H., A 'rule of three' for fragment-based lead discovery? *Drug Discov. Today* **2003**, 8, (19), 876-877.
62. Phillips, V. Molecular twist compounds: A novel strategy for structure-selective nucleic acid targeting. University of Bradford, Bradford, 2003.
63. Wheelhouse, R. T.; Jennings, S. A.; Phillips, V. A.; Pletsas, D.; Murphy, P. M.; Garbett, N. C.; Chaires, J. B.; Jenkins, T. C., Design, synthesis, and evaluation of novel biarylpyrimidines: A new class of ligand for unusual nucleic acid structures. *J. Med. Chem.* **2006**, 49, (17), 5187-5198.

References

64. Choudry, G. A.; Hamilton Stewart, P. A.; Double, J. A.; Krul, M. R. L.; Naylor, B.; Flannigan, G. M.; Shah, T. K.; Brown, J. E.; Phillips, R. M., A novel strategy for NQO1 (NAD(P)H:quinone oxidoreductase, EC 1.6.99.2) mediated therapy of bladder cancer based on the pharmacological properties of EO9. *Br. J. Cancer* **2001**, 85, (8), 1137-1146.
65. Rong, D. Structure-activity relationships of a series of aryl heterocycles as anti-tumour drugs University of Bradford, Bradford, 2009.
66. Menna, M., Antitumor potential of natural products from mediterranean ascidians. *Phytochem. Rev.* **2009**, 8, (2), 461-472.
67. Gonzalez-Santiago, L.; Suarez, Y.; Zarich, N.; Munoz-Alonso, M. J.; Cuadrado, A.; Martinez, T.; Goya, L.; Iradi, A.; Saez-Tormo, G.; Maier, J. V.; Moorthy, A.; Cato, A. C. B.; Rojas, J. M.; Munoz, A., Aplidin induces JNK-dependent apoptosis in human breast cancer cells via alteration of glutathione homeostasis, Rac1 GTPase activation, and MKP-1 phosphatase downregulation. *Cell Death Differ.* **2006**, 13, (11), 1968-1981.
68. Brandon, E. F. A.; Sparidans, R. W.; van Ooijen, R. D.; Meijerman, I.; Lazaro, L. L.; Manzanares, I.; Beijnen, J. H.; Schellens, J. H. M., In vitro characterization of the human biotransformation pathways of aplidine, a novel marine anti-cancer drug. *Invest. New Drugs* **2007**, 25, (1), 9-19.
69. Urdiales, J.; Morata, P.; De Castro, I. N.; Sánchez-Jiménez, F., Antiproliferative effect of dehydrodidemnin B (DDB), a depsipeptide isolated from Mediterranean tunicates. *Cancer Lett.* **1996**, 102, (1-2), 31-37.

70. Nagle, D.; Zhou, Y.-D., Marine natural products as inhibitors of hypoxic signaling in tumors. *Phytochem. Rev.* **2009**, 8, (2), 415-429.
71. Cuadrado, A.; Garcia-Fernandez, L. F.; Gonzalez, L.; Suarez, Y.; Losada, A.; Alcaide, V.; Martinez, T.; Fernandez-Sousa, J. M. I. a.; Sanchez-Puelles, J. M. I. a.; Munoz, A., Aplidin Induces apoptosis in human cancer cells via glutathione depletion and sustained activation of the epidermal growth factor receptor, Src, JNK, and p38 MAPK. *J. Biol. Chem.* **2003**, 278, (1), 241-250.
72. Erba, E.; Serafini, M.; Gaipa, G.; Tognon, G.; Marchini, S.; Celli, N.; Rotilio, D.; Brogгинi, M.; Jimeno, J.; Faircloth, G. T.; Biondi, A.; D'Incalci, M., Effect of Aplidin in acute lymphoblastic leukaemia cells. *Br. J. Cancer* **2003**, 89, (4), 763-773.
73. Garcia-Fernandez, L. F.; Losada, A.; Alcaide, V.; Alvarez, A. M.; Cuadrado, A.; Gonzalez, L.; Nakayama, K.; Nakayama, K. I.; Fernandez-Sousa, J. M.; Munoz, A.; Sanchez-Puelles, J. M., Aplidin induces the mitochondrial apoptotic pathway via oxidative stress-mediated JNK and p38 activation and protein kinase C delta. *Oncogene* **2002**, 21, (49), 7533-7544.
74. Munoz-Alonso, M. J.; Gonzalez-Santiago, L.; Zarich, N.; Martinez, T.; Alvarez, E.; Rojas, J. M.; Munoz, A., Plitidepsin has a dual effect inhibiting cell cycle and inducing apoptosis via Rac1/c-Jun NH₂-terminal kinase activation in human melanoma cells. *J. Pharmacol. Exp. Ther.* **2008**, 324, (3), 1093-1101.
75. Brogгинi, M.; Marchini, S. V.; Galliera, E.; Borsotti, P.; Taraboletti, G.; Erba, E.; Sironi, M.; Jimeno, J.; Faircloth, G. T.; Giavazzi, R.; D'Incalci,

- M., Aplidine, a new anticancer agent of marine origin, inhibits vascular endothelial growth factor (VEGF) secretion and blocks VEGF-VEGFR-1 (flt-1) autocrine loop in human leukemia cells MOLT-4. *Leukemia* **2003**, 17, (1), 52-59.
76. Straight, A.; Oakley, K.; Moores, R.; Bauer, A.; Patel, A.; Tuttle, R.; Jimeno, J.; Francis, G., Aplidin reduces growth of anaplastic thyroid cancer xenografts and the expression of several angiogenic genes. *Cancer Chemoth. Pharm.* **2006**, 57, (1), 7-14.
77. Hart, C. P., Finding the target after screening the phenotype. *Drug Discov. Today* **2005**, 10, (7), 513-519.
78. Page, M. J.; Amess, B.; Rohlff, C.; Stubberfield, C.; Parekh, R., Proteomics: a major new technology for the drug discovery process. *Drug Discov. Today* **1999**, 4, (2), 55-62.
79. Hood, B. L.; Veenstra, T. D.; Conrads, T. P., Mass spectrometry-based proteomics. *Int. Cong. Ser.* **2004**, 1266, 375-380.
80. Jeffery, D. A.; Bogyo, M., Chemical proteomics and its application to drug discovery. *Curr. Opin. Biotech.* **2003**, 14, (1), 87-95.
81. Yates, J. R.; Ruse, C. I.; Nakorchevsky, A., Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* **2009**, 11, (1), 49-79.
82. Parekh, R. B.; Rohlff, C., Post-translational modification of proteins and the discovery of new medicine. *Curr. Opin. Biotech.* **1997**, 8, (6), 718-723.
83. Barrett, J.; Brophy, P. M.; Hamilton, J. V., Analysing proteomic data. *Int. J. Parasitol.* **2005**, 35, (5), 543-553.

References

84. Patterson, S. D.; Aebersold, R. H., Proteomics: the first decade and beyond. *Nat. Genet.* **2003**, 33, 311-323.
85. Chen, G.; Pramanik, B. N., Application of LC/MS to proteomics studies: current status and future prospects. *Drug Discov. Today* **2009**, 14, (9-10), 465-471.
86. Kruse, U.; Bantscheff, M.; Drewes, G.; Hopf, C., Chemical and pathway proteomics: Powerful tools for oncology drug discovery and personalized health care. *Mol. Cell. Proteomics* **2008**, 7, (10), 1887-1901.
87. Mocellin, S.; Riccardo Rossi, C.; Traldi, P.; Nitti, D.; Lise, M., Molecular oncology in the post-genomic era: the challenge of proteomics. *Trends Mol. Med.* **2004**, 10, (1), 24-32.
88. Liang, X.; Bai, J.; Liu, Y.-H.; Lubman, D. M., Characterization of SDS-PAGE-separated proteins by matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* **1996**, 68, (6), 1012-1018.
89. O'Farrell, P. H., High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **1975**, 250, (10), 4007-4021.
90. Edman, P., Method for determination of the amino acid sequence in peptides. *Acta Chem. Scand.* **1950**, 4, 283-293.
91. Bradshaw, R. A.; Burlingame, A. L., From proteins to proteomics. *IUBMB Life* **2005**, 57, (4-5), 267-72.
92. El-Aneed, A.; Cohen, A.; Banoub, J., Mass spectrometry, Review of the basics: Electrospray, MALDI, and commonly used mass analyzers. *Appl. Spectrosc. Rev.* **2009**, 44, (3), 210-230.
93. Dempster, A. J., A new method of positive ray analysis. *Phys. Rev.* **1918**, 11, (4), 316.

References

94. Hoffmann, E. d.; Stroobant, V., *Mass spectrometry: Principles and applications*. 3rd ed.; Wiley: **2007**; p 502.
95. Schiller, J.; Süß, R.; Arnhold, J.; Fuchs, B.; Leßig, J.; Müller, M.; Petkovic, M.; Spalteholz, H.; Zschörnig, O.; Arnold, K., Matrix-assisted laser desorption and ionization time-of-flight (MALDI-TOF) mass spectrometry in lipid and phospholipid research. *Prog. Lipid Res.* **2004**, 43, (5), 449-488.
96. Ragoussis, J.; Elvidge, G. P.; Kaur, K.; Colella, S., Matrix-assisted laser desorption/ionisation, time-of-flight mass spectrometry in genomics research. *PLOS Genet.* **2006**, 2, (7), e100.
97. Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* **1988**, 60, (20), 2299-2301.
98. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Sp.* **1988**, 2, (8), 151-153.
99. Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F., Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int. J. Mass Spectro.* **1987**, 78, 53-68.
100. Gobom, J.; Schuerenberg, M.; Mueller, M.; Theiss, D.; Lehrach, H.; Nordhoff, E., α -Cyano-4-hydroxycinnamic acid affinity sample preparation. A protocol for MALDI-MS peptide analysis in proteomics. *Anal. Chem.* **2001**, 73, (3), 434-438.

101. Stephens, W., A pulsed mass spectrometer with time dispersion. *Phys. Rev.* **1946**, 69, 691.
102. Cotter, R. J., *Time-of-flight mass spectrometry: Instrumentation and applications in biological research*. Illustrated Edition ed.; American Chemical Society: Washington, DC, **1997**.
103. Mano, N.; Goto, J., Biomedical and biological mass spectrometry. *Anal. Sci.* **2003**, 19, (1), 3-14.
104. Cotter, R. J., The new time-of-flight mass spectrometry. *Anal. Chem.* **1999**, 71, (13), 445a-451a.
105. Mamyrin, B.; Karataev, V.; Shmikk, D.; Zagulin, V., The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Sov. Phys. JETP* **1973**, 37, 45.
106. Bruker, D. G., Ultraflex II operator manual. In **2004**.
107. Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A., A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* **2003**, 376, (7), 952-965.
108. Liska, A. J.; Shevchenko, A., Combining mass spectrometry with database interrogation strategies in proteomics. *Trends Anal. Chem.* **2003**, 22, (5), 291-298.
109. Henzel, W. J.; Watanabe, C.; Stults, J. T., Protein identification: the origins of peptide mass fingerprinting. *J. Am. Soc. Mass Spectr.* **2003**, 14, (9), 931-942.
110. Eng, J. K.; McCormack, A. L.; Yates Iii, J. R., An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectr.* **1994**, 5, (11), 976-989.

References

111. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature*. **2003**, 422, (6928), 198-207.
112. Damodaran, S.; Wood, T. D.; Nagarajan, P.; Rabin, R. A., Evaluating peptide mass fingerprinting-based protein identification. *Genomics, Proteomics & Bioinformatics* **2007**, 5, (3-4), 152-157.
113. Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T., Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Analytical Chemistry* **1987**, 59, (21), 2621-2625.
114. Johnson, R. S.; Martin, S. A.; Biemann, K., Collision-induced fragmentation of $(M + H)^+$ ions of peptides. Side chain specific sequence ions. *Int. J. Mass Spectro.* **1988**, 86, 137-154.
115. Mascott, www.matrixscience.com.
116. Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J., Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* **1993**, 3, (6), 327-332.
117. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, 20, (18), 3551-3567.
118. Rix, U.; Superti-Furga, G., Target profiling of small molecules by chemical proteomics. *Nat. Chem. Biol.* **2009**, 5, (9), 616-624.
119. Wissing, J.; Godl, K.; Brehmer, D.; Blencke, S.; Weber, M.; Habenberger, P.; Stein-Gerlach, M.; Missio, A.; Cotten, M.; Muller, S.; Daub, H., Chemical proteomic analysis reveals alternative modes of action for pyrido[2,3-d]pyrimidine kinase inhibitors. *Mol. Cell. Proteomics* **2004**, 3, (12), 1181-1193.

References

120. Brehmer, D.; Greff, Z.; Godl, K.; Blencke, S.; Kurtenbach, A.; Weber, M.; Muller, S.; Klebl, B.; Cotten, M.; Keri, G.; Wissing, J.; Daub, H., Cellular targets of gefitinib. *Cancer Res.* **2005**, 65, (2), 379-382.
121. Weber, P. C.; Ohlendorf, D. H.; Wendoloski, J. J.; Salemme, F. R., Structural origins of high-affinity biotin binding to streptavidin. *Science* **1989**, 243, (4887), 85-88.
122. Freitag, S.; Le Trong, I.; Klumb, L.; Stayton, P. S.; Stenkamp, R. E., Structural studies of the streptavidin binding loop. *Protein Sci.* **1997**, 6, (6), 1157-66.
123. Livnah, O.; Bayer, E. A.; Wilchek, M.; Sussman, J. L., Three-dimensional structures of avidin and the avidin-biotin complex. *Proc. Natl. Acad. Sci. USA* **1993**, 90, (11), 5076-80.
124. Hyre, D. E.; Le Trong, I.; Merritt, E. A.; Eccleston, J. F.; Green, N. M.; Stenkamp, R. E.; Stayton, P. S., Cooperative hydrogen bond interactions in the streptavidin-biotin system. *Protein Sci.* **2006**, 15, (3), 459-67.
125. Freitag, S.; Le Trong, I.; Chilkoti, A.; Klumb, L. A.; Stayton, P. S.; Stenkamp, R. E., Structural studies of binding site tryptophan mutants in the high-affinity streptavidin-biotin complex. *J. Mol. Biol.* **1998**, 279, (1), 211-221.
126. Wang, J. H.; Hewick, R. M., Proteomics in drug discovery. *Drug Discov. Today* **1999**, 4, (3), 129-133.
127. Murphy, P. M.; Phillips, V. A.; Jennings, S. A.; Garbett, N. C.; Chaires, J. B.; Jenkins, T. C.; Wheelhouse, R. T., Biarylpyrimidines: a new class of ligand for high-order DNA recognition. *Chem. Commun.* **2003**, (10), 1160-1.

References

128. Herbert, C. G.; Bass, R. G.; Watson, K. A.; Connell, J. W., Preparation of poly(arylene ether pyrimidine)s by aromatic nucleophilic substitution reactions. *Macromolecules* **1996**, 29, (24), 7709-7716.
129. Wilson, W. D.; Strekowski, L.; Tanious, F. A.; Watson, R. A.; Mokrosz, J. L.; Strekowska, A.; Webster, G. D.; Neidle, S., Binding of unfused aromatic cations to DNA. The influence of molecular twist on intercalation. *J. Am. Chem. Soc.* **1988**, 110, (25), 8292-8299.
130. Miyaura, N.; Suzuki, A., Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **1995**, 95, (7), 2457-2483.
131. Miyaura, N.; Yamada, K.; Suzuki, A., A new stereospecific cross-coupling by the palladium-catalyzed reaction of 1-alkenylboranes with 1-alkenyl or 1-alkynyl halides. *Tetrahedron Lett.* **1979**, 20, (36), 3437-3440.
132. Blouin, N.; Michaud, A.; Gendron, D.; Wakim, S.; Blair, E.; Neagu-Plesu, R.; Belletete, M.; Durocher, G.; Tao, Y.; Leclerc, M., Toward a rational design of Poly(2,7-Carbazole) derivatives for solar cells. *J. Am. Chem. Soc.* **2008**, 130, (2), 732-742.
133. Cheng, F.; Adronov, A., Suzuki coupling reactions for the surface functionalization of single-walled carbon nanotubes. *Chem. Mater.* **2006**, 18, (23), 5389-5391.
134. Sabat, M.; Johnson, C. R., Synthesis of unnatural amino acids via suzuki cross-coupling of enantiopure vinyloxazolidine derivatives. *Org. Lett.* **2000**, 2, (8), 1089-1092.
135. Neumann, H.; Brennfürher, A.; Beller, M., A general synthesis of diarylketones by means of a three-component cross-coupling of aryl and

- heteroaryl bromides, carbon monoxide, and boronic acids. *Chem.-Eur. J.* **2008**, 14, (12), 3645-3652.
136. Aliprantis, A. O.; Canary, J. W., Observation of Catalytic Intermediates in the Suzuki Reaction by Electrospray Mass-Spectrometry. *J. Am. Chem. Soc.* **1994**, 116, (15), 6985-6986.
137. Gniewek, A.; Ziolkowski, J. J.; Trzeciak, A. M.; Zawadzki, M.; Grabowska, H.; Wrzyszczyk, J., Palladium nanoparticles supported on alumina-based oxides as heterogeneous catalysts of the Suzuki-Miyaura reaction. *J. Catal.* **2008**, 254, (1), 121-130.
138. Beletskaya, I. P.; Kashin, A. N.; Khotina, I. A.; Khokhlov, A. R., Efficient and recyclable catalyst of palladium nanoparticles stabilized by polymer micelles soluble in water for suzuki-miyaura reaction, ostwald ripening process with palladium nanoparticles. *Synlett* **2008**, 2008, (10), 1547-1552.
139. Köhler, K.; Heidenreich, R. G.; Soomro, S. S.; Pröckl, S. S., Supported palladium catalysts for suzuki reactions: Structure-property relationships, optimized reaction protocol and control of palladium leaching. *Adv. Synth. Catal.* **2008**, 350, (18), 2930-2936.
140. Fairlamb, I. J. S., Palladium catalysis in synthesis: where next? *Tetrahedron* **2005**, 61, (41), 9661-9662.
141. Clayden, J., *Organic chemistry*. Oxford University Press: Oxford, **2001**; p 1512.
142. Tsuji, J., *Palladium reagents and catalysts: new perspectives for the 21st century*. 2 ed.; Wiley: **2004**; p 670.

References

143. Casado, A. L.; Espinet, P., On the configuration resulting from oxidative addition of RX to Pd(PPh₃)₄ and the mechanism of the cis-to-trans isomerization of [PdRX(PPh₃)₂] complexes (R equals aryl, X equals halide). *Organometallics* **1998**, 17, (5), 954-959.
144. Smith, G. B.; Dezeny, G. C.; Hughes, D. L.; King, A. O.; Verhoeven, T. R., Mechanistic studies of the Suzuki cross-coupling reaction. *J. Org. Chem.* **1994**, 59, (26), 8151-8156.
145. Matos, K.; Soderquist, J. A., Alkylboranes in the Suzuki-Miyaura coupling: stereochemical and mechanistic studies. *J. Org. Chem.* **1998**, 63, (3), 461-470.
146. Suzuki, A., Recent advances in the cross-coupling reactions of organoboron derivatives with organic electrophiles, 1995-1998. *J. Organomet. Chem.* **1999**, 576, (1-2), 147-168.
147. Heck, R. F.; Nolley, J. P., Palladium-catalyzed vinylic hydrogen substitution reactions with aryl, benzyl, and styryl halides. *J. Org. Chem.* **1972**, 37, (14), 2320-2322.
148. Milstein, D.; Stille, J. K., A general, selective, and facile method for ketone synthesis from acid chlorides and organotin compounds catalyzed by palladium. *J. Am. Chem. Soc.* **1978**, 100, (11), 3636-3638.
149. Hatanaka, Y.; Hiyama, T., Cross-coupling of organosilanes with organic halides mediated by a palladium catalyst and tris(diethylamino)sulfonium difluorotrimethylsilicate. *J. Org. Chem.* **1988**, 53, (4), 918-920.
150. Sonogashira, K.; Tohda, Y.; Hagihara, N., A convenient synthesis of acetylenes: catalytic substitutions of acetylenic hydrogen with

- bromoalkenes, iodoarenes and bromopyridines. *Tetrahedron Lett.* **1975**, 16, (50), 4467-4470.
151. King, A.; Okukado, N.; Negishi, E., Highly general stereo-, regio-, and chemo-selective synthesis of terminal and internal conjugated enynes by the Pd-catalysed reaction of alkynylzinc reagents with alkenyl halides. *Chem. Commun.* **1977**, (19), 683-684.
152. Newkome, G. R.; Garbis, S. J.; Majestic, V. K.; Fronczek, F. R.; Chiari, G., Chemistry of heterocyclic-compounds .61. Synthesis and conformational studies of macrocycles possessing 1,8-naphthyridino or 1,5-naphthyridino subunits connected by carbon-oxygen bridges. *J. Org. Chem.* **1981**, 46, (5), 833-839.
153. Montalbetti, C. A. G. N.; Falque, V., Amide bond formation and peptide coupling. *Tetrahedron* **2005**, 61, (46), 10827-10852.
154. Coste, J.; Frerot, E.; Jouin, P., Coupling N-methylated amino acids using PyBroP and PyCloP halogenophosphonium salts: mechanism and fields of application. *J. Org. Chem.* **1994**, 59, (9), 2437-2446.
155. Kurtan, T.; Nesnas, N.; Li, Y.-Q.; Huang, X.; Nakanishi, K.; Berova, N., Chiral recognition by CD-sensitive dimeric zinc porphyrin host. 1. Chiroptical protocol for absolute configurational assignments of monoalcohols and primary monoamines. *J. Am. Chem. Soc.* **2001**, 123, (25), 5962-5973.
156. Schomaker, J. M.; Delia, T. J., Arylation of halogenated pyrimidines via a Suzuki coupling reaction. *J. Org. Chem.* **2001**, 66, (21), 7125-7128.

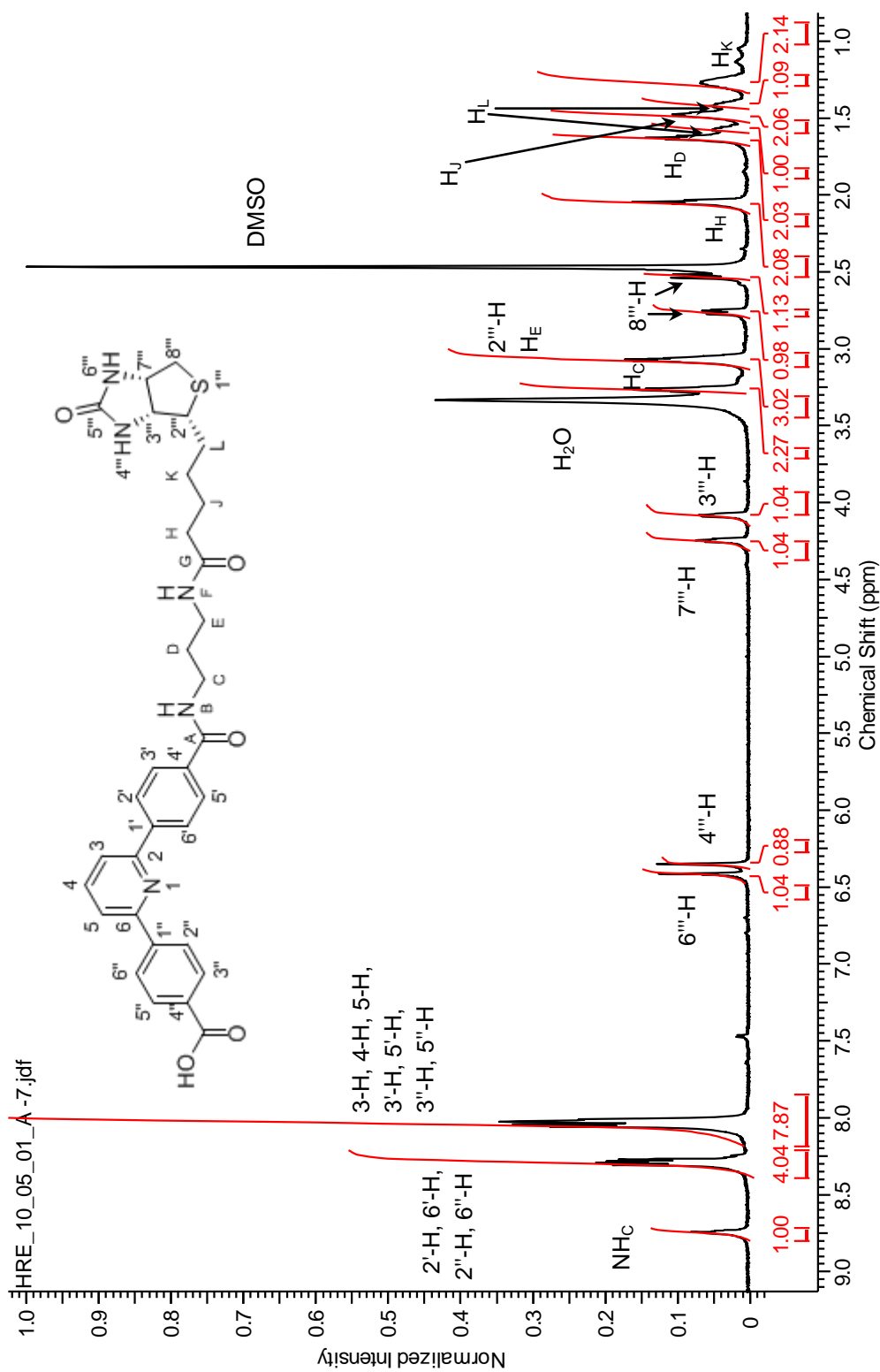
References

157. Delia, T. J.; Schomaker, J. M.; Kalinda, A. S., The synthesis of substituted phenylpyrimidines via Suzuki coupling reactions. *J. Heterocyclic Chem.* **2006**, 43, (1), 127-131.
158. Zhou, B. S.; Taylor, B.; Kornau, K., Synthesis of unsymmetrical 4,6-diarylpyrimidines. *Tetrahedron Lett.* **2005**, 46, (23), 3977-3979.
159. Gong, Y.; Pauls, H. W., A convenient synthesis of heteroaryl benzoic acids via Suzuki reaction. *Synlett* **2000**, (6), 829-831.
160. Zhang, H.; Cai, Q.; Ma, D., Amino acid promoted CuI-catalyzed C-N bond formation between aryl halides and amines or N-containing heterocycles. *J. Org. Chem.* **2005**, 70, (13), 5164-5173.
161. Cohen, T.; Wood, J.; Dietz, A. G., Organocopper intermediates in the exchange reaction of aryl halides with salts of copper(I). The possible role of copper(III). *Tetrahedron Lett.* **1974**, 15, (40), 3555-3558.
162. Paine, A. J., Mechanisms and models for copper mediated nucleophilic aromatic-substitution .2. A single catalytic species from 3 different oxidation-states of copper in an Ullmann synthesis of triaryl amines. *J. Am. Chem. Soc.* **1987**, 109, (5), 1496-1502.
163. Katayama, H.; Oda, Y., Chemical proteomics for drug discovery based on compound-immobilized affinity chromatography. *J. Chromatogr. B* **2007**, 855, (1), 21-27.
164. Ford, L. P.; Wright, W. E.; Shay, J. W., A model for heterogeneous nuclear ribonucleoproteins in telomere and telomerase regulation. *Oncogene* **2002**, 21, (4), 580-583.
165. Carpenter, B.; MacKay, C.; Alnabulsi, A.; MacKay, M.; Telfer, C.; Melvin, W. T.; Murray, G. I., The roles of heterogeneous nuclear

- ribonucleoproteins in tumour development and progression. *BBA - Rev. Cancer* **2006**, 1765, (2), 85-100.
166. He, Y.; Brown, M. A.; Rothnagel, J. A.; Saunders, N. A.; Smith, R., Roles of heterogeneous nuclear ribonucleoproteins A and B in cell proliferation. *J. Cell. Sci.* **2005**, 118, (14), 3173-3183.
167. STRING www.string-db.org.
168. Kozu, T.; Henrich, B.; Schäfer, K. P., Structure and expression of the gene (HNRPA2B1) encoding the human hnRNP protein A2/B1. *Genomics* **1995**, 25, (2), 365-371.
169. Lorimer, A. V.; O'Connor, P. D.; Brimble, M. A., Buchwald-Hartwig mono-N-arylation with 2,6-dihaloisonicotinic acid derivatives: a convenient desymmetrization method. *Synthesis* **2008**, 2008 (17), 2764-2770.
170. Still, W. C.; Kahn, M.; Mitra, A., Rapid chromatographic technique for preparative separations with moderate resolution. *J. Org. Chem.* **1978**, 43, (14), 2923-2925.
171. Salama, I.; Hocke, C.; Utz, W.; Prante, O.; Boeckler, F.; Hubner, H.; Kuwert, T.; Gmeiner, P., Structure-selectivity investigations of D-2-like receptor ligands by CoMFA and CoMSIA guiding the discovery of D-3 selective PET radioligands. *J. Med. Chem.* **2007**, 50, (3), 489-500.

APPENDIX

APPENDIX 1: NMR DATA FOR COMPOUND 92

Figure 1. ¹H NMR spectrum for compound 92

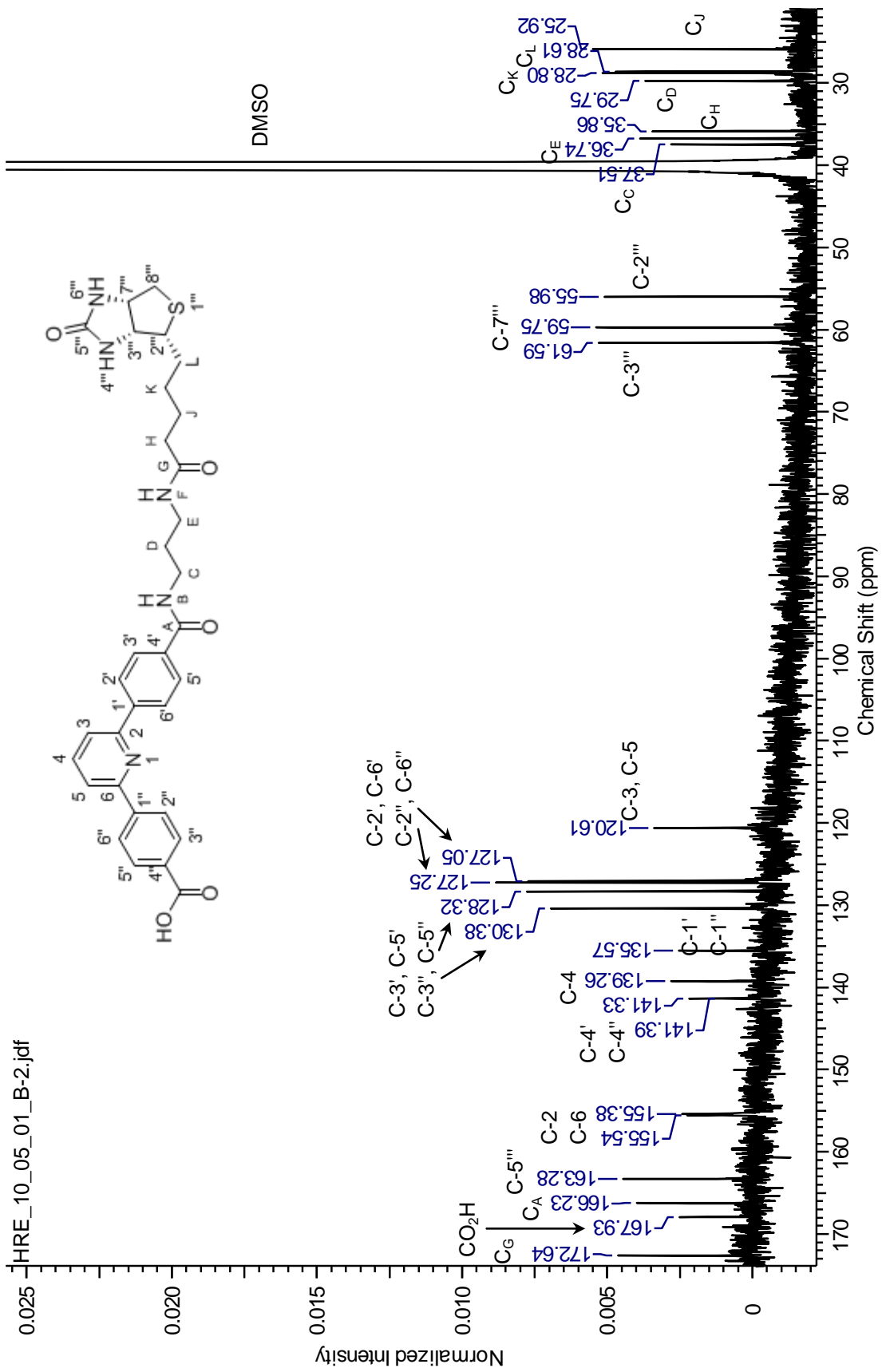
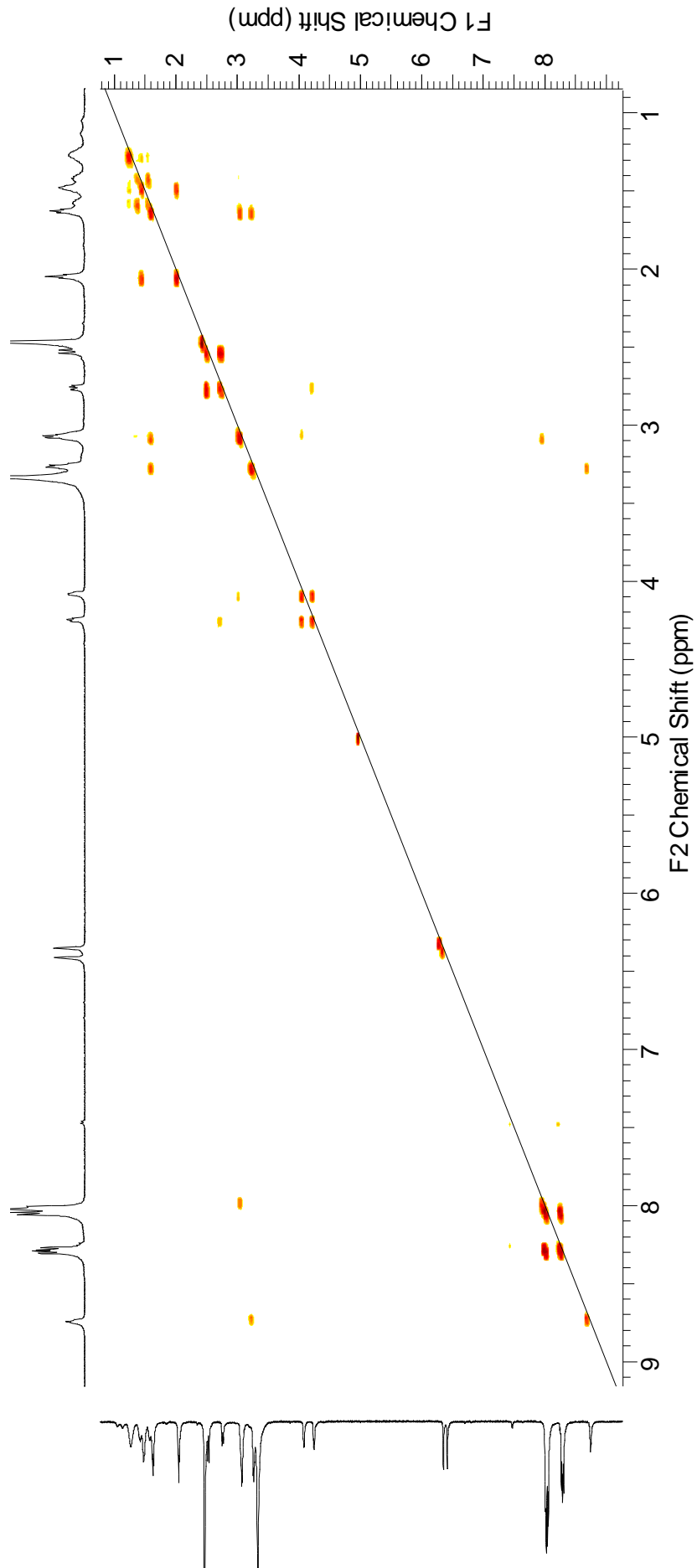


Figure 2. ^{13}C NMR spectrum for compound **92**



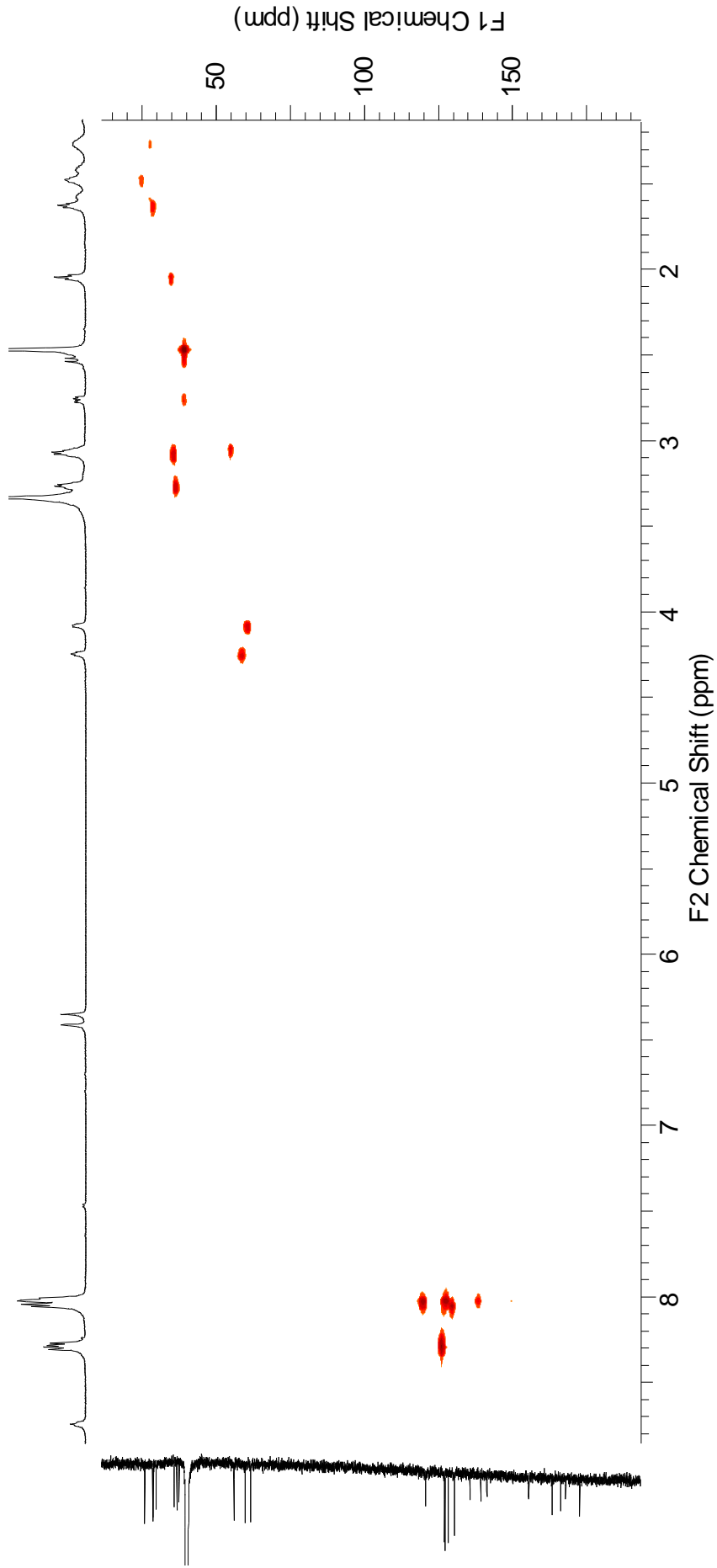


Figure 4. ^1H - ^{13}C HSQC NMR spectrum for compound **92**

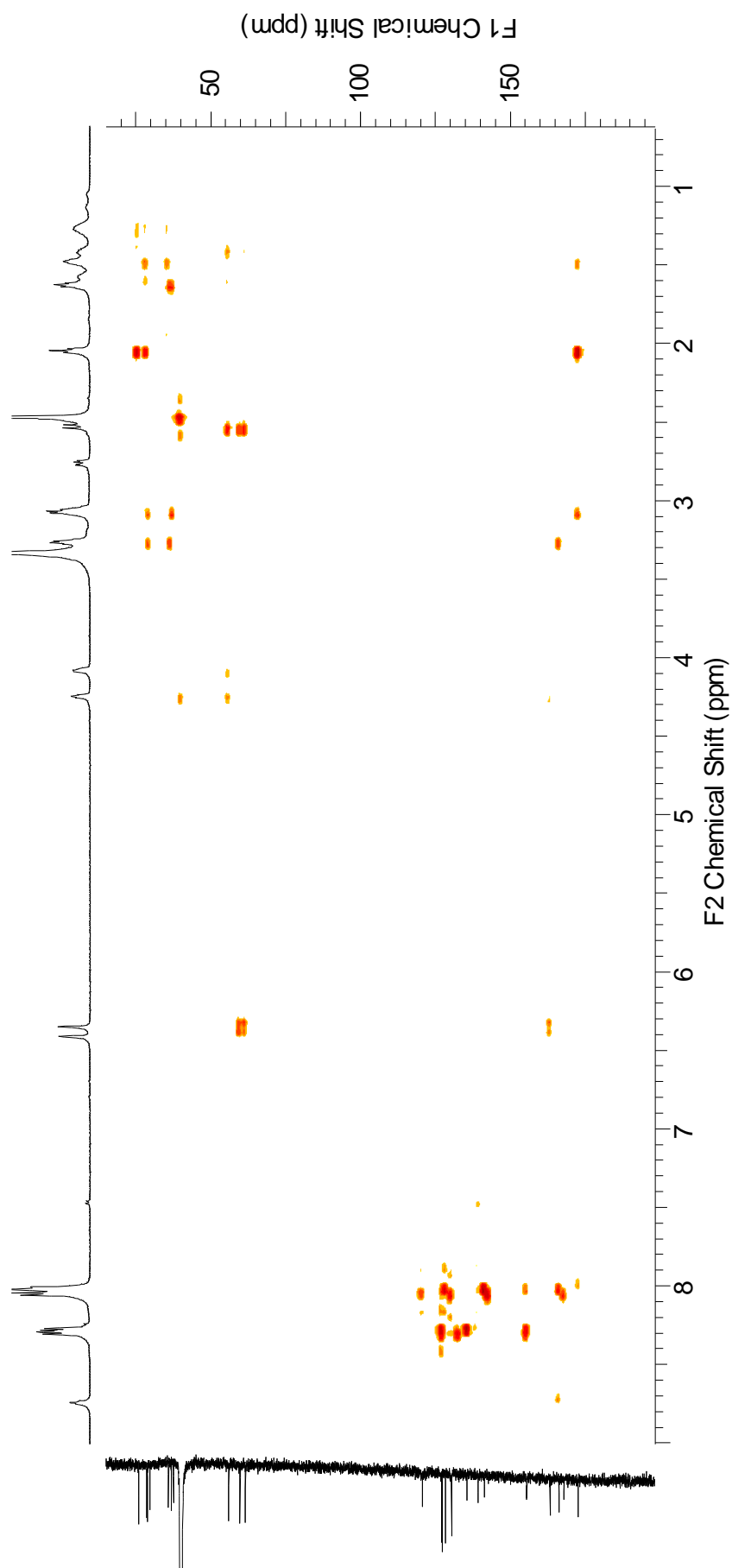


Figure 5. ^1H - ^{13}C HMBBC NMR spectrum for compound **92**

APPENDIX 2: MASS SPECTROMETRY DATA FOR PEPTIDES IDENTIFIED

AS hnRNP A2/B1

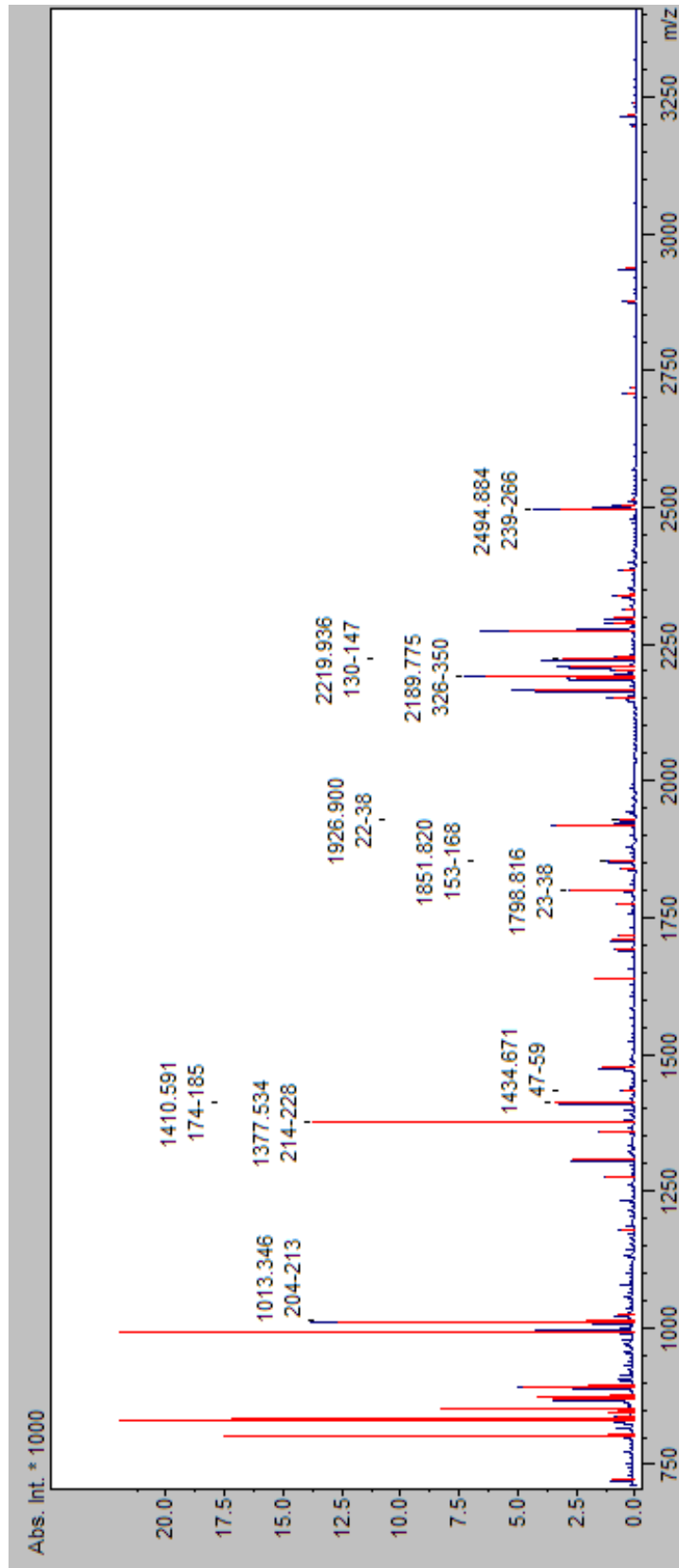


Figure 1 MS PMF of hnRNP A2/B1, showing matched peptide peaks.

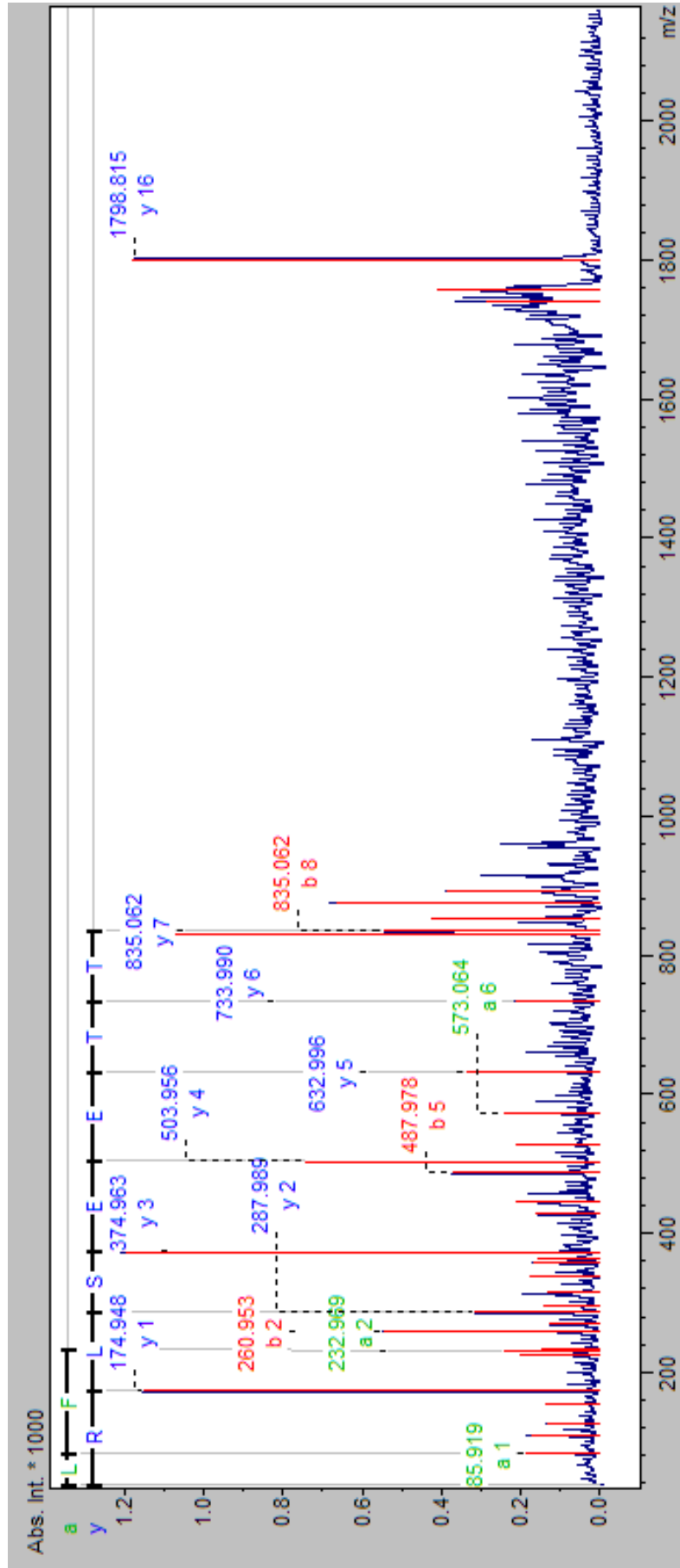


Figure 2. MS/MS Spectrum of peptide m/z 1798.8, identified as hnRNPA2/B1.

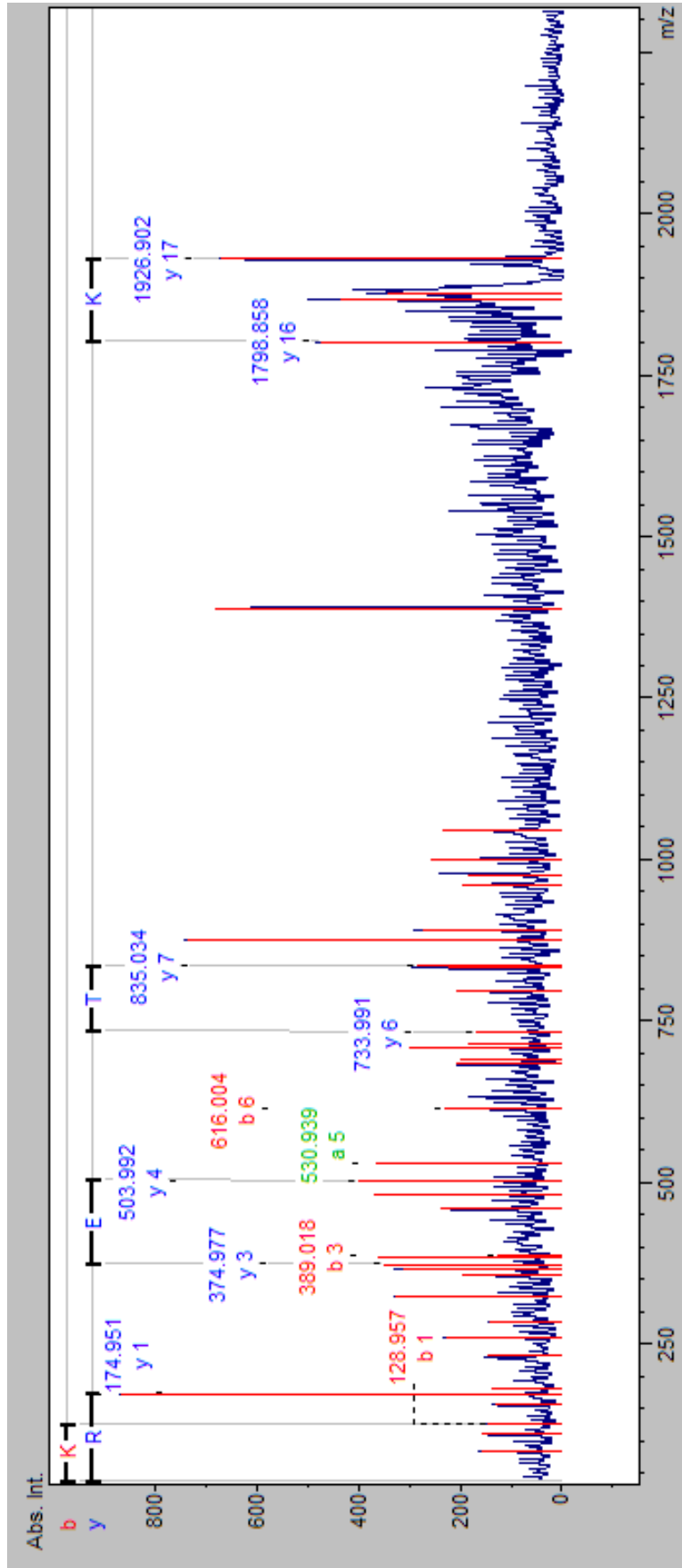


Figure 3. MS/MS Spectrum of peptide m/z 1926.9, identified as hnRNPA2/B1.

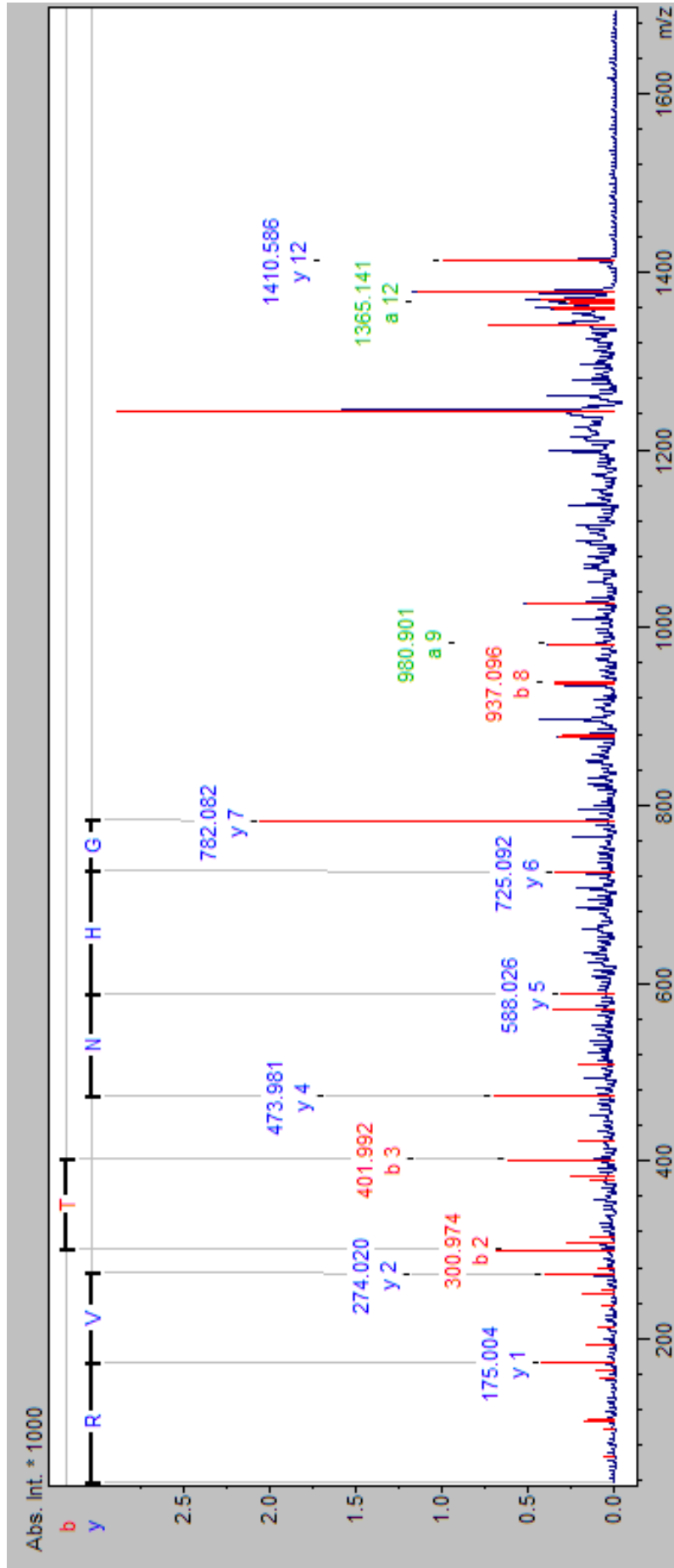


Figure 4. MS/MS Spectrum of peptide m/z 1410.5, identified as hnRNPA2/B1.

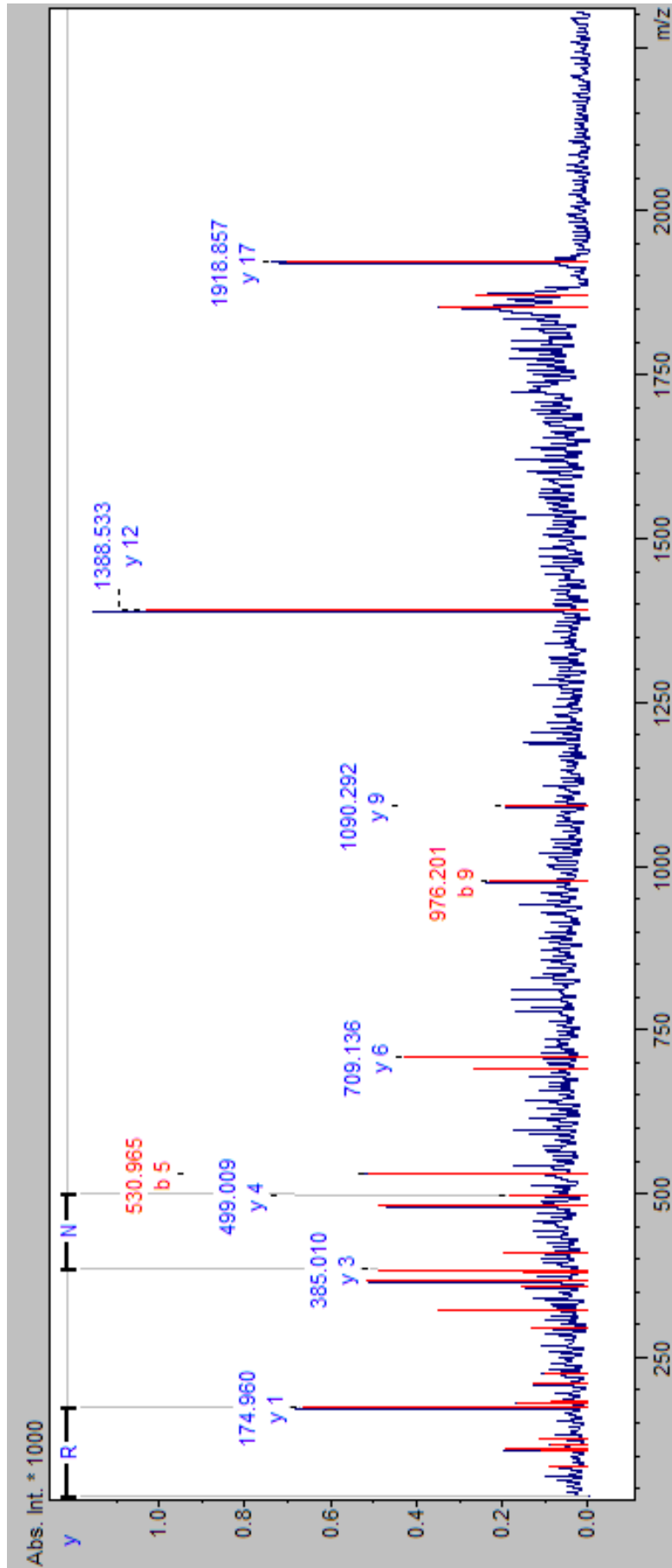


Figure 5. MS/MS Spectrum of peptide m/z 1918.8, identified as hnRNPA2/B1.

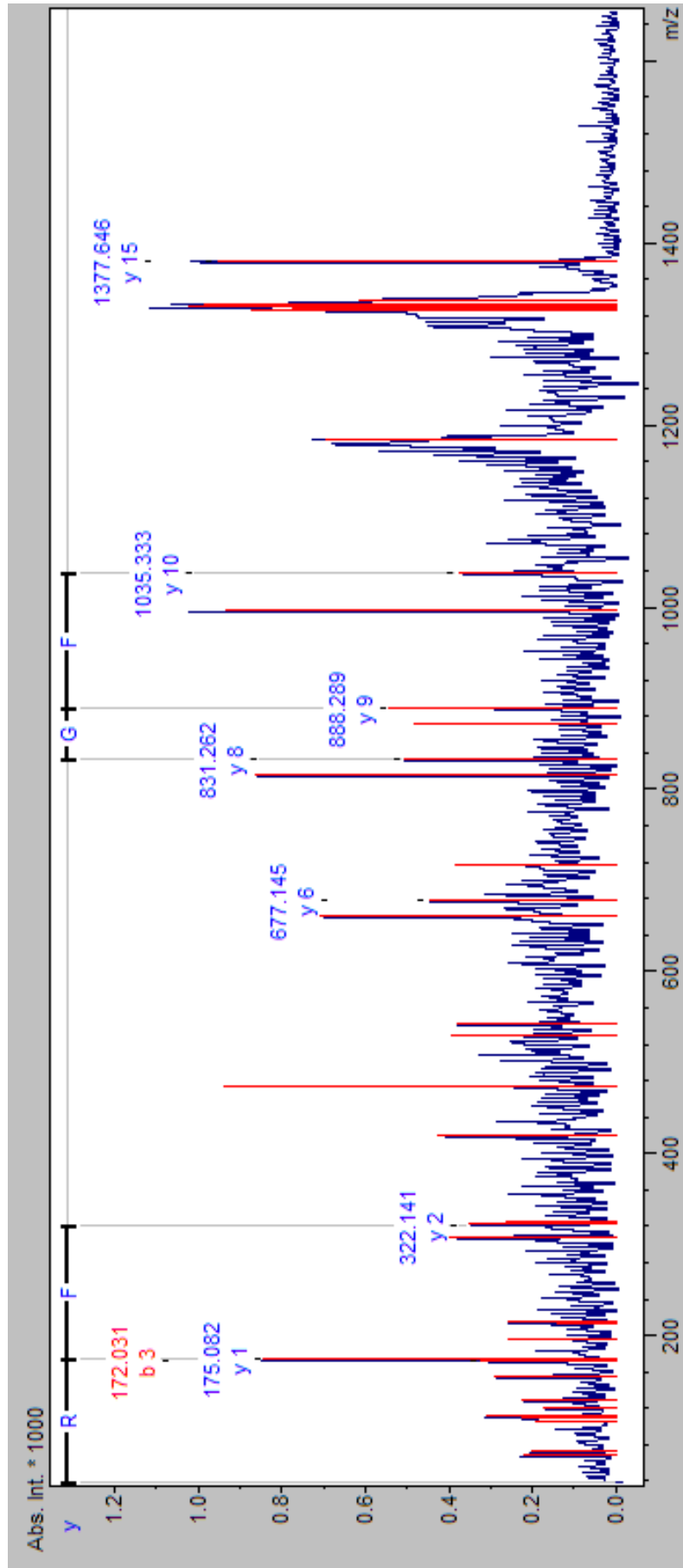


Figure 6. MS/MS Spectrum of peptide m/z 1377.6, identified as hnRNPA2/B1.

APPENDIX 3: PUBLICATIONS

Novel small molecule inhibitors of telomerase. A. Adekunle, A., H.R. Evans, V.A. Phillips, D. Pletsas, R.T. Wheelhouse, S.M. Parkin, D.A.L. Watt, D.T.S. Sharpe, R.M. Phillips. 20th EORTC-NCI-AACR Symposium on Molecular Targets and Cancer Therapeutics, October 2008 Geneva, Switzerland. *Eur. J. Cancer* (Supl) **6** (2008) 139.

Investigating a novel antitumor molecular target using a chemical proteomic strategy. H.R. Evans, D.W. Rong, R.M. Phillips, C.W. Sutton and R.T. Wheelhouse. 21st EORTC-NCI-AACR Symposium on Molecular Targets and Cancer Therapeutics, November 2009, Boston, Massachusetts. *Mol. Cancer Ther.* (Supl) **8** (2009) A204.

Targeting DNA•RNA hybrids. R.T. Wheelhouse, H.R. Evans, N.C. Garbett and J.B. Chaires. BACR 50th Anniversary Celebration meeting 'Hallmarks of Cancer: from Mechanisms to Therapies' June 2010, Glasgow.

Investigating a novel antitumor molecular target using a chemical proteomic strategy. H.R. Evans, D.W. Rong, R.M. Phillips, C.W. Sutton and R.T. Wheelhouse. Yorkshire Cancer Research Annual Scientific Meeting June 2010, Harrogate.

Targeting DNA•RNA hybrids. R.T. Wheelhouse, H.R. Evans, N.C. Garbett and J.B. Chaires. RSC Nucleic Acids Forum, July 2010, Liverpool