

University of St Andrews



Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

Synthetic Studies Towards Target Identification in Parasitology



Thesis presented for the degree of Doctorate in Philosophy

Rachel Elizabeth Morgan

January 2006

School of Chemistry

University of St Andrews



TH
F141

Declarations

I, Rachel Elizabeth Morgan, hereby certify that this thesis, which is approximately 60,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date: 20/1/2006 Signature of candidate:

I was admitted as a research student in September, 2002 and as a candidate for the degree of Doctor of Philosophy in January 2006; the higher study for which this is a record was carried out at the University of St Andrews between 2002 and 2006.

Date: 20/1/2006 Signature of candidate:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 20/1/2006 Signature of supervisor:

(Dr. N. J. Westwood)

Copyright Declaration

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date: 20/1/2006 Signature of candidate:

Abstract

Toxoplasma gondii is an obligate intracellular parasite which is thought to infect over a quarter of the world's population. This parasite causes the potentially fatal disease, toxoplasmosis. Invasion is an essential part of the parasite's life cycle. Despite this, the mechanistic details of the process of invasion are far from being completely understood.

Chemical genetics is the use of compounds to modify a biological system in order to enhance our understanding at a molecular level. This approach facilitates the study of essential components of that system. *Forward* chemical genetics begins by identifying compounds from a high throughput screen and culminates in target identification studies.

A high throughput screen against *T. gondii* invasion identified several small molecules which both inhibit and enhance invasion. This thesis presents experimental work towards the identification of the cellular target of a compound that was identified in this screen and which enhances *T. gondii* invasion.

Confirmation of the identity of the original compound is described. Derivatives of the compound have been prepared and structure activity relationship data obtained. This information guided the selection of compounds to be synthesised for the purposes of target identification and studies towards compounds with increased potency.

The parallel synthesis of a small library of analogues gave an insight into both the physical and biological properties of these types of compounds. Compounds were successfully identified with improved physical and biological properties.

Affinity chromatography resins containing a derivative of the original hit were synthesised. Development of suitable controls provided a valuable set of reagents with which target identification could be achieved. Promising initial results from the use of these reagents are also described.

Acknowledgements

Thank you to.....

Nick Westwood for the opportunity to work for him, for his support and enthusiasm over the course of my PhD studies, for encouraging and inspiring me to continue with research and for introducing me to the cool subjects of parasitology and chemical genetics.

The Westwood research group members past and present for all their support and help: Cristina Lucas-Lopez, Russell Pearson, David Taddei, Neil McIntyre, Richard Riggs, Nick Voute, Nicolas Isambert and Helmut Kraus. In particular to: Stephen Patterson, Jon Hollick, Kathryn Evans, Federica Catti and Alan Jones for reading this thesis.

Gary Ward for the opportunity to work for him in Vermont and for all his support and encouragement.

The Ward research group for making me feel welcome during my time in their lab and in particular Jeralyn Haraldsen for assay data and Kim Carey for sharing her time and expertise.

The following people for all their help with analytical data:

Catherine Botting (mass spectrometry, LCMS and QSTAR MS/MS), Tomas Lebl (NMR), Melanja Smith (NMR), Caroline Horsburgh (mass spectrometry), Alex Slawin (X-ray crystallography), Sylvia Williamson (elemental analysis).

Prof. I. Samuels for the use of his fluorimeter and Mark Goosens for his help with this.

The NIH, EPSRC and the Russell Trust for financial support.

Mum, Dad and Paul for all their support, it would not have been possible without you.

My good friends, Matthew, Denzil, Sarah, Teleri, Claire, Kirsty, Tessa, Kelcey, Juliet, Kathryn, Darren, Laura, for helping me to remain slightly sane over the past three years and for reminding me how lucky I am.

Standard Abbreviations and Acronyms

Å	Angstrom
ACD	available chemical database
ADF	actin depolymerising factor
Arg	arginine (amino acid)
Asp	aspartic acid (amino acid)
ADDP	azodicarboxylic dipiperidine
br	broad (spectral)
<i>n</i> -Bu	normal (primary) butyl
°C	degrees Celsius
CDCl ₃	deuterated chloroform
CI	chemical ionisation
CDK	cyclin-dependent kinase
COSY	proton-proton correlation spectroscopy
Cys	cysteine (amino acid)
CSD	Cambridge structural database
δ	chemical shift in parts per million downfield from tetramethylsilane
d	doublet
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DEAD	diethyl azodicarboxylate
DHFR	dihydrofolate reductase
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethylsulfoxide
DMSO- <i>d</i> ₆	deuterated dimethylsulfoxide
EDC	1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride
EI	electron impact
ε	extinction coefficient
eq	equivalents
<i>et al.</i>	<i>et alia</i> , latin, and others

ES+	electrospray ionisation, operating in positive mode
ES-	electrospray ionisation, operating in negative mode
FCG	forward chemical genetics
fM	femtomolar
GPI	glycosylphosphatidyl inositol
GR	glutathione reductase
GSH	glutathione
g	gram(s)
His	histidine (amino acid)
HMBC	Heteronuclear Multiple Bond Connectivity
hrs	hour(s)
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
HSQC	Heteronuclear Single Quantum Coherence
HTS	high throughput screen
Hz	Hertz
ID	identification
IMC	inner membrane complex
IR	infrared
<i>J</i>	coupling constant (in NMR spectrometry)
kDa	kiloDaltons (mass)
λ_{\max}	absorbance maximum
LAC	lowest active concentration
LCMS	liquid chromatography-mass spectrometry
lit.	literature
m	multiplet (spectral); metre(s); milli
M	molar (moles per litre); mega
M ⁺	parent molecular ion
MALDI	matrix-assisted laser desorption ionisation
MHz	megaHertz
min	minute(s); minimum
mM	millimolar (millimoles per litre)
mol	mole(s)
mmol	millimole(s)

mp	melting point
MS	mass spectrometry
μM	micromolar (micromoles per litre)
m/z	mass-to-charge ratio
NMR	nuclear magnetic resonance
NTPase	nucleoside triphosphate hydrolase
Nu	nucleophile
PEG	polyethyleneglycol
Ph	phenyl
Phe	phenylalanine
ppm	part(s) per million
PV	parasitophorous vacuole
q	quartet (spectral)
RCG	reverse chemical genetics
RNAi	RNA interference
rt	room temperature
s	singlet (spectral); second(s)
t	triplet (spectral)
TFA	trifluoroacetic acid
THF	tetrahydrofuran
Thr	threonine (amino acid)
TLC	thin-layer chromatography
TR	trypanothione reductase
Tr	trityl
Trp	tryptophan (amino acid)
Tyr	tyrosine (amino acid)
UV	ultraviolet
v/v	volume per unit volume
v	wavenumber, cm^{-1}
w/w	weight per unit weight (weight-to-weight ratio)
YFP	yellow fluorescent protein

Contents

Abstract.....	I
Acknowledgements.....	II
Standard Abbreviations and Acronyms.....	IV
Contents.....	VII
Chapter 1: Introduction.....	1
1.1 Chemical genetics.....	1
1.1.1 Forward chemical genetics.....	1
1.1.2 Reverse chemical genetics.....	3
1.1.3 The advantages of using chemical genetics.....	4
1.2 The chemistry inherent in chemical genetics.....	5
1.2.1 Synthesis of novel compound libraries.....	5
1.2.2 Synthesis and further chemical modifications of compounds identified through the screening process.....	6
1.2.2.1 Resynthesis of the hit compound.....	6
1.2.2.2 Modification of the hit compound.....	7
1.3 Target Identification Studies.....	8
1.3.1 Chemistry based methods of target ID.....	8
1.3.1.1 Affinity chromatography.....	8
1.3.1.2 Radiolabelling.....	9
1.3.1.3 Tagging.....	10
1.3.1.4 Photoaffinity labelling.....	11
1.3.1.5 Yeast-3-Hybrid.....	12
1.4 Protozoan parasites.....	14
1.4.1 Apicomplexan parasites.....	14
1.5 <i>Toxoplasma gondii</i>	15
1.5.1 <i>T. gondii</i> life cycle.....	17
1.5.2 Invasion.....	19
1.5.2.1 Motility.....	21
1.5.2.2 Surface recognition.....	23
1.5.2.3 Conoid extension.....	24
1.5.2.4 Secretion.....	24
1.5.2.4.1 Proteases involved in invasion.....	25
1.5.2.5 Summary.....	26
1.6 Forward chemical genetics approach to studying <i>T. gondii</i> invasion.....	26
1.6.1 Overall project goals.....	26

1.7 High Throughput screening in parasitology.....	27
1.7.1 Screening against purified proteins.....	28
1.7.2 In silico screening.....	29
1.8 High throughput screen against <i>T. gondii</i> invasion.....	30
1.8.1 Secondary assays.....	34
1.8.1.1 Motility.....	35
1.8.1.2 Secretion.....	36
1.8.1.3 Conoid extension.....	37
1.8.2 The activity of the compounds against other apicomplexans.....	37
1.8.3 Selection of the compound to be studied.....	38
1.8.4 Literature search on the selected compound 44	39
1.8.4.1 Synthetic precedent.....	39
1.8.4.2 Known biological properties.....	40
1.8.5 Observation of the same phenotype in the literature.....	41
1.8.6 Project aims.....	41
Chapter 2: Synthesis of active compound and derivatives.....	43
2.1 Synthesis of the bioactive compound 44 identified from the high throughput screen...	43
2.1.1 Investigating the solubility of 44	48
2.1.2 The fluorescence properties of 44	51
2.2 Preliminary structure activity relationship data.....	52
2.2.1 Changing the aromatic group.....	54
2.2.2 N-alkylation.....	56
2.2.2.1 N-methylation.....	56
2.2.2.1.1 A route to an <i>N</i> -methylated derivative of 44 via 63	57
2.2.2.1.2 Direct methylation of compound 44	59
2.2.2.2 <i>N</i> -butenylation.....	61
2.2.3 Reduction of the exocyclic double bond in 44	63
2.2.4 Extension of CH ₃	64
2.2.5 Replacement of the pyridone oxygen for a sulphur.....	67
2.2.8 SAR Conclusions.....	67
2.3 Additional chemical reactivity associated with structures of type 44	68
2.3.1 Reaction of the lactone moiety with amines.....	68
2.3.2 Reaction of lactone in 44 with alkoxides.....	69
2.4 Chapter Summary.....	71
Chapter 3: Parallel synthesis of derivatives.....	72
3.1 Library synthesis.....	72
3.2 Available technology.....	72

3.2.1	Parallel Synthesis.....	73
3.2.2	Microwaves.....	74
3.2.3	Solid phase synthesis.....	74
3.2.4	Automated purification.....	75
3.2.5	High Throughput compound analysis.....	76
3.3	Selection of the reaction for library synthesis.....	77
3.4	Adaptation of the reaction protocol for use in parallel synthesis.....	78
3.5	The selected aldehydes.....	79
3.6	Success of the parallel synthesis.....	80
3.7	Reactions which were unsuccessful in the parallel synthesis format.....	86
3.8	Determination of solubility.....	89
3.9	Biological activities.....	91
3.9.1	Benzaldehyde derivatives.....	92
3.9.2	5-membered heterocyclic rings.....	94
3.9.3	Fused aromatic rings.....	95
3.9.4	Summary.....	96
3.10	Fluorescent properties.....	97
3.11	Summary.....	102
Chapter 4: Developing reagents for affinity chromatography.....		103
4.1	Target ID in protozoan parasites.....	103
4.1.1	Affinity chromatography.....	103
4.1.2	Radiolabelling.....	105
4.1.3	Photoaffinity labelling.....	106
4.2	Designing a reagent for affinity chromatography.....	107
4.2.1	Immobilisation of a derivative of 44 onto an affinity resin.....	107
4.2.2	Controls for affinity chromatography experiments.....	108
4.2.2.1	Competitive enhancers/inhibitors.....	109
4.2.2.2	Control resins.....	110
4.2.3	Protein extracts.....	110
4.3	Synthesis of an affinity chromatography reagent.....	111
4.3.1	Synthesis of an analogue of 44 containing a linker unit.....	112
4.3.1.1	Selection of a linker system and point of attachment.....	112
4.3.1.2	Synthesis of a compound containing a short linker unit.....	113
4.3.1.2.1	Light transformation.....	114
4.3.2	Loading of compounds onto affinity resins.....	118
4.3.2.1	Affinity resins.....	118
4.3.2.2	The loading level of the affinity resin.....	119
4.3.2.2.1	Significance of the loading level.....	119

4.3.2.2.2	Determining the loading level.....	119
4.3.2.2.3	Loading 177 onto an affinity resin.....	120
4.3.2.3	Blocking of unreacted sites.....	121
4.3.2.3.1	NMR experiment to determine the potential success of blocking with ethanolamine.....	122
4.3.2.3.2	Attempts to analyse the resin.....	123
4.4	Initial affinity chromatography experiments performed in the laboratory of Prof. G. E. Ward at the University of Vermont.....	124
4.5	Developments for 2 nd generation affinity chromatography experiments.....	129
4.5.1	Attempts to reduce non-specific binding.....	129
4.5.1.1	Removing the non-specifically bound proteins from the protein lysate.....	129
4.5.1.2	The use of compounds with differing distances from the resin.....	129
4.5.1.2.1	Synthesis of a longer linker system.....	130
4.5.1.2.2	Different internal resin linker lengths.....	131
4.5.2	Improvement of experimental controls.....	132
4.5.2.1	Competitive controls.....	132
4.5.2.2	Control resins.....	134
4.6	Chapter summary.....	135
 Chapter 5: Affinity chromatography and covalent modification.....		137
5.1	Covalent modification.....	137
5.1.1	Known covalent modifiers.....	137
5.1.2	Target identification for covalent modifiers.....	138
5.1.2.1	Tagging.....	138
5.1.2.3	Affinity chromatography.....	139
5.1.3	Identifying the protein binding partners.....	140
5.2	Evidence for the covalent reaction of 44 with its protein target(s).....	140
5.3	NMR investigations into the mechanism of action of 44	146
5.3.1	The use of IR to verify the outcome of the NMR investigations.....	150
5.4	Covalent affinity chromatography experiments.....	152
5.5	Inflammatory profilin as a possible target.....	155
5.5.1	Mammalian profilin.....	155
5.5.1.1	Profilin actin-binding activity.....	156
5.5.1.2	Other binding proteins of profilin.....	157
5.5.1.3	The crystal structure of human profilin.....	157
5.5.1.3.1	The actin binding site.....	158
5.5.1.3.2	Proline-rich binding site.....	159
5.5.1.3.3	Other conserved sites.....	160
5.5.1.3.4	Summary from examining the crystal structure.....	160

5.52	Summary.....	161
5.6	Chapter summary.....	161
Chapter 6:	Experimental.....	163
6.1	General.....	163
6.2	Synthetic procedures and Analysis.....	165
6.2.63	General procedure for determining solubility.....	211
6.2.64	General Resin loading procedure.....	212
6.2.65	NMR experiments to determine a suitable blocking procedure.....	212
6.2.66	General procedure for the NMR competition experiments.....	213
6.3	Biological Procedures.....	214
6.3.1	Assay against <i>T. gondii</i> invasion.....	214
6.3.1.1	Cell Lines and Parasites.....	214
6.3.1.2	High-throughput invasion assay.....	214
6.3.2	Affinity chromatography experiments using <i>T. gondii</i> parasites and HFF cells.....	215
6.3.3.1	<i>T. gondii</i> lysate.....	215
6.3.3.2	HFF cell lysate.....	216
6.3.3.3	General affinity chromatography procedure.....	216
6.3.3.4	Affinity chromatography procedure for covalent modification.....	216
Chapter 7:	Conclusions and Future work.....	218
7.1	Conclusions.....	218
7.2	Future work.....	219
7.2.1	Validation of the cellular target.....	219
7.2.2	Completing chemical investigations.....	220
References	222
Appendices	234

Chapter 1: Introduction

1.1 Chemical Genetics

Chemical genetics is the name given to a research technique which uses small molecules (or compounds) to modify the function of biological systems in order to advance the level of understanding associated with each system.^{1,2,3,4,5,6,7,9,10} These compounds are usually selected through the use of high throughput screening (HTS), a method that is used to assess large compound collections. There are two types of chemical genetics approaches: forward and reverse. Forward chemical genetics (FCG) involves a screen that has as its readout a *phenotype* of interest with the aim of linking a compound that induces the desired phenotype with the compound's cellular target (usually a protein). FCG often implicates novel cellular components in the selected phenotype. Reverse chemical genetics (RCG) involves screening against a *protein* of interest. Compounds that alter function of the protein can then be used to study the role of that protein within the biological system.

1.1.1 Forward chemical genetics

FCG has the potential to implicate new cellular components (e.g. proteins) in a particular biological process or discover new functions of known components. This approach is therefore capable of providing novel and potentially valuable information about cellular mechanisms. Identification of a compound that alters the function of a new protein can also provide exciting opportunities in drug discovery. The majority of research in drug discovery remains directed towards the “rational design” of compounds against known targets; FCG complements this approach. As there are only a limited number of targets that are compatible with the so called “rational approach”, the discovery of ligands for new targets using FCG is important if progress is to be made in therapeutic development. The increasing occurrence of resistance seen against existing drugs particularly in the area of parasitology also contributes to the importance of identifying compounds that act on new targets.^{11,12,13,14,15} The strength

of FCG lies is its “unbiased nature” as many potential targets are screened at once. This important advantage is also where the most significant challenge lies – the need to identify the cellular target of the compound. The methods available for such investigations are highlighted in Section 1.3.

The identification of monastrol (**1**) is an example of a successful FCG study. Mayer *et al.* developed a screen in order to identify compounds which affect mitosis.¹⁶ Mitosis is the process of nuclear division in eukaryotic cells. The screen involved whole-cell immunodetection to identify compounds which increased the phosphorylation of the protein nucleolin. Nucleolin is known to be specifically phosphorylated in cells entering mitosis. Therefore compounds that cause cells to arrest their cell cycle in mitosis would be expected to display increased levels of phosphorylated nucleolin in a defined number of cells (compared with control). From a library of 16,320 compounds, 139 were identified that caused cells to arrest in mitosis. Secondary assays were used to eliminate those compounds that were affecting tubulin polymerisation dynamics, affecting the cells at interphase and disrupting actin filaments. From the secondary assays one compound of particular interest was identified, monastrol (**1**).

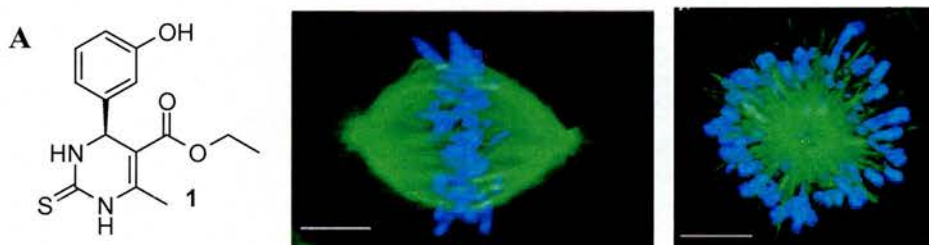


Figure 1. Monastrol induced alteration of the cellular spindle in mitotic cells. **A** The chemical structure of (*S*)-monastrol (**1**). **B** Normal bipolar spindle. **C** Monoastral spindles caused by the addition of **1**. Image reproduced from T. Mayer *et al.*, *Science*, **1999**, *286*, 971-974.¹⁶

Monastrol (**1**) was observed to induce a monoastral array of the spindle at mitosis (Fig. 1). Treatment with monastrol (**1**) over 4 hours resulted in the majority of cells being arrested at mitosis and 90% of arrested cells displaying the monoastral array of microtubules. This particular phenotype has been observed before when the mitotic kinesin Eg5 was disrupted by the addition of Eg5-specific antibodies.¹⁷ Further investigations led Mayer *et al.* to conclude that the cellular target of monastrol

(1) was Eg5.¹⁶ Monastrol (1) has been used in further research, in fact a search of the literature reveals that the original paper concerning the discovery of monastrol (1) has been cited over 279 times since the paper was published.

1.1.2 Reverse chemical genetics

RCG, the screening of compounds against a purified protein of interest is a strategy which is widely employed, not only in academia but also in the pharmaceutical industry. This technique enables the identification of new lead compounds or core structures around which novel drug types can potentially be developed, making this approach appealing to the pharmaceutical industry.

Proteins can have multiple binding partners, the range of which can cause them to affect multiple phenotypes within a singular organism. RCG therefore provides chemical tools to help identify at best, all the possible functions of a given protein. Understanding the multiple functions of a given protein is important especially in drug discovery. Inhibiting a particular protein and disrupting one of its functions may have the desired effect, but it may also result in unwanted side effects due to the inhibition of a secondary function of the protein.

Blebbistatin is an example of a successful RCG screen.¹⁸ Straight *et al.* were interested in determining the role of non-muscle myosin II in cytokinesis. They screened a library of over 16,000 compounds using an assay based on observing the myosin-dependent hydrolysis of ATP through the use of luciferase and luciferin (Fig. 2). An inhibitor of non-muscle myosin II was identified in a well containing the purchased compound **2**. **2** was repurchased and reassayed in order to confirm that it was active in the assay. However, on reassaying the compound no activity was observed. **2** was dissolved in DMSO and left in the light in the presence of air and after a time was reassayed. This time the DMSO solution contained the active component. This active component of the solution was subsequently identified as blebbistatin (**4**). Later studies showed that the active enantiomer of blebbistatin is in fact (*S*)-(-)-blebbistatin (**3**).¹⁹

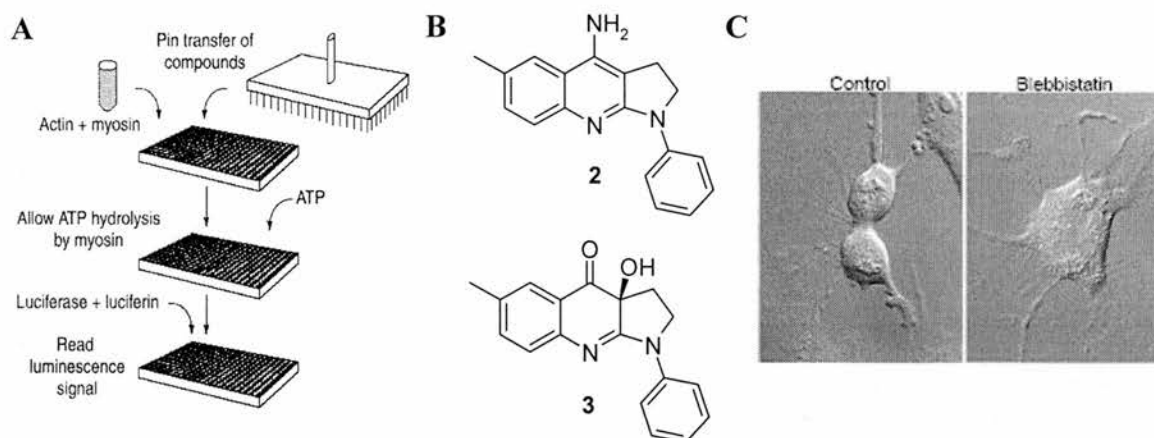


Figure 2. Identification of (*S*)-(-)-blebbistatin (**3**) as a non-muscle myosin II inhibitor and its effects on cytokinesis. **A** The assay used to identify (*S*)-(-)-blebbistatin (**3**). Actin and myosin are placed in 1536 well plates. Compounds are transferred to the plates as DMSO stock solutions. ATP is added to the wells. If myosin function is unaffected by a compound myosin will ‘walk’ along the actin filaments, hydrolysing ATP in the process. Inhibited myosin will not hydrolyse ATP. Luciferase and luciferin are then added to the wells. Any remaining ATP will be used by the luciferase which processes luciferin leading to the observed luminescence. Wells which luminesce are identified as containing compounds which inhibit myosin. **B** Compound **2** was the purchased compound a stored DMSO solution of which scored in the assay. (*S*)-(-)-blebbistatin (**3**), a decomposition product of **2**, has been identified as the active component. **C** Addition of (\pm) blebbistatin (**4**) to cells inhibits cleavage furrow contraction during cytokinesis. Image reproduced from A. F. Straight *et al.*, *Science*, **2003**, 299, 1743-1747.¹⁸

The addition of (\pm) blebbistatin (**4**) to cells was shown to inhibit cleavage furrow contraction and hence cytokinesis (Fig. 2), thus demonstrating the role of non-muscle myosin II in this process.¹⁸ This ability to inhibit furrow contraction was used to enable the study of how cytokinesis is spatially controlled.

(*S*)-(-)-Blebbistatin (**3**) is an excellent example of a valuable tool which can be identified through RCG approaches. Since the publication of the original article over 104 publications have cited the original article published on blebbistatin, indicating the extent of use of this compound by the research community.²⁰

1.1.3 The advantages of using chemical genetics

The increasing number of review articles written on the subject of chemical genetics is an illustration of the importance of this technique and its application to

many different areas of biology.^{1,2,3,4,5,6,7,9,10} Chemical genetics has certain advantages over alternative methods such as classical genetics and RNAi approaches and these are briefly discussed below.

Many of the advantages stem from the ability to add compound to a system at any time point and to any number of cellular environments. Limitations can arise with the classical genetic techniques when wishing to study essential genes. As a compound can be added to the biological system after the species has been cultured and can be removed rapidly (as most of the compounds identified act reversibly) essential genes can be readily studied. Chemical genetics can also be used in other situations where it is difficult to use the classical genetics approach. Creating multiple knockouts is difficult, especially if one or both of the genes is important (if the presence of one is made essential by the absence of the other).²¹ Compounds can be added to species containing existing knockouts or used in combination with other compounds to facilitate such experiments.¹⁸ In addition, the compounds can be used in a variety of environmental set ups from whole cell assays with differing cell lines, in crude protein extracts or with purified protein samples. Compounds also present a potentially closer link to drug discovery as some of the functionality required for activity has already been identified.

1.2 The chemistry inherent in chemical genetics

There are two main areas in which a chemist can impact on a chemical genetics project: i) synthesis of novel compound libraries to be screened and ii) synthesis of and further chemical modifications and/or optimisation of compounds identified during the screening stage.

1.2.1 Synthesis of novel compound libraries

Selection of a suitable compound library is necessary in order to achieve a good hit rate. The larger the diversity of the compound collection, the greater the chance of achieving the identification of novel bioactive compound classes. In the case of FCG the compounds should be diverse enough to identify different compounds with different cellular targets.²²

There are three different approaches which can be taken to obtain a compound collection for screening: 1) purchasing a compound library, 2) collating of compounds already synthesised in-house to create a library, 3) synthesis of a diverse library.

The synthesis of a library is a method by which chemists can significantly impact on chemical genetic projects.⁶ While this is clearly important to the research field of chemical genetics this thesis deals with the results from a screen using a commercial library

Whilst there are many commercially available compound libraries, they are all expensive. Various compound collections are available that range from compounds selected purely on the basis of diverse structure, to those with drug like properties, selected in accordance to Lipinski's rule of 5.^{23,24,25} There are also collections available which consist of known drugs, this allows for the identification of hit compounds which already have chemical properties suitable for drug development and for which the protein target is known. The size of commercial libraries varies from 10,000 compounds through to much larger libraries of 200,000. Larger libraries cost less per compound but generally the average price is £1 per 1mg of compound.⁶ Purchasing compound libraries is therefore an investment, resulting in institutions collectively buying a collection as it can be used in multiple assays.

1.2.2 Synthesis and further chemical modifications of compounds identified through the screening process.

Identifying an active compound using high-throughput screening is only the beginning of a chemical genetics project. Depending on if the screen is a forward or reverse screen there are differences in the manner in which the project continues. What is true of both instances is that further synthetic work is required. This often includes: 1) resynthesis of the hit compound, 2) synthesis of derivatives of the hit compound.

1.2.2.1 Resynthesis of the hit compound

Confirmation of the identity of the active component is important, it has already been seen that the compound purchased may not be the active compound (e.g.

(S)-(-)-blebbistatin (**3**), Section 1.1.2). Therefore resynthesis of the hit and biological evaluation allows the identity of the active component to be confirmed. In addition, there are often stereochemical questions relating to the hit compound, (especially if a purchased library was used), resynthesis of the compound (and its stereoisomers) allows the use of analytical techniques such as NMR and X-ray crystallography to answer such questions. The physical properties of the compound are also of interest as questions of solubility or cell permeability will also affect the biological activity of the compound and its derivatives. With a large supply of the hit compounds in hand, these can be easily investigated.

A further advantage of these studies is that once the synthetic route to the hit compound has been established, methods of introducing structural modifications can be assessed more easily.

1.2.2.2 Modification of the hit compound

Synthesis of analogues of the compound has two purposes: i) optimisation of the physical and biological properties of the compound and ii) identification of suitable sites to incorporate structural modifications for the purposes of target identification (ID) (of importance in *forward* chemical genetics).

Optimisation of biological properties is often required as the hit compound may not be sufficiently potent, selective, or may lack the desired physical properties (such as solubility). Importantly, the technology now available for the synthesis of focused collections of derivatives facilitates this goal (see section 3.2).

Various target ID methods require the modification of the hit compound. This is often in the form of attaching a linker unit (see section 1.3). The positioning of the linker is important as it must not interfere significantly with the activity of the compound. Therefore, initial studies are required to identify possible positions for attachment before the synthesis of a suitable reagent can occur (see section 2.2).

Selection of a suitable method of target ID is required in order to direct such synthesis. There are a number of methods which require the use of chemistry and these are outlined in Section 1.3.1.

1.3 Target Identification Studies

There are various methods by which potential cellular target(s) of a compound can be identified. It is important to note that many of these methods only identify the protein binding partners of a compound and follow up validation work must always be carried out. The available methods can be split into two main categories: i) those which do not require the use of any synthetic chemistry and ii) those which do.

Methods which do not require the use of synthetic chemistry include: generation of mutants that are resistant to the compound followed by complementation cloning,²⁶ transcriptional profiling,²³ protein micro-arrays^{28,29,30} and phage displays.³¹ These techniques have been very successfully employed.^{32,33} However, the focus in this section is on the methods which require chemical input.

1.3.1 Chemistry based methods of target ID

A number of target ID techniques which require the use of chemistry are described below. These methods require the synthesis of analogues either with a linker unit for attachment to a solid support or tag, or an analogue containing a label, (e.g. biotin, a radiolabel).

1.3.1.1 Affinity chromatography

Affinity chromatography takes advantage of the interaction between the compound and protein in order to 'pull out' potential target protein(s) from a crude or prefractionated cell lysate (Fig. 3). The compound is immobilised onto an affinity resin (usually agarose based so as to be water-compatible). Crude protein extract is then incubated with the resin. The target protein(s) binds to the compound on the resin and other proteins are washed off the resin. Elution of the target bound resin, with a solution able to disrupt the stronger interactions between compound and target(s), yields a purified sample of the target(s) which can then be identified using techniques such as mass spectrometry and validated in a cellular context.

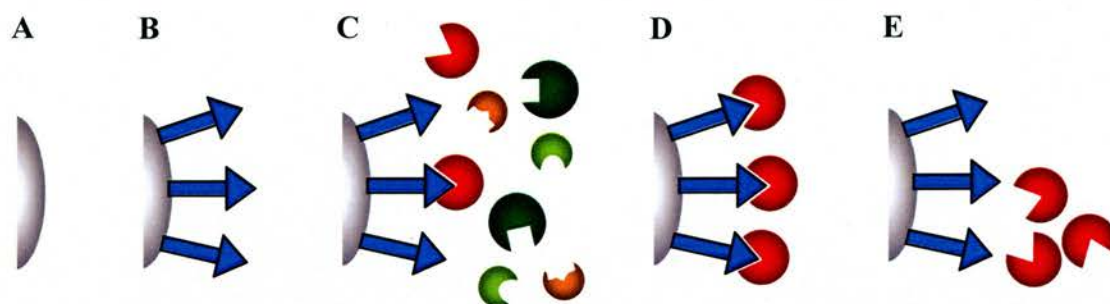


Figure 3. Affinity chromatography. **A** Commercially available water compatible resin, usually with an activating group to promote the binding of ligands to the resin. **B** The compound which has been modified to contain a linker system is loaded onto the resin. **C** Resin loaded with the compound is incubated with crude protein extract. The target protein(s) binds to the compound on the resin. **D** Unbound protein is washed off the resin. This leaves only the strongly binding target protein(s) attached to the resin. **E** The target protein(s) is washed off the resin under stronger elution conditions, to yield the isolated target protein(s).

— = linker system, ▲ = compound.

Affinity chromatography has been successfully applied in many different biological systems (see section 4.1.1 for a specific example).^{34,35,36,37,38,39,40,41,42,43} However, it is not without problems, the most significant of which is probably non-specific binding of proteins to the resin. This requires a balance to be reached with regards to the elution conditions at the varying stages, the goal being to insure selective elution of *all* the non-specific binding proteins in the first elution stages without the removal of the target protein(s). Controls are also used to highlight proteins that are non-specifically bound (see section 4.1 for full discussion on requirements for affinity chromatography and possible limitations/problems).

1.3.1.2 Radiolabelling

Radiolabelling provides a method by which the location on a chromatogram or 1- or 2-D gel of the compound can be followed. The bioactive compound is labelled using a radioactive isotope. This can be achieved, for example, through synthesis of an aryl bromide derivative followed by substitution of the bromine atom for tritium (^3H).⁴⁴ Alternatively it is possible to synthesise the compound from starting materials that contain radioactive atoms such as: ^{14}C , ^3H , ^{125}I , $^{32/33}\text{P}$ and ^{35}S . However the

radioactive atom should be added as late as possible in the synthesis due to cost and safety issues.

Once a radioactive analogue has been synthesised several options are available for identifying its protein target(s). These methods include separating the target protein(s) from a crude protein mixture with the use of gels or chromatography and using the radioactivity of the compound to highlight protein fractions of interest (see section 4.1.2 for an example of the successful use of this approach).⁴⁵ Bands/spots/fractions containing potential target protein(s) will be radiolabeled and can therefore be identified, initiating validation studies.^{45,46} Alternatively, the crude proteins can be displayed on either a non-denaturing gel or undergo renaturation after running the gel.⁴⁷ The radiolabeled compound can then be washed over the gel providing an opportunity for the compound to bind to its target. Visualising the gel will then highlight the band(s) corresponding to binding partner(s) (technique akin to drug westerns⁴⁸).⁴⁶

1.3.1.3 Tagging

The techniques used in conjunction with radiolabeled compounds can also be used with other labelling reagents. Compounds which have been modified to contain a linker unit ending in a reactive functional group can be used to link the compound to a variety of tags. Possible tags include those that contain a radiolabel (e.g. **5**) (useful if incorporating a radioactive atom into the compound itself is not possible); are fluorescent (e.g. Bodipy (**6**)) or biotin (**7**) (see Fig. 4).

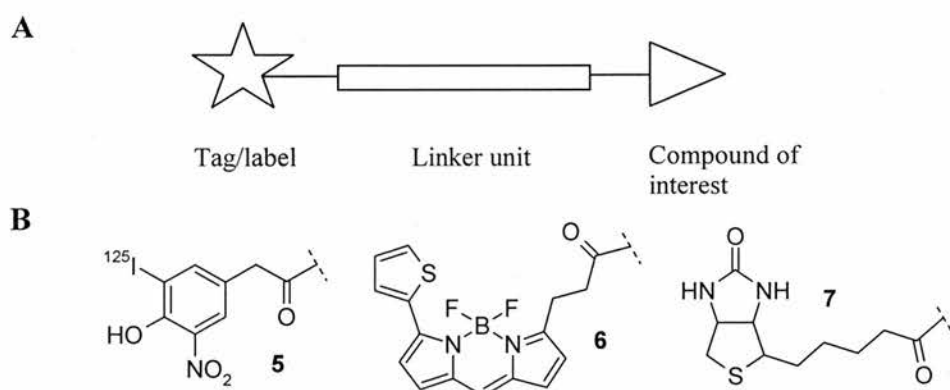


Figure 4. Tagging or labeling the compound of interest. **A** Schematic showing the reagent produced by labelling the compound of interest. **B** The structures of tags which can be attached to the compound

of interest: biotin (7), the fluorescent compound Bodipy (6) and radiolabelled 1-(4-hydroxy-3-iodo-5-nitrophenyl)acetone (5).

These tagged compounds can be used as above (section 1.3.2.2), with corresponding alterations in the detection method. For example, the use of fluorescence microscopy will highlight the presence of a Bodipy (6)-tagged compound.

Biotin (7) has a strong affinity to avidin (or streptavidin). This high affinity has been used to great advantage.⁴⁹ A crude protein extract is incubated with biotinylated compound and then run out on a gel. Fluorescently labelled avidin can then be added to the gel, where upon it will bind to the biotin (7) and highlight the positions of the protein(s) of interest.⁵⁰ Agarose beads loaded with avidin can also be purchased. Affinity chromatography based experiments can then be achieved using biotinylated analogues bound onto the resin through the biotin-avidin complex.

1.3.1.4 Photoaffinity labelling

Photoaffinity labelling integrates the use of labelling the compound with the use of a controllable covalent interaction between the compounds and potential protein targets. Covalent interactions which can be initiated at any desired time point will not be disrupted by harsh washing or identification techniques such as mass spectrometry (section 5.1.3). A further advantage of photoaffinity labelling is that it also has the potential to provide information on the binding site of the compound.

Photoaffinity labelling requires that the hit compound be modified to contain both a photoactivatable group (e.g. aryl azide, diazirine) and a means of detection (e.g. radiolabel, biotin). The modified compound is incubated with a crude protein extract, where upon the compound binds to potential target protein(s) (Fig. 5). The mixture is irradiated with light at a wavelength suitable to cause activation of the photoactivatable group. This causes a covalent bond to be formed between the protein and the compound. Isolation of the protein-compound complex is achieved through separation guided by the label. Mass spectrometric analysis leads to identification of proteins (labelled peptide sequences and even amino acid residues to which the compound is bound can be identified, through fragmentation techniques in conjunction with mass spectra analysis (Fig. 5)).

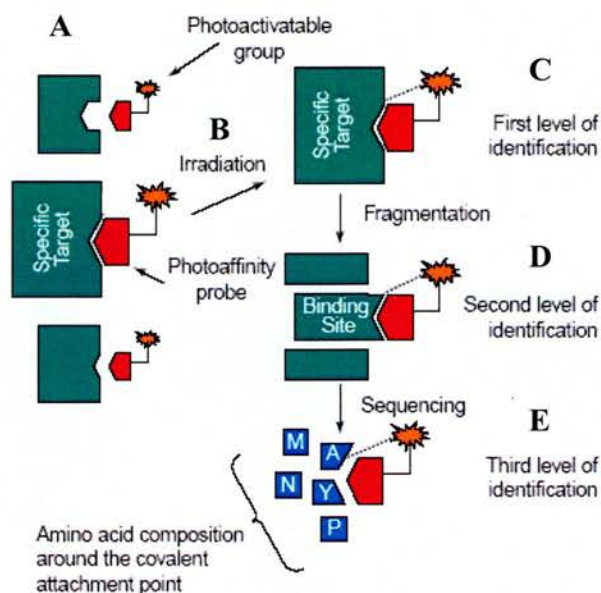


Figure 5. Schematic representation of the photoaffinity labelling approach to target ID. **A** The compound containing a photoactivatable group is incubated with a crude protein extract and the compound binds to its target(s). **B** The compound-protein mixture is irradiated with light at a set wavelength. This causes the photoactivatable group to form a covalent bond to the protein. **C** The compound-protein complex is isolated and identified (using mass spectrometry). **D** The protein is broken down into short peptide sequences. Comparison with the peptide sequences expected for the native protein identifies a region of the protein where the compound is bound. **E** The labelled peptide can then be broken down further into the individual amino acids providing information as to the specific amino acid residue labelled by the photoactivatable group. Image adapted from G. Dormán *et al.*, *Trends Biotech.*, **2000**, *18*, 64-77.⁵¹

There are various criteria which have to be met for these experiments to be successful, the most important of which is probably that the life time of the activated group is shorter than the time taken for the compound-protein complex to dissociate, but long enough to allow for reaction to occur with the protein.⁵¹ Despite many potential problems, there are a number of cases where this technique has been used successfully (see section 4.1.3 for a detailed example).^{51,52,53,54,55,56}

1.3.1.5 Yeast-3-Hybrid

The yeast-3-hybrid system is a variation on the more established yeast-2-hybrid system.⁵⁷ This technique is relatively new so there are only a few examples of

its use in the literature.^{58,59,60,61,62} Like the yeast-2-hybrid, reporter genes are used as the read out. The compound for which the protein target(s) is to be identified is attached through a linker unit to a second known compound which has a known protein binding partner (for example methotrexate and its binding partner dihydrofolate reductase DHFR, Fig. 6). This compound hybrid is then used to fish out potential protein target(s). The protein binding partner for the known compound is linked to a DNA-binding domain and all the possible binding partners for the compound under investigation are each bound to a transactivation domain. Bringing the DNA-binding domain in close proximity with the transactivation domain causes activation of the reporter gene. Therefore, when the protein target binds to the compound, the reporter gene is activated. This process requires the use of a cDNA library to transform yeast already expressing the DNA-binding domain-known protein target complex.

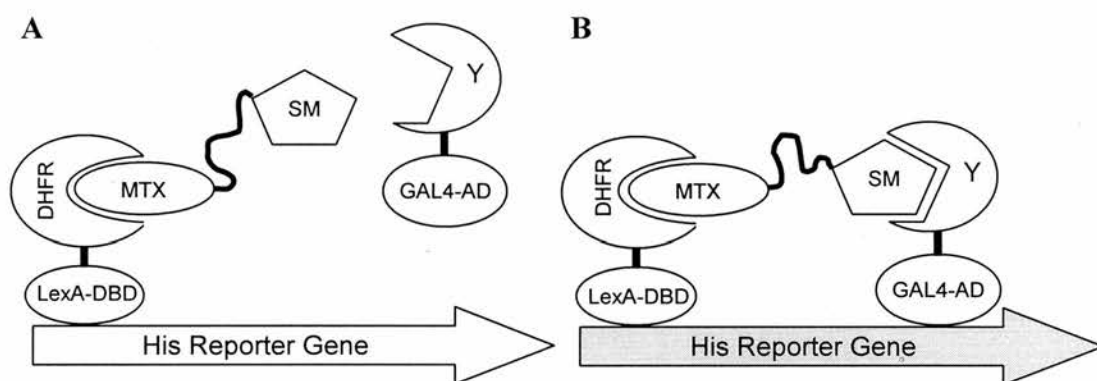


Figure 6. An example of the yeast 3-hybrid approach. **A** Two compounds methotrexate and the compound of interest (SM) are linked together. Methotrexate is used as it is a ligand for a known binding partner dihydrofolate reductase (DHFR). DHFR is expressed as a fusion protein with a DNA-binding domain (LexA-DBD). This DNA-binding domain binds to an operator for a reporter gene (His3). Activation of the reporter gene requires the DNA-binding domain (LexA-DBD) and the transactivation domain (GAL4-AD) of a transcription factor to be in close proximity. **B** This close proximity will only be achieved if protein (Y) bound to the transactivation domain binds to SM. Therefore, when the target protein(s) (Y) of SM bind to it activation of the reporter gene occurs. Image modified from F. Becker *et al.*, *Chemistry and Biology*, **2004**, *11*, 211-223.⁶²

Becker *et al.* have recently used the yeast 3-hybrid system to identify binding partners for purvalanol B using compound hybrid **8**.⁶⁰

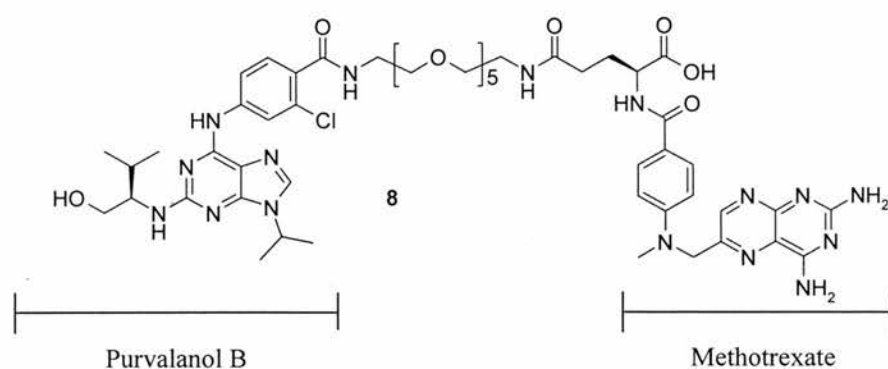


Figure 7. Purvalanol B-methotrexate hybrid.⁶⁰

The identification of CDK2, CDK1 and CK1, (known targets of purvalanol B⁶³), as binding partners of purvalanol B validated the experiment. However several new binding partners were also discovered: CLK1, CLK2, CLK3, PCTK1, PCTK2. This experiment highlighted the possibilities for yeast 3-hybrid experiments though it may be some time before this technique is used routinely.

1.4 Protozoan parasites

Protozoan parasites are responsible for some of the more devastating diseases seen in the world. Figures have been put forward that estimate that in the region of 3 billion people suffer from diseases caused by these parasites.⁶⁴ The better known diseases of protozoan parasites are: malaria, leishmaniasis, sleeping sickness and Chagas' disease.

Within the kingdom protozoa there are various groups and phyla. *Leishmania* and *Trypanosoma* species, (responsible for leishmaniasis, sleeping sickness and Chagas' disease), belong to the kinetoplasta phylum and *Plasmodium* (the genus responsible for malaria) and *Toxoplasma*, (which causes toxoplasmosis) belong to the apicomplexa phylum (also known as sporozoa).⁶⁵

1.4.1 Apicomplexan parasites

Members of the apicomplexa phylum are identified by the mechanism of locomotion employed by the parasite and the presence of an apical complex.⁶⁵ These

parasites do not use a flagellum or cilia to move, but they do exhibit a gliding type of motility. The apical complex consist of secretory organelles, microtubules and a conoid (see section 1.5.2.4). Members of this phylum include: *Plasmodium*, *Toxoplasma*, *Eimeria*, *Isopora*, *Sarcocystis*, *Cryptosporidium*, *Theileria* and *Babesia*.

There are other similarities between the members of this phylum. Invasion is a prerequisite for establishing an infection across the phylum.⁶⁶ Unlike many other infective organisms, the apicomplexans invade cells through an active invasion process driven by the parasite, with little or no contribution to the process from the host cell. Our current level of understanding of the invasion mechanism suggests that although it differs subtly across some species, (*Plasmodium* and *Toxoplasma*), it is generally conserved.^{67,68}

1.5 *Toxoplasma gondii*

Toxoplasma gondii is a protozoan parasite of the apicomplexa phylum. This parasite is an obligate intracellular parasite. This results from the fact that survival of the tachyzoites (the infective parasites within humans) outside of a host cell is restricted to time periods of between 6 and 12 hours. In the region of a quarter of the world's population are infected with *T. gondii*. However, only when individuals are immunocompromised and in the case of the unborn foetus does the disease known as Toxoplasmosis become evident.

In *T. gondii*, the apical complex includes the conoid, consisting of spirally bound α -tubulin filaments. The conoid protrudes in the early stages of invasion (see section 1.5.2.3). The *T. gondii* subpellicular cytoskeleton consists of 22 microtubules that are attached to a polar ring at the anterior of the parasite (a polar ring also exists at the posterior). This polar ring is part of the apical complex. There are three known specialist organelles involved in the secretion of proteins by the parasites: the micronemes, rhoptries and dense granules. The micronemes and rhoptries are part of the apical complex and secretion from these organelles is important in motility and host cell invasion (see section 1.5.2.4). Secretion from the dense granules occurs after host cell invasion and is connected to the modification of the parasite vacuole.

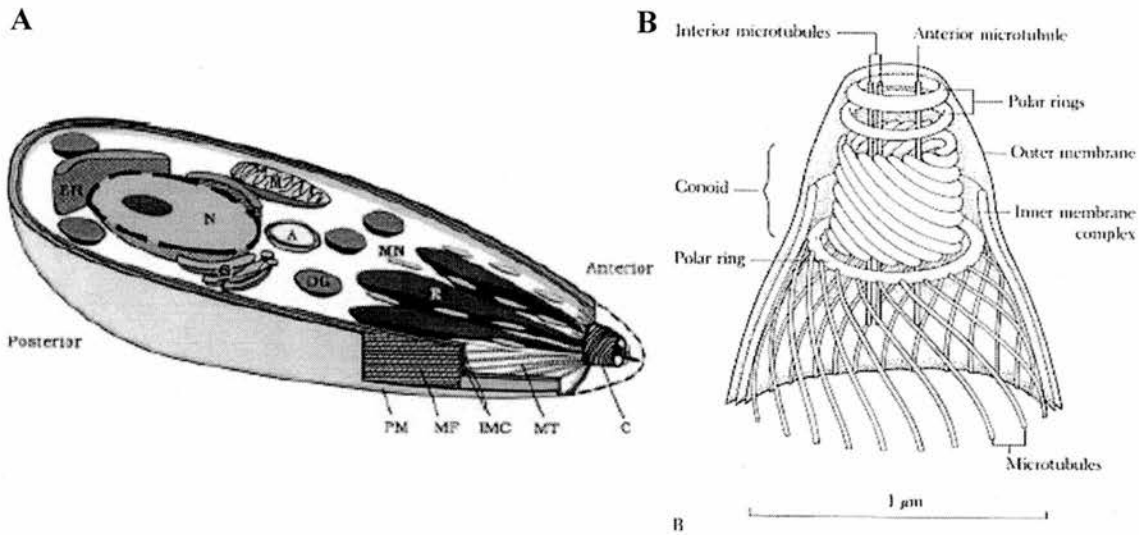


Figure 8. Cellular structure of *T. gondii*. **A** The structure of a tachyzoite of *T. gondii*. The single celled parasite contains the organelles: golgi apparatus (G), endoplasmic reticulum (ER), nucleus (N), mitochondria (M). Organelles specific to *T. gondii* and other apicomplexans include: micronemes (MN), rhoptries (R), dense granules (DG), conoid (C) and an apicoplast (A). Microtubules (MT) radiate out from the anterior of the parasite. The parasite is encased in a plasma membrane (PM) (also called outer membrane), inside this membrane is a inner membrane complex (IMC) and sandwiched between the two are microfilaments (MF). Image reproduced from V. B. Carruthers, *Parasitology International*, **1999**, *48*, 1-10.⁶⁹ **B** Detailed view of the apical complex. The conoid consist of spirally wound fibres, the protrusion of which is a key step in invasion. 22 microtubules make up the subpellicular cytoskeleton and are anchored to the polar ring at the apical end of the parasite. Image reproduced from Fig. 2.2, chapter 2, *Cell Biology of Toxoplasma gondii*, E. R. Pfefferkorn, *Modern Parasite Biology*, David J. Wyler editor, W.H. Freeman & Co, **1990**.⁷⁰

Like other eukaryotic cells, *T. gondii* tachyzoites contain a Golgi apparatus, an endoplasmic reticulum, mitochondria and a nucleus. The membrane which surrounds the parasite consists of both an outer membrane and an inner membrane complex. Located between the inner and outer membrane complex is a protein complex made up of actin, myosin and other associated proteins which collectively constitute a molecular machine known as the *glideosome*. The glideosome is responsible for the movement of the parasite (see section 1.5.2.1). *T. gondii* also contains an apicoplast, an extrachromosomal genome (25 to 37 kb) contained within a multi-walled organelle. The exact function of this plastid is not fully understood and is under intense investigation.⁷¹

1.5.1 *T. gondii* life cycle.

The *T. gondii* lifecycle, like many members of the apicomplexan phylum can generally be considered as a repetition of three events: 1) host cell invasion, 2) parasite replication and 3) host cell lysis (Fig. 9).

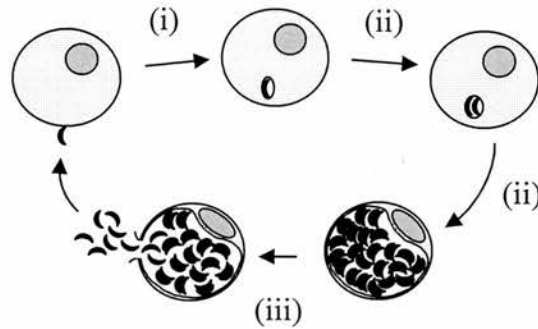


Figure 9. Overview of *T. gondii* lifecycle. *T. gondii* lifecycle consists of (i) host cell invasion, (ii) parasite replication and (iii) host cell lysis.

Although *T. gondii* invades cells and establishes an infection in humans, cats are the definitive host of the parasite. This means that within the cat the parasite replication component of the life cycle is both sexual and asexual.⁷² The lifecycle observed within the definitive hosts is a coccidia type lifecycle (Fig. 10), similar to that of *Eimeria*.⁶⁵ Invasion of the epithelial cells lining the gut of the definitive host occurs after ingestion of the parasite oocyst. The parasites (in the sporozoite form) then replicate asexually by schizogony and eventually leave the host cell causing lysis. This asexual replication occurs a number of times. Eventually the parasites invade a cell and differentiate into male (microgametes) or female (macrogametes) gametocytes. The microgametes have flagella which they use to locate and fertilise the macrogametes to produce a zygote (sexual replication). A cyst wall forms around the zygote which is then excreted from the host, eventually forming the infective oocyst, which goes on to infect another host and continues the cycle.

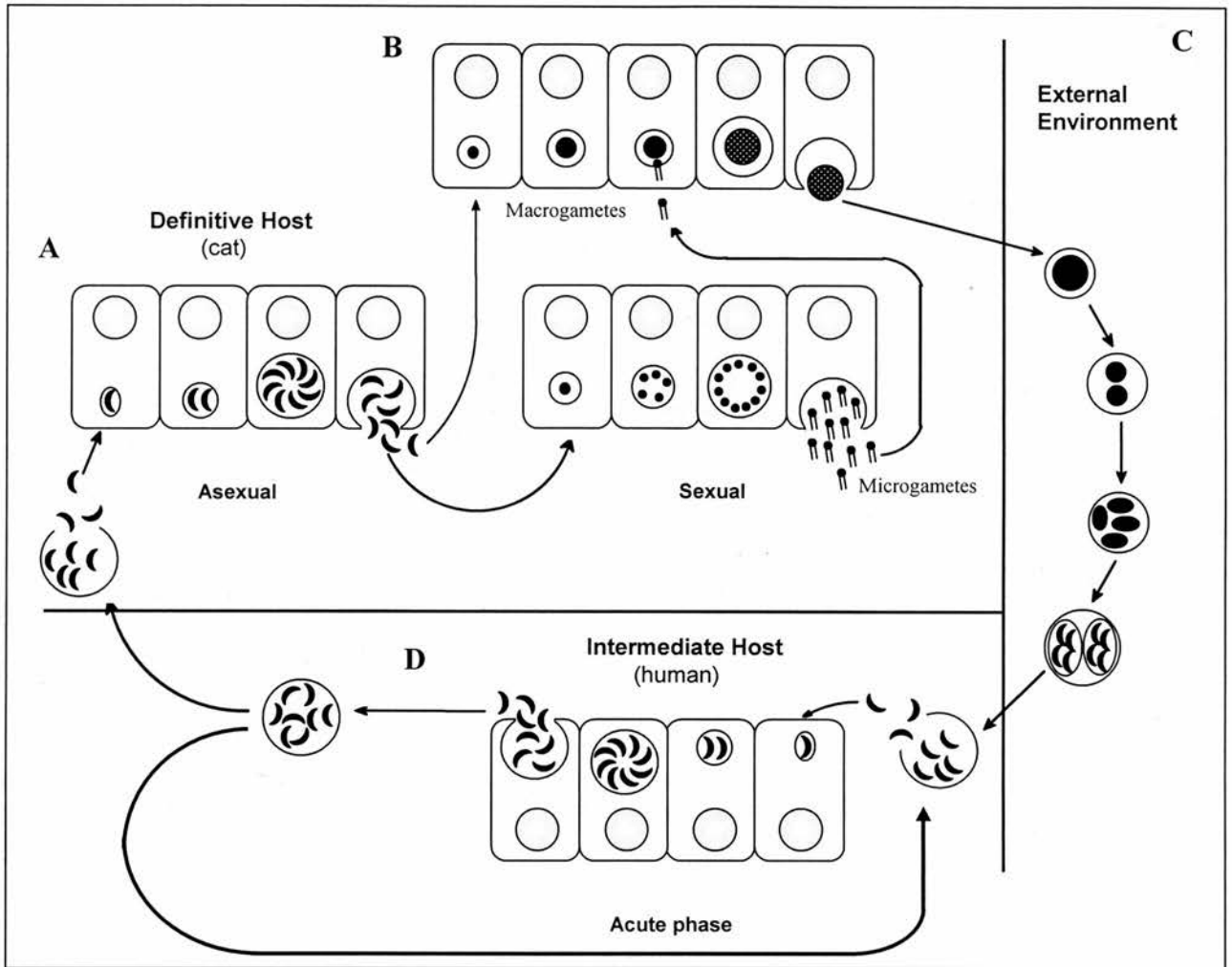


Figure 10. The lifecycle of *T. gondii*. **A** Within the definite host the parasites invade the epithelial cells of the gut. The parasite then under goes repeated events of asexual replication. **B** Once a number of cycles of replication has been achieved the parasite then forms microgametes and macrogametes and replicates sexually. The microgametes fertilise the macrogametes to form zygotes. These zygotes develop a cyst wall and then are secreted from the host. **C** While the parasite is in the external environment the zygote divides to form two sporocysts each containing four multiple sporozoites which are released upon ingestion of the oocyst. **D** Intermediate host can be come infected by the parasite through ingestion of ooysts or infected individuals. Within the intermediate host the parasite undergoes asexual replication within any nucleated cell during the acute phase of infection. Once a chronic phase is reached the parasite takes the form of cysts located in the brain or central nervous system. The parasite will remain dormant there unless the host is ingested by another intermediate or a definitive host where upon it will establish an infection in that species.

If the oocysts are ingested by another organism capable of being an intermediate host (this includes humans, sheep, rodents, etc.) then an alternative life cycle is invoked. Within the intermediate host the parasite is able to invade any

nucleated cell. The parasite invades the cell and replicates asexually (also known as endodyogeny). After approximately 2 days the parasites (in the region of 64 to 256 parasites per cell⁷⁰) leave the host, a process called egress, which causes lysis of the host cell.

This replication is part of the acute stage of infection and as the parasite load increases, the infection enters a chronic stage (see Fig. 10). During this stage the parasites form cysts in the brain or central nervous system of the intermediate hosts. These cyst are persistent and within some intermediate hosts they will remain for the rest of the host's life.⁷³ When ingested (by either definitive or intermediate hosts) these cysts are infectious and so the parasite's life cycle continues.

The transmission route of this parasite to humans occurs through the ingestion of the parasite through the consumption of unpasteurised milk or infected meat, or alternatively by coming in contact with oocyst containing cat faeces. This infection causes a disease known as toxoplasmosis which in more serious cases can lead to death in immunocompromised individuals and the unborn foetus when transferred through the placental barrier from the mother.

In summary, *T. gondii* is an obligate intracellular parasite, its infective nature in humans is due to its ability to invade any nucleated cell in the human body in order to replicate. This necessity to invade host cells illustrates the importance of the invasion process for parasite survival.

1.5.2 Invasion

Whilst progress towards understanding invasion has increased dramatically, the process is far from completely understood. Invasion and the cellular components involved in the process have been the focus of many studies and a series of excellent reviews have been written in this area.^{67,68,69,74,75,76,77,78,79,80,81,82}

The invasion process (as shown in Fig. 11) involves the parasite moving along the surface of the host cell until invasion is initiated. During invasion, interactions between proteins on the surface of both the parasite and host cell are clearly important to the successful completion of this process. After initiation of invasion, conoid extension occurs⁸³ followed by an up-regulation in secretion from the micronemes.⁸⁴ Secretion from the rhoptries occurs along with active internalisation of the parasite

driven by a myosin-based motor.^{74,85} The parasite modifies the host membrane to form a parasitophorous vacuole (PV) around itself as it enters the host cell.⁸⁶

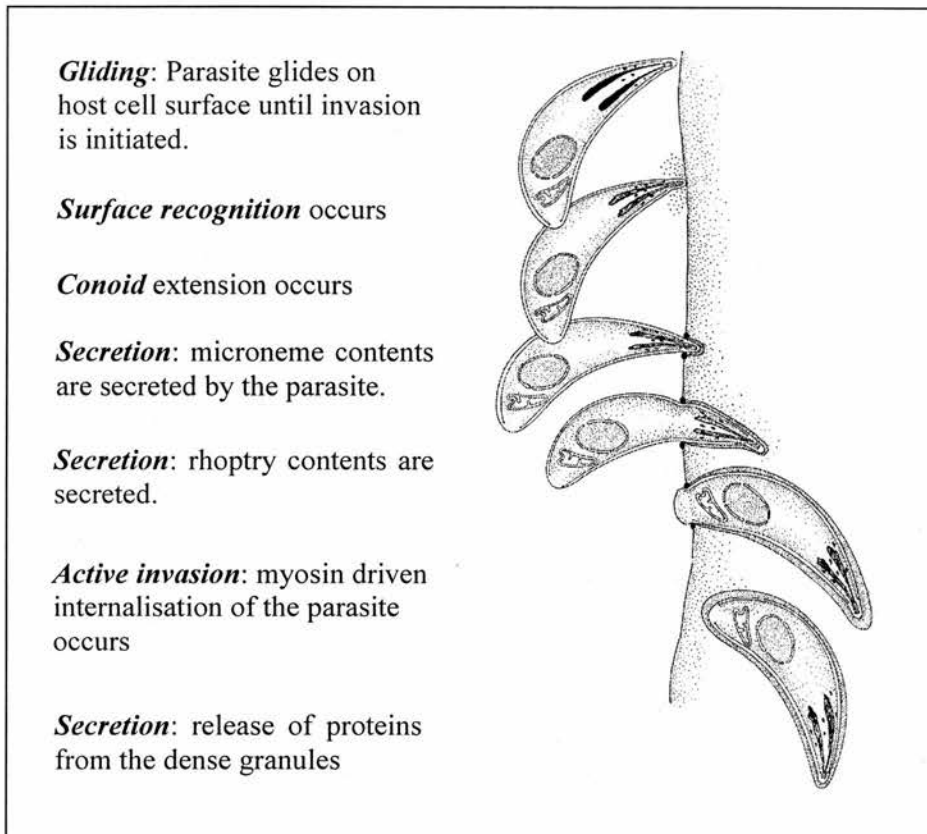


Figure 11. Overview of the invasion process of *T. gondii*. The parasite glides along the surface of the host cell. Conoid extension occurs followed by an increase in secretion of proteins from the micronemes and rhoptries. The parasite is driven into the developing parasitophorous vacuole (PV) by a myosin motor. Proteins (and possibly lipids) secreted from the dense granules modify the functional properties of the PV. Image reproduced from Figure 2.4 from chapter 2, *Cell Biology of Toxoplasma gondii*, E. R. Pfefferkorn, *Modern Parasite Biology*, David J. Wyler editor, 1990, W.H. Freeman and Company.⁷⁰

The events which occur during invasion can therefore generally be broken down into four categories: motility, surface recognition, conoid extension and secretion. Each of these categories and the proteins involved in them will be discussed.

1.5.2.1 Motility

Movement of *T. gondii* across cell surfaces and into cells is known as gliding. As discussed previously the glideosome is the name given to the protein complex believed to be responsible for this movement (Fig. 12). To date, investigations have identified four components of the glideosome: TgMyoA, TgMLC1, TgGAP50 and TgGAP45. These are all located in-between the inner membrane complex and the plasma membrane (Fig. 12).

TgMyoA is a class XIV myosin. Through the use of a tetracycline-inducible transactivator system, TgMyoA was shown to be essential for invasion and motility.⁸⁷ Severe impairment in both processes was observed when the levels of TgMyoA were reduced. TgMLC1 is a myosin light chain which is known to associate with TgMyoA.⁸⁸

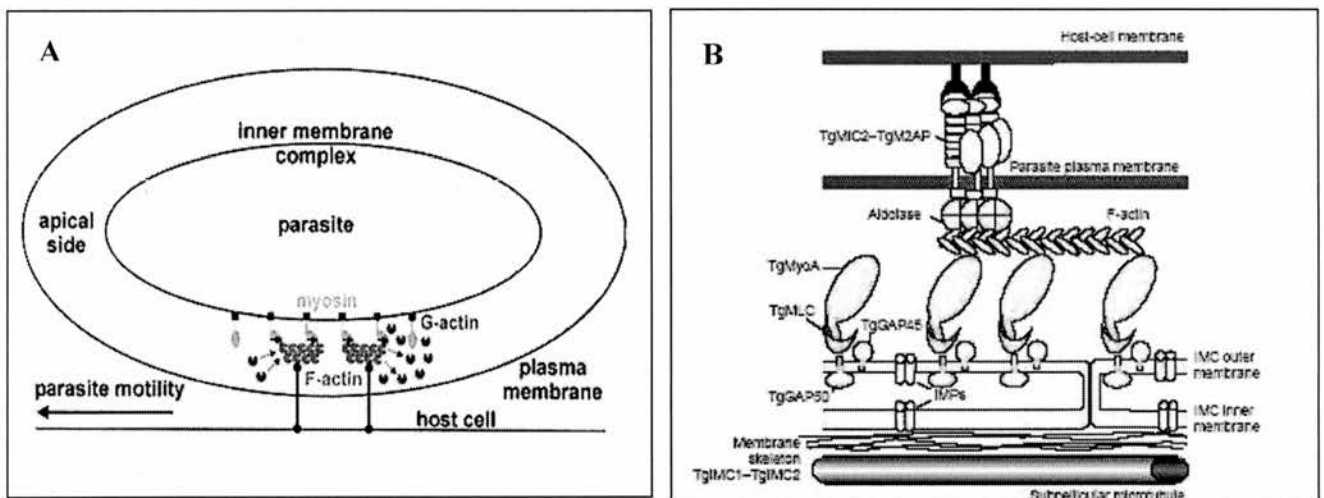


Figure 12: Components of the glideosome. **A** Schematic illustrating the suggested cellular location of the glideosome within apicomplexan parasites. Image taken from S. Schmitz *et al.*, *J. Mol. Biol.*, **2005**, 349, 113-25.⁸⁹ **B** Diagram showing the individual components of the glideosome. Image taken from A. Keeley *et al.*, *Trends in Cell. Biol.*, **2004**, 14, 528-532.⁷⁴

TgGAP50 and TgGAP45 are both attached to the inner membrane complex.⁹⁰ Immunoprecipitation studies have shown that both TgGAP45 and TgGAP50 bind to TgMyoA and TgMLC1. TgGAP50 acts as an anchor connecting the glideosome to the inner membrane complex. Both TgGAP45 and TgGAP50 appear to be essential

as judged by the inability of researchers to make viable knockouts of either of these proteins.⁹⁰

Movement is thought to be activated through the myosin TgMyoA ‘walking’ along actin filaments and hydrolysing ATP in the process using similar cellular components that induce movement in mammalian systems. Actin from *T. gondii* has recently been isolated⁹¹ and immunofluorescent methods have located actin to the conoid and plasma membrane.^{91,92} *T. gondii* actin filaments are shorter than actin found in other organisms (such as rabbit).⁹³ Jasplakinomide (**9**), an actin stabilising reagent, has been shown to increase actin polymerisation in *T. gondii* (as it does in mammalian cells^{94,95}), leading to inhibition of invasion. This illustrates the importance of tight regulation of actin polymerisation in parasite invasion.⁹⁴

Several apicomplexan actin binding proteins have been identified that may control actin polymerisation dynamics in the parasite, including a pair of 32/34 kDa proteins associated with heat shock protein 70, homologues of coronin, α -actinin, tropomyosin and spectrin-like proteins as well as an actin depolymerising factor (ADF) and a protein named toxophilin from *T. gondii* (for a recent review of this area see ⁹⁶). Recent studies have also identified an aldolase in both *T. gondii*⁹⁷ and *P. falciparum*⁹⁸ that provides the essential link between TgMIC2/TRAP and the parasite’s actin filaments. Pharmacological disruption of gliding motility through targeted disruption of aldolase function may prove to be important.

It is clear that we are at an early stage of understanding the control of actin in parasites. Control systems exist in mammalian systems, the equivalent of which has yet to be identified in parasites. An example of this is profilin which is known to bind actin and control polymerisation in mammalian systems,⁹⁹ although *T. gondii* profilin has been identified¹⁰⁰ its significance in the control of actin, (and therefore potentially movement), has yet to be determined (see section 5.5 for further discussion).

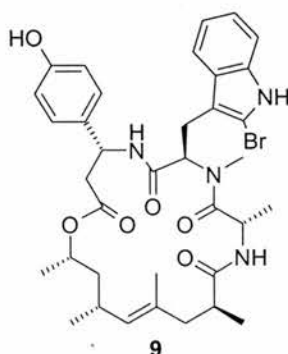


Figure 13. Actin stabilising agent jasplakinomide (**9**).

Gliding of the parasite is achieved through the interaction of all of these components with each other. Translation of these interactions into movement of the parasite is achieved through the interconnectivity of the glideosome to the parasite and the host cell membrane. The anchor which connects the glideosome to the inner membrane complex (IMC) of the parasite is TgGAP50. This is connected to TgMLC1 and in turn TgMyoA. The myosin in turn is believed to bind directly to the actin filaments. An aldolase links the actin filaments to a microneme protein known as TgMIC2 (1.5.2.4). These microneme proteins have an extracellular domain on the surface of the parasite which binds to proteins on the host cell surface (see section 1.5.2.4). In the current working model as the myosin 'walks' along the actin filaments the microneme proteins are moved from the anterior to the posterior of the parasite. This translates into movement along or into the host cell as the microneme proteins are bound to the essentially stationary host cell.

1.5.2.2 Surface recognition

There are two types of components present on the surface of *T. gondii*: surface antigens and apical invasion proteins. Apical invasion proteins are the proteins secreted from the micronemes and rhoptries which are transported to the parasite surface so they can bind to receptors on the host cell (see section 1.5.2.1 and 1.5.2.4). These proteins are either transmembrane proteins or associate with transmembrane proteins in order to be displayed on the parasite surface.¹⁰¹

Surface antigens cover the parasites surface secured by a glycosylphosphatidyl inositol (GPI) anchor.^{101,102} This anchor provides lateral fluidity but it also means the surface antigens are not connected to any of the parasites internal machinery. The role of the surface antigens is therefore recognition only. These proteins form the initial reversible contact of the parasite to the host cell. Investigations using knockouts of SAG 1¹⁰³ and SAG 3¹⁰⁴ have shown that removing either of these proteins reduces the ability of the parasite to invade host cells. Dzierszinski *et al.*¹⁰⁴ were unable to produce double knockout mutants to observe the effect of removing SAG 1 and 3, demonstrating their essential role. However, incubating SAG3 knockouts with monoclonal antibodies for SAG1 (previously shown to inhibit parasite invasion¹⁰⁵) caused nearly 90% inhibition of invasion.¹⁰⁴

The surface antigens are cleaved off the parasites surface with proteases throughout the invasion process.¹⁰¹ SAG1 has been shown to be cleaved from the parasite surface, however, the protease responsible has not been identified.¹⁰¹

1.5.2.3 Conoid extension

Conoid extension is one of the first steps of the invasion process. The exact function of the conoid remains unclear,⁸³ although it is thought to be an important component in the mechanics of invasion.⁸³

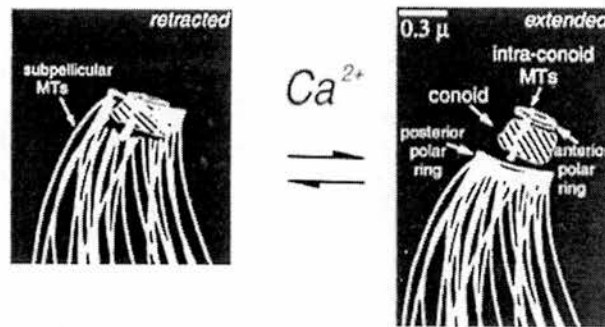


Figure 14. Conoid movement alternating between retraction and extension in response to cytoplasmic Ca^{2+} concentration. Adapted from K. Hu *et al.*, *J. Cell Biol.*, **2002**, *156*, 1039-1050.¹⁰⁶

The conoid consists of α -tubulin polymerised in such a way that it forms tight spirals. Extension and retraction of the conoid can be triggered in the laboratory by an increase or decrease in intracellular calcium levels through the addition of calcium chelators or ionophores (Fig 14).^{106,107}

1.5.2.4 Secretion

During invasion apicomplexan parasites secrete proteins from three different types of organelle: the micronemes, rhoptries and dense granules.

Micronemes are small organelles that release proteins from the apical end of the parasite. Some of these proteins are thought to form bridges between host receptors and the internal motility machinery of the parasite (see Fig. 12, section 1.5.2.1). The subsequent “capping” of these proteins from the anterior to posterior of the parasite is an integral part of the motility mechanism in apicomplexans (as shown

schematically in Fig. 12 and described above). On arrival at the posterior of the parasite, these proteins are cleaved from the parasite by proteases and shed into the extracellular media. Microneme secretion is a constitutive process, consistent with the continuous requirement for microneme proteins in motile parasites and appears to be upregulated during invasion.¹⁰⁸ Microneme secretion can also be experimentally upregulated by the addition of calcium ionophores, and inhibited by intracellular calcium chelators.^{108,109,110,111} Many of the micronemal proteins undergo maturation through regulated proteolytic processing and the formation of heterooligomeric complexes, prior to their transport to the parasite surface. This maturation has been shown in some instances to be required for invasion to occur.^{112,113}

The function of most of the rhoptry proteins is unknown, but a recent proteomic analysis of the contents of the rhoptries offers great promise for understanding the role these proteins play in invasion and establishment of infection.¹¹⁴ A subset of the rhoptry proteins (e.g., ROP2^{115,116}) is thought to function in the recruitment of host cell organelles (endoplasmic reticulum, mitochondria) to the membrane of the parasitophorous vacuole (PV). Unlike microneme secretion, it appears that secretion from the rhoptries is not triggered by changes in calcium levels.

Although microneme and rhoptry proteins are considered to be essential for invasion, dense granule proteins are thought to be involved primarily in modification of the PV. Whilst there is evidence that these organelles contribute to the parasite's constitutive secretory system,¹¹⁷ it has been shown that a large burst of protein secretion from the dense granules occurs within the first hour after complete formation of the PV.^{118,119,120} What triggers this release remains unclear but as in the case of the rhoptries, changes in calcium levels are not thought to be involved.¹¹⁷

1.5.2.4.1 Proteases involved in invasion

The secretion of proteins is clearly important to the invasion process however, equally important are the parasite proteases. This importance is reflected in the large amount of interest recently shown in proteases which have been the subject of several reviews.^{117,121,122,123}

Most microneme and rhoptry proteins are known to be multiply processed: during intracellular transport, at the time of secretion and (in the case of microneme

proteins) on the surface of the parasite.^{124,125,126,127,128} A variety of candidate microneme and rhoptry protein proteases have been identified in *T. gondii*.^{129,130,131,132} In particular, there is currently a great deal of interest in parasite proteases of the rhomboid family, which appear to be involved in cleaving the transmembrane domain of microneme proteins on the surface of the parasite, thereby severing the links between the parasite and the host cell at the end of invasion.^{133,134,135}

1.5.2.5 Summary

In this section, the invasion process has been split into 4 distinct phases and each has been overviewed. Existing studies clearly indicate that many of the components discussed are essential for invasion to occur. Invasion is a complex process and we are only just beginning to integrate the knowledge we have. It is also clear that many pieces of the puzzle have yet to be identified. There is therefore large scope and indeed the need for the identification of new proteins and/or new connections between known proteins and the invasion process.¹³⁶ This view was the primary driving force of the work described in this thesis and is discussed in more detail in the rest of this chapter.

1.6 Forward chemical genetics approach to studying *T. gondii* invasion.

Given that invasion is essential in the life cycle of *T. gondii* and FCG allows the study of essential processes, a high throughput screen was developed to identify compounds which affect invasion.¹³⁷ This work was initially carried out by Dr Nicholas Westwood in the laboratory of Professor Gary Ward (University of Vermont) and at ICCB, Harvard Medical School in the laboratory of Professor T. J. Mitchison.

1.6.1 Overall project goals

This forward chemical genetics project on *T. gondii* invasion set out to achieve a number of goals both in the long and short term:

1. To advance our understanding of the invasion process in apicomplexa in general through the use of *T. gondii* as a representative parasite within the phylum.
2. To use a FCG approach to identify compounds that are of relevance to goal 1.
3. To use our skills in synthetic organic chemistry to drive these studies through to identification of proteins involved in parasite invasion.

This includes:

- i. Synthesis of original hits to confirm their activity (chapter 2).
 - ii. Synthesis of derivatives to determine a position which can be altered but allowing for the biological activity to be retained (chapter 2).
 - iii. Synthesis of a small collection of derivatives in order to optimise the physical and biological properties (chapter 3).
 - iv. Synthesis of compounds suitable for loading onto an affinity resin (chapter 4).
 - v. Synthesis of the resin and its use in affinity chromatography experiment (chapter 4 and 5).
4. To deliver protein-compound pairs of relevance to *T. gondii* and possibly for other apicomplexa invasion research. These compounds may also be of relevance to drug discovery studies.

In order to achieve these goals the first step is the successful development and implementation of a high through-put screen.

1.7 High Throughput screening in parasitology

Chemical genetics is not a technique which has been frequently employed in parasitology. High throughput screens have been performed, but more frequently using collections of compounds based on known inhibitor scaffolds against known targets. There is to our knowledge only one example of a whole cell FCG screen using an *unbiased* library, that is the work done by Carey *et al.* to identify inhibitors of *T. gondii* invasion (see 1.8 for full discussion on this work).¹³⁷

1.7.1 Screening against purified proteins

Screening against a purified protein (RCG) is a more common approach to screening but there are only a few instances where unbiased compound libraries have been used.^{106,138} An excellent example of this approach is the work by Asai *et al.*¹³⁸ *T. gondii* expresses the protein nucleoside triphosphate hydrolase (NTPase) in two isoforms NTPase-I and NTPase-II. These enzymes are released into the parasite-containing vacuole and are essential for tachyzoite replication within the host cell. Approximately 150,000 compounds were assayed against NTPase in *T. gondii* in an automated 96-well format. The activity of the compounds was monitored by the presence of orthophosphate, formed by hydrolysis of ATP. Five compounds **10-14**, were shown to inhibit both NTPase-I and NTPase-II (see Fig. 15). One of the compounds, **10**, was a selective inhibitor of NTPase-I. Importantly on further investigation all five compounds were found to inhibit parasite replication.

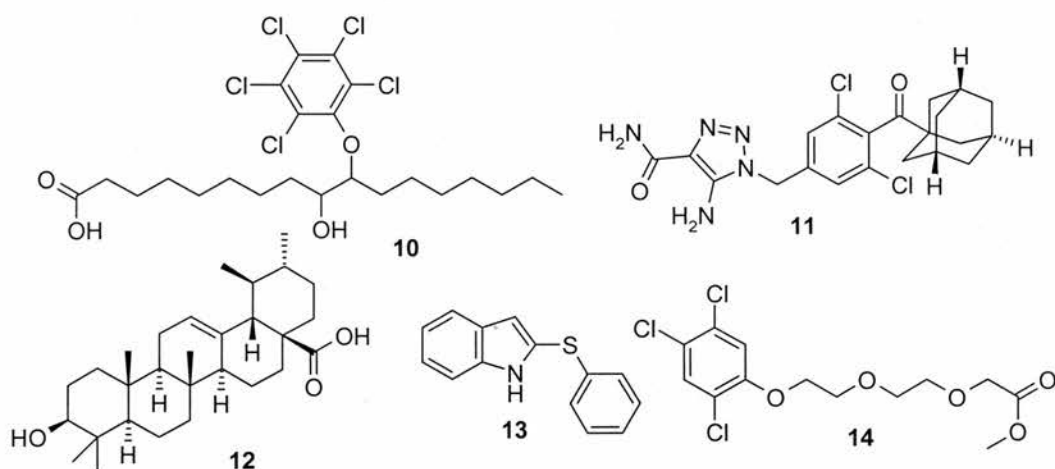


Figure 15. Compounds identified as inhibiting both NTPase-I and NTPase-II.

More recently a screen against *P. falciparum* dihydroorotate dehydrogenase identified 1249 compounds, out of a library of 220,000 compounds, over 12 of which inhibited the enzymes activity at submicromolar levels.¹³⁹

1.7.2 In silico screening

A second class of HTS projects that use an *unbiased* compound set does exist but this set is a “virtual” one.¹⁴⁰ Table 1 summarises reports of projects that have started with an “*in silico*” screen against a specific protein. In the majority of cases, these studies are aided by the availability of structural information on the chosen protein, although homology models are also used.

Protozoan Parasite	Protein Target	No. of Compounds; Hit Rate [†]	Reference
<i>Leishmania donovani</i>	Dihydrofolate reductase	25,694; 0/5	141
<i>Leishmania major</i>	Cysteine proteases	150,00; 18/69	142
<i>Plasmodium falciparum</i>	Cysteine proteases	55,313; 4/31	143
<i>Plasmodium falciparum</i>	Dihydrofolate reductase	230,000; 12/24	144
<i>Plasmodium falciparum</i>	Dihydrofolate Reductase- Thymidylate Synthase	n.a.; 2/21	145
<i>Tritrichmomonas foetus</i>	Hypoxanthine-Guanine- Xanthine Phosphoribosyl transferase	599; 2/32	146
<i>Trypanosoma cruzi</i>	Dihydrofolate reductase (DHFR)	56, 868; 2/12	147
<i>Trypanosoma cruzi</i>	Hypoxanthine phosphoribosyltransferase	3,400,000; 16/22	148
<i>Trypanosoma cruzi</i>	Trypanothione Reductase	2,500; 9/13	149
<i>Trypanosoma cruzi</i>	Trypanothione Reductase	n.a., 1/1	110

Table 1. Target-based approaches using high throughput *in silico* screening. [†] In this table the hit rate is defined as the number of compounds which showed *in vitro* activity/ the number of compounds screened *in vitro*. The examples were chosen according to the following criteria i) the species studied was a protozoan parasite, ii) an initial *in silico* screen was performed followed by a functional screen using the purified protein and the computationally identified hits. The following compound databases were used: Available Chemical Database (ACD), Cambridge Structural Database (CSD) and the Fine Chemical Directory.

For example Bond *et al.* used this approach to study the protein target trypanothione reductase (TR) in the protozoan parasite *T. cruzi*.¹¹⁰ This protein was selected due to its potential importance in the survival of the parasite. TR reduces trypanothione disulfide (T[S]₂) to dihydrotrypanothione, which can in turn reduce glutathione disulfide (GSSH) to glutathione (GSH). GSH protects cells by removing potential oxidising agents. Bond *et al.* obtained a crystal structure of TR with its T[S]₂ substrate bound. An area of the active site of TR which is the most structurally dissimilar to the human equivalent glutathione reductase (GR) was chosen to screen the Cambridge Structural Database (CSD) for potential inhibitors.

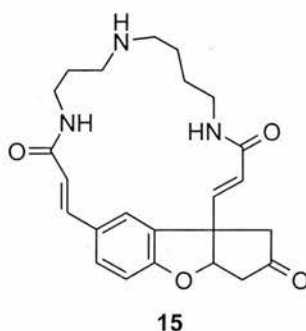


Figure 16. The natural product lunarine identified as inhibiting TR.

The natural product lunarine (**15**), was identified from this screen and considered suitable for further investigation. Lunarine (**15**), which is commercially available, showed inhibition kinetics indicating non-covalent binding to TR followed by covalent binding to the reduced form of TR.

1.8 High throughput screen against *T. gondii* invasion.

A commercially available library of 12,160 compounds (Chembridge DiverSet) was screened in a high throughput screen against *T. gondii* invasion.¹³⁷ The invasion assay involved observing the number of parasites that were inside host cells after a given “invasion time”. Parasites which express yellow fluorescent protein (YFP) were added to a 384 well plate containing host cells. DMSO stock solutions of the compounds had already been added to these wells (one compound per well). The plates were then incubated at 22°C for 15 min to allow the parasites to be exposed to the compounds prior to invasion. Parasites are not able to invade at 22°C, but when the plates are then warmed to 37°C, the parasites are at the optimum temperature for

invasion. The parasites were then incubated with the host cells at 37°C for 1 hour. Removal of the extracellular media was followed by the addition of a fluorescently labelled antibody to SAG1, a surface antigen on *T. gondii*. All the parasites were YFP-expressing parasites (pseudo-coloured in green Fig. 17), those parasites which were external to the host cells were labelled by the antibody and so appear red (Fig. 17). Merging the two images means internal parasites appear green and external parasites appear yellow (combination of green and red). The amount of internal (green) parasites in a given field of view compared with the number in the same size area of a control well indicates whether the compound is an inhibitor or an *enhancer* of invasion (Fig. 17 & Fig. 18).

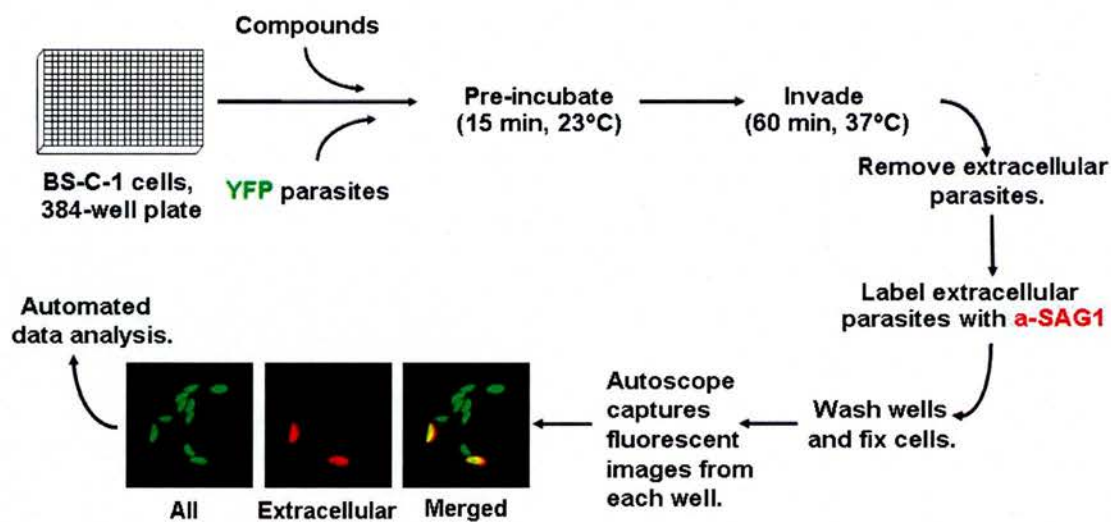


Figure 17. A chemical genetics high throughput screen for inhibitors of *Toxoplasma gondii* invasion. Each well of a 384-well plate contained either a unique compound or an equivalent volume of DMSO (control). Intracellular and extracellular parasites were distinguished by fluorescence microscopy; extracellular parasites were red and green, appearing yellow in merged images; intracellular parasites were green only. Compounds that reduced the invasion levels to <20% compared to control wells were considered inhibitors.

The effect produced when the invasion was inhibited was dramatic with virtually no green parasites evident (Fig. 18). In addition to the identification of compounds which inhibit invasion, compounds which *enhanced* invasion were also identified. The images from the relevant wells show many more internalised parasites per field compared with the control (Fig. 18). Compounds which enhanced invasion were

observed as having a dramatic effect on the number of parasites contained within an individual cell with an increase from 1 parasite to 50-70 parasites in some cases.

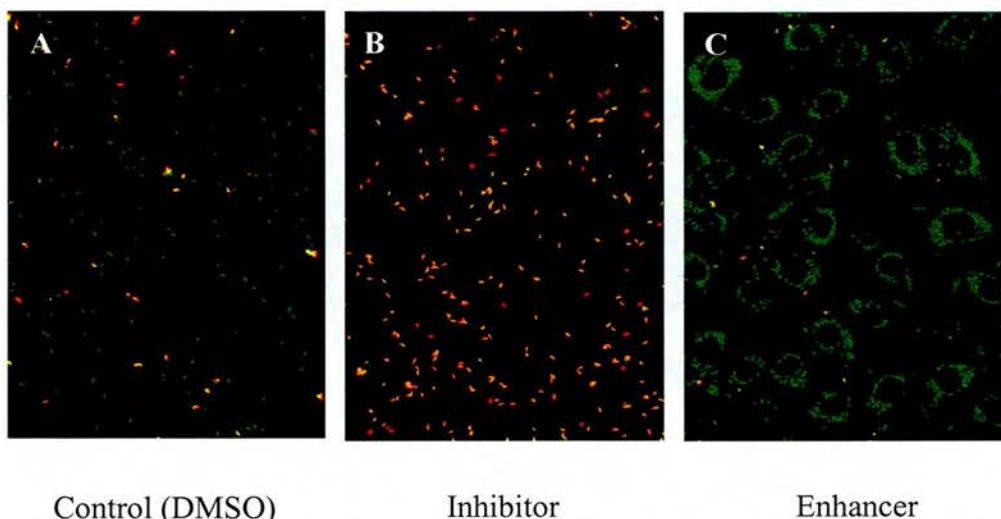


Figure 18. Images produced from the high through-put screen. **A** Image from a control well where DMSO has been added. **B** Image from a well containing an inhibitor assigned as such due to the absence of green parasites and a large increase in yellow parasites. **C** Image from a well containing an enhancers, assigned as such due to the large number of green parasites and the very small number of yellow parasites.

Of the 12,160 compounds assayed, 24 inhibitors (**16-40**) and 6 enhancers (**41-46**) of invasion were identified (Fig. 19 and Fig. 20). Some of these compounds can be assigned to similar structural classes. For example compounds **16**, **18**, **28**, **29** and **38** all contain a trichloromethylaminal group. Compounds **22** and **39** both contain a piperazine group. Interestingly inhibitor **35** and enhancer **42** are both derivatives of hexahydro-1H-pyrazino[3,2,1-jk]carbazole. Whilst having a similar structure does not guarantee that the compounds have the same mechanism of action, it is suggestive of this. Identifying the protein target(s) of any of these compounds would lead to interesting conclusions with regards to mechanism of action. There are questions of stereochemistry which also need to be investigated relating to a number of these compounds (information not available from the commercial supplier).

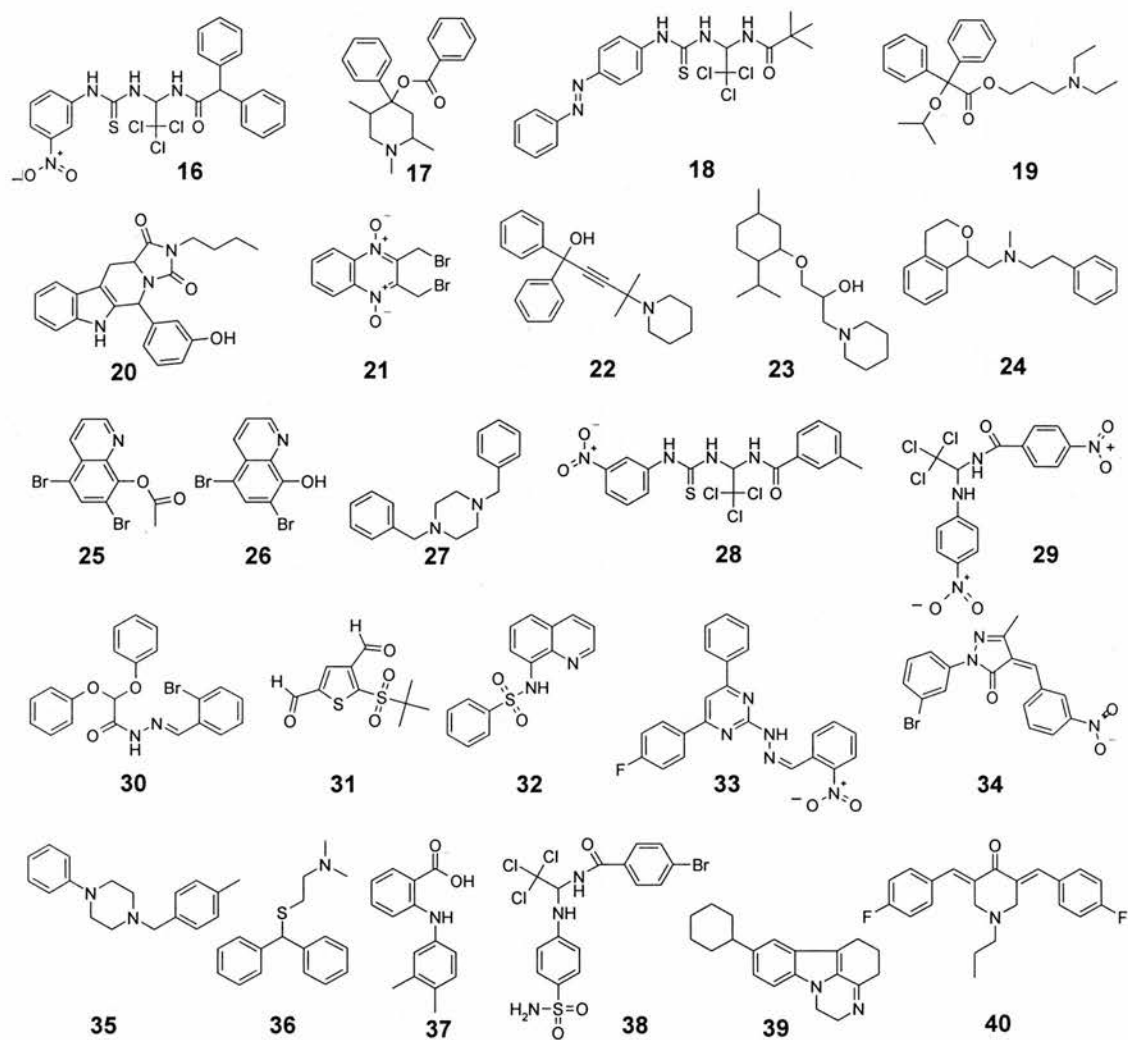


Figure 19. 24 inhibitors of invasion identified. The well for compound 25 was contaminated with 26. Where stereochemistry is not shown, it is not provided by Chembridge.

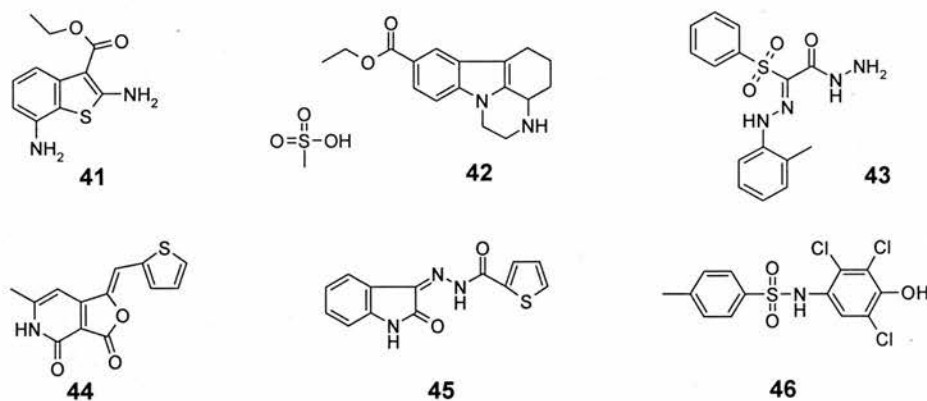


Figure 20. Compounds identified as enhancers of invasion.

Having identified the bioactive compounds, dose response assays were used to determine their potency (Fig. 21). The potency was defined by the lowest active concentration (the concentration within the dilution series at which a significant difference in levels of inhibition/enhancement compared with DMSO control was observed).

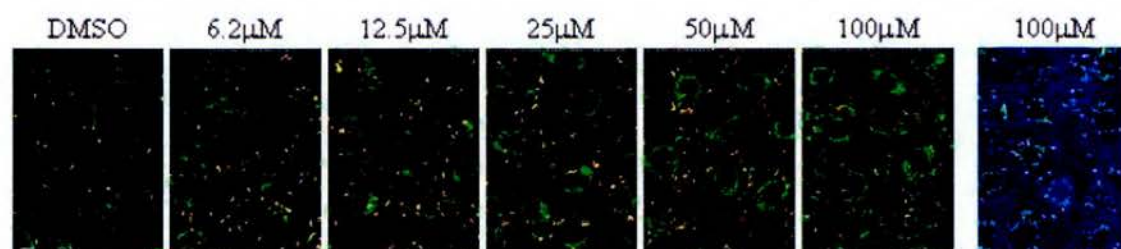


Figure 21. Dose-dependent enhancement of invasion by **41**. In the merged images shown here, intracellular parasites are green and extracellular parasites appear yellow. The enhancement by **41** is concentration dependent, with a lowest active concentration of $12.5\mu\text{M}$.

Cytotoxicity assays were also used to confirm that the 24 inhibitors and 6 enhancers were not cytotoxic to either the cells or the parasites on the timescale of the assay. In addition, liquid chromatography-MS analysis of all the hits confirmed the samples purity as $>95\%$ as judged by UV absorbance.

1.8.1 Secondary assays

The reversibility of the compounds was assessed by washing the compound out after the initial pre-incubation period. Of the 30 compounds, 25 were found to act reversibly. Target cell identification of the five irreversible compounds **16**, **18**, **20**, **21** and **26** was achieved by pre-treatment of the parasites with compound, washing the parasites, then addition of the parasites to the host cell in the usual manner. The experiment was repeated this time pretreating the host cells rather than the parasites. Three of the compounds targeted the parasite only: **20**, **21** and **26**. Compounds **16** and **18** targeted both the host cell and the parasite.

As discussed in section 1.5.2, several well defined processes are known to play a key part in the invasion process. In order to gain further information as to the

possible mechanism of action of all 30 compounds, secondary assays were used to determine the effect of the compounds on motility, secretion and conoid extension.

1.8.1.1 Motility

T. gondii deposits “slime trails” consisting of parasite plasma membranes and proteins during the gliding process, these can be visualised by the addition of an antibody to the major surface protein SAG1. The effect of each of the compounds on parasite motility was determined by examining the effect on the slime trails produced (Fig. 22).

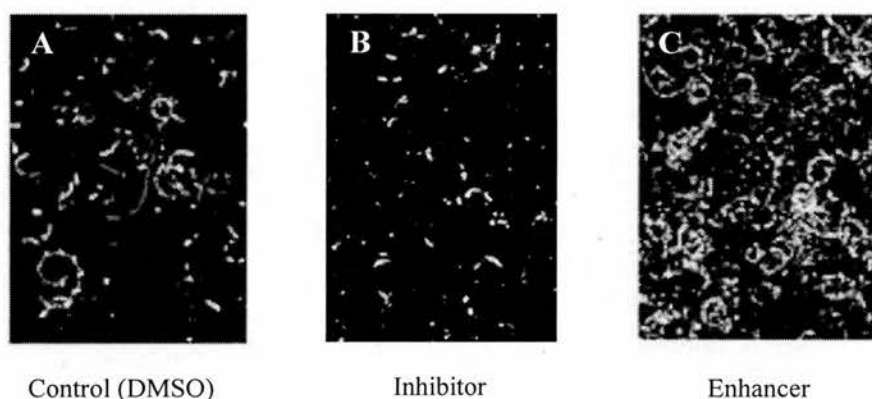


Figure 22. Motility trail assay. Trails deposited by the parasite are imaged using an antibody against the surface protein SAG1. **A** Trails produced by parasites imaged by adding a fluorescently labelled antibody to SAG1. **B** The trails produced by parasites inhibited with compounds. **C** Trails produced by a parasites incubated with an enhancer. The trails are much more extended than in the control.

21 of the invasion inhibitors also inhibited motility. The dose dependency of this effect was also observed for these compounds and it was found that the majority of these compounds inhibited invasion by inducing a state of rigor. This was determined by observing the effect of the compounds on live parasites using videomicroscopy. However three of these compounds, **22**, **31** and **36**, induced an odd and non-productive form of motility. All of the compounds which enhanced invasion also enhanced motility apparently increasing the time that they are motile rather than inducing an increase in speed of gliding.

1.8.1.2 Secretion

One of the secondary assays performed on hits from the *T. gondii* invasion assay asked whether the bioactive compounds had an effect on constitutive microneme secretion. The extent of constitutive secretion was assessed by incubation of the parasites in the presence of the compound for 90 minutes. The parasites were then pelleted and the supernatant probed for the presence of the microneme protein MIC5¹⁵⁰ by western blotting.¹³⁷ Of the 24 invasion inhibitors tested, 18 were shown to inhibit secretion of MIC5 in this assay. Four of the invasion inhibitors had no effect and, surprisingly, two of the inhibitors **27** and **34** resulted in an increase in the secreted levels of MIC5. All of the enhancers caused an enhancement in constitutive secretion.

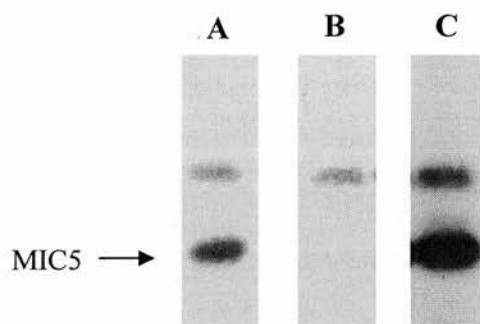


Figure 23. Compound effect on microneme secretion determined by western blotting. **A** Control levels of MIC5 secretion. **B** Inhibition of secretion after addition of an inhibitor. **C** Enhancement of secretion after addition of an enhancer.

Use of MIC2 as a marker of microneme secretion gave similar results. Rhoptry and dense granule secretion were not affected by any of these compounds, as judged by monitoring the levels of ROP4 and GRA8 respectively.

The effect of the compounds on calcium ionophore-induced secretion was also determined. Surprisingly, only 11 of the 18 inhibitors of constitutive microneme secretion inhibited induced secretion. This pharmacological uncoupling of constitutive and induced secretion demonstrated that it was possible to influence parasite secretion with a degree of precision through the use of compounds. The detailed mechanism by which these compounds are influencing secretion is the source of on-going studies within the Westwood and Ward laboratories.

1.8.1.3 Conoid extension

The effect of the 30 compounds on the extension and contraction of the conoid was determined by visualising the parasites by phase microscopy. Induced conoid extension was achieved by the addition of the calcium ionophore, ionomycin (**47**). Compounds were added and the parasites visualised by phase microscopy in order to determine if they inhibited induced conoid extension. Out of all the compounds only one compound had an effect on conoid extension, **21**. **21** inhibited ionomycin induced conoid extension.

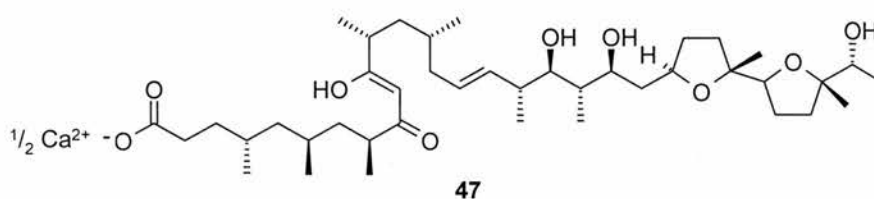


Figure 24. The calcium ionophore Ionomycin (**47**).

In summary, the secondary assays have provided important information as to the mechanism of action of these compounds which will help to direct target identification studies. Dissection of the invasion process can be achieved using the different inhibitory properties of these compounds to investigate the interdependence of motility, secretion and conoid extension.

1.8.2 The activity of the compounds against other apicomplexans

As discussed previously, the similarities between the members of the apicomplexan phylum are well established. Having identified compounds which inhibited the invasion of *T. gondii*, a large collaborative effort involving N. J. Westwood, G. E. Ward, K. L. Carey, D. Wirth, K. Matuschewski, F. Tomley, G. Widmer, E. Suvarnapunya and M. Stein assessed the 30 hit compounds in invasion assays using *P. knowlsei*, *P. falciparum*, *P. berghei*, *E. tenella*, *C. parvum* and *Salmonella typhimurium*. *S. typhimurium* was chosen as a control species in order to show that the compounds were selectively interacting with a part of the apicomplexan invasion.

Importantly in the context of the work described in this dissertation, the results for *P. knowlesi* showed that out of the 30 compounds, 11 inhibitors and 1 enhancer (**44**) also inhibited/enhanced invasion in this system. Assays against *P. berghei* showed that 11 of the inhibitors also inhibited invasion of *P. berghei*. Seven of these compounds were the same as those which inhibited *P. knowlesi*. Although **44** did not enhance the invasion of *P. berghei*, it did result in an odd motile behaviour. This identification of compounds which inhibit and enhance both *T. gondii* and two *Plasmodium* species provides a set of tools which can be used to study apicomplexan parasites in general. Identification of the cellular target(s) of these compounds will therefore provide information not only of relevance to *T. gondii* but also for other apicomplexan parasites. Intriguingly, compound **37** did not score in any of the secondary assays or against any other parasites, this is apparently a selective compound which only targets *T. gondii* invasion by a mechanism that does not inhibit motility, secretion or conoid extension.

1.8.3 Selection of the compound to be studied

High throughput screens of the type described in section 1.8 almost invariably lead to the identification of a large number of hits (30 in this case). The next major challenge is to decide which of the hits to prioritise for further study. The selection of a compound for target ID studies can be driven by either biological or chemical issues but more likely by an balance of both. Compounds which effect invasion in other apicomplexan parasites in addition to *T. gondii* were of highest priority. In addition to this, the phenotype of enhancement has not been frequently observed before therefore identification of the cellular target(s) involved in this phenotype could provide interesting insights into the invasion process.

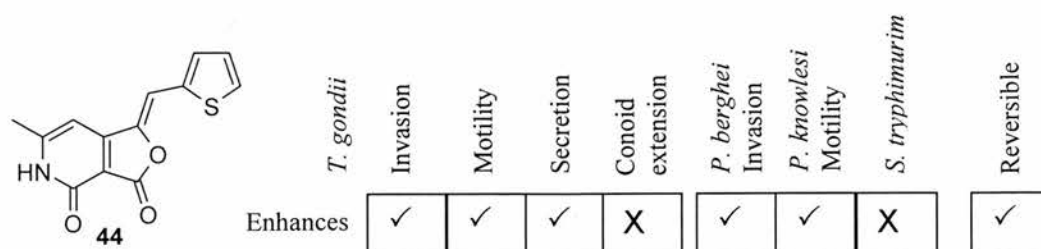


Figure 25. Compound **44** and a profile of its biological activity.

Compound **44** enhances invasion in *T. gondii* and *P. knowslei* this makes it an interesting compound to study for both of the reasons discussed above. The phenomena of enhanced invasion is one which was observed as being both dramatic and interesting and could in fact be used to aid the culturing of parasites by facilitating the invasion process. The effect on the host cell of many parasites invading it has not previously been established and raises interesting questions as to whether there might be a limit to the number of parasites which can replicate successfully within one host cell. A compound which enhances invasion could be used to investigate this and may even be considered of use in drug discovery. The activity of this compound across species means that any information about this parasite could be widely applicable and provide valuable information with regards to the invasion process of apicomplexa. From a chemical perspective **44** would appear to be a potential covalent modifier of proteins (due to the strained furanone) however, the secondary assay revealed that the compound has a *reversible* effect on invasion. This seeming contradiction means investigations into the mechanism of action of **44** are required.

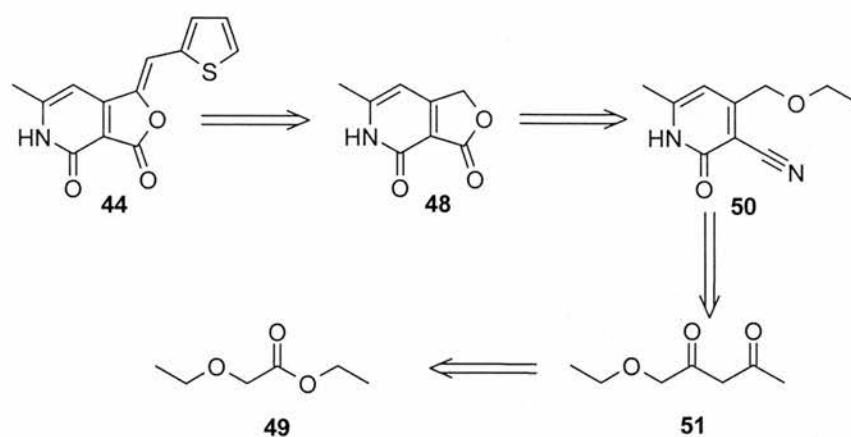
Literature searches into the synthetic route to **44** indicated that the synthesis of this compound could be amenable to alterations required for target ID studies (see section 1.8.4.1). For all of these reasons, **44** was selected as the compound to study.

1.8.4 Literature search on the selected compound **44**

Informed approaches to target ID experiments or optimisation of a compound's physical and biological properties require assessment of any prior synthetic work on the chosen compound. In addition, any known biological activity and other possible cellular targets already linked to the desired phenotype must be checked.

1.8.4.1 Synthetic precedent

Compound **44** has been synthesised before from **48**.¹⁵¹ The synthesis of furanone **48** has been well documented from ethyl ethoxyacetate **49**.^{152,153} Therefore, the retrosynthetic analysis of **44** led to the route shown in scheme 1. A more detailed discussion is given in Chapter 2.



Scheme 1. Retrosynthetic analysis of **44**.

Synthesis of derivatives with alternative aromatic rings in place of the thiophene ring has also been achieved using **48** as the starting material.¹⁵⁴ This may provide a useful starting point for the synthesis of analogues for target ID studies.

1.8.4.2 Known biological properties

Biological activities pertaining to compound **44** have not been published in the literature. However, a high throughput screen identified **52** as an inhibitor of mammalian arginine *N*-methyltransferase.⁵⁵ Although not identical to **44** the similarity in chemical structure between **44** and **52** may mean that the mechanism of action of these compounds are related.

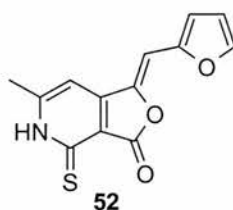


Figure 26. Compound identified as an inhibitor of mammalian arginine *N*-methyltransferase.

Experiments designed to investigate this possibility have been discussed but have not yet been pursued in the Westwood/Ward laboratories. Arginine *N*-methyltransferase is a protein which either monomethylates or dimethylates the guanidine nitrogen atoms of arginine side chains on a wide array of proteins.⁵⁵ This

modification is implicated in various protein trafficking, signal transduction and transcription regulation. BLAST searching the *T. gondii* genome did not highlight a homologue in the parasite.

1.8.5 Observation of the same phenotype in the literature

When determining the target of a compound prior occurrences of the phenotype produced need to be investigated. There may be a protein which has been previously reported to cause the same phenotype, therefore making it a likely target. A recent example of this is the identification of the protein target of monastrol (**1**) (section 1.1.1).

Enhancement of invasion of apicomplexa has only been reported in a few instances. ROP1 has been shown to be part of a penetration enhancing factor.⁵⁸ The antibody Tg49 inhibits this factor and has been shown to recognise ROP1. This penetration enhancing factor appears to enhance the efficiency with which the parasite invades cells.^{155,156} Addition of antibodies for SAG2, T4 3GI1 and T4 IF5 to the parasite is known to cause an enhancement of invasion.¹⁵⁷ This effect is time dependent, with very little enhancement being observed after 1 hour. However, 69% enhancement was reported after 3 hours. Enhanced attachment was also observed after the first hour and is thought to be the cause of enhanced invasion. There may be a possible explanation to the target of **44** within these examples but further secondary assays (e.g. an attachment assay) would be required to reach any such conclusion. It is unlikely that these would lead to the identification of a signal protein target.

1.8.6 Project aims

Having selected the compound of interest the main aim of the research is to identify the cellular target(s) of **44**. Affinity chromatography was selected as the target ID method of choice. It was therefore decided that initial investigations should involve the synthesis of a number of derivatives of **52** selected to determine positions in which **44** can be modified without a loss in activity.

These novel derivatives would then be tested in the same assay against *T. gondii* invasion.

It was reasoned that once a position where modification of **44** could be tolerated was established, a compound for target ID studies could be proposed. This compound will have to contain a linker system by which to attach the compound to an affinity resin. Synthesis of this compound and loading it onto a resin to create a reagent for affinity chromatography was imagined as a key mid-term goal of the project. Affinity chromatography experiments using the reagent with the goal of identifying the cellular target(s) of **44** were then planned to take place.

Whilst assessing the synthesis of derivatives of **44** for affinity chromatography it was also decided to prepare a small library of compounds in order to identify derivatives with optimised physical and biological properties.

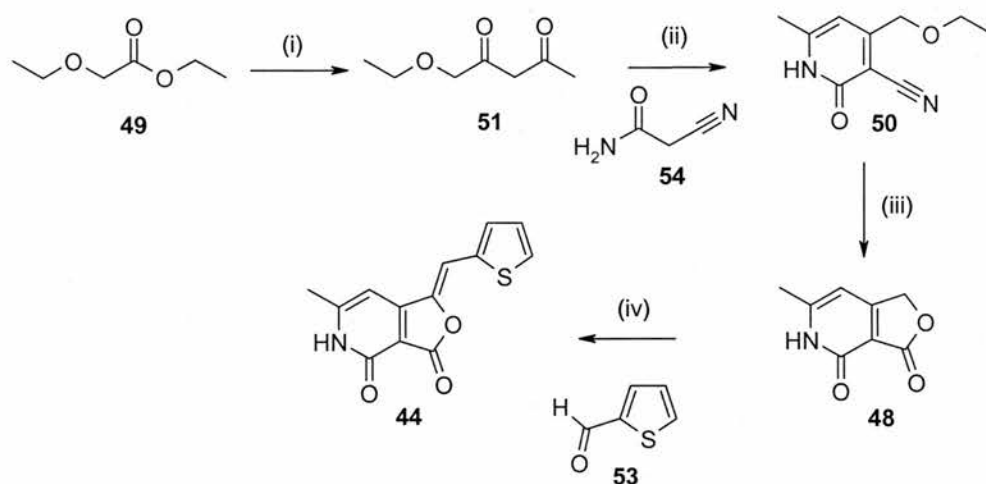
The specific goals this research set out to address were therefore:

- I. Resynthesis of **44** to confirm its biological activity.
- II. Synthesis of derivatives of **44** to determine a position which can be altered and the biological activity retained.
- III. Synthesis of a collection of derivatives of **44** in order to optimise its the physical and biological properties.
- IV. Synthesis of a derivative of **44** for loading onto an affinity resin.
- V. Synthesis of the resin based on **44** and its use in affinity chromatography experiments.

Chapter 2: Synthesis of active compound and derivatives.

2.1 Synthesis of the bioactive compound 44 identified from the high throughput screen.

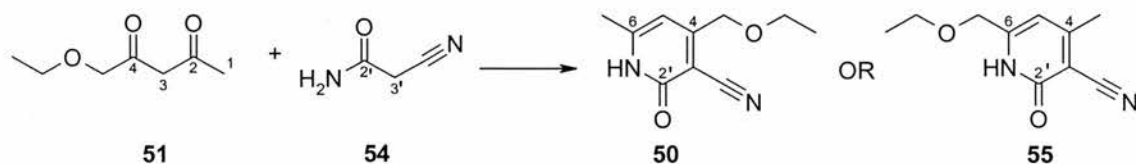
The initial goal of any chemical genetics project is to resynthesise the bioactive compound identified from the screen and confirm its biological activity. As previously discussed 6-methyl-1-[2-thienylmethylidene]furo[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**44**) was identified as an enhancer of host cell invasion by *Toxoplasma gondii*.¹³⁷ Therefore resynthesis of **44** was identified as the initial goal in this research. Investigation of the literature revealed a published synthetic route to **44** as shown in scheme 2. The final step in the sequence is a Knöevenagel-type condensation between furanone **48** and thiophene carboxaldehyde (**53**) in the presence of piperidine. Arustamova *et al.*¹⁵⁴ reported the use of this reaction to make a small number of derivatives of **44**. The product of the condensation reaction is insoluble under the reaction conditions, the implication of this is that purification can be achieved via filtration and careful washing alone.



Scheme 2. Synthesis of 6-methyl-3,4-dioxo-1-(1-thienylidene)furo[3.4] pyridine (**44**). (i) NaH, acetone, toluene, -10°C, 54%, (ii) piperidine, ethanol, 65%, (iii) 50% aqueous H₂SO₄, 125°C, 86%, (iv) piperidine, ethanol, reflux, 72%.

Synthesis of the furanone **48** was achieved using literature procedures.¹⁵³ 1-Ethoxypentane-2,4-dione (**51**) was synthesised by addition of analytical grade acetone to a solution of the anion generated by treatment of ethyl ethoxyacetate (**49**) with sodium hydride at -10 °C.

The pyridone ring in **50** was formed by condensation of **51** with cyanoacetamide (**54**) in the presence of piperidine. Hot methanol was required to solubilise the cyanoacetamide (**54**). Once all the reactants were in solution, the reaction mixture was then cooled and the desired product **50** precipitated and was recovered by filtration. Two possible products could have been formed in this condensation reaction: 6-methyl-4-ethoxymethyl-3-cyano-2(*1H*)-pyridone (**50**), or 6-ethoxymethyl-4-methyl-3-cyano-2(*1H*)-pyridone (**55**) (scheme 3). The major product isolated from this reaction in our hands was compound **50**, which is consistent with the literature.¹⁵⁸ Harris *et al.* reported that this was the only product formed. However Wenner *et al.*¹⁵⁸ were also able to isolate a small quantity (15%) of **55** by acidification and concentration of the filtrate. Bardhan *et al.*¹⁵⁹ put forward a rationale for the formation of **50** as the major product, postulating that the relative reactivity of the two carbonyl groups (C2, C4) of the diketone **51** is responsible for the selectivity and that the most reactive carbonyl will react with C3' of the cyanoacetamide (**54**) first.



Scheme 3. Formation of two possible pyridones.

Following this explanation, if the most reactive carbonyl of the diketone is C2 then C1 will end up at position 4 of the pyridone. In the case of 1-ethoxypentane-2,4-dione (**51**) the carbonyl α to the ethoxymethyl group (C4) could be considered as the most reactive carbonyl relative to the carbonyl α to the methyl (C2) due to the slightly greater level of electron donation of the methyl resulting in the C2 carbon of the carbonyl being less δ^+ (hence being less electrophilic). Therefore the product formed is **50**. This is only one possibility and the subtle differences in expected reactivity of C2 and C4 suggest that other influences may be involved. One possibility is that the

reaction may be under thermodynamic control. The mechanism of the formation is such that individual steps may be reversible, given that the yield of the reaction is 65% the possibility of additional side reactions occurring cannot be excluded.

Formation of furanone **48** involved refluxing the pyridone **50** in 50% aqueous H_2SO_4 .^{152,153} Pouring of the acid mixture onto ice resulted in the formation of colourless crystals of **48**, which were isolated by filtration in 86% yield.

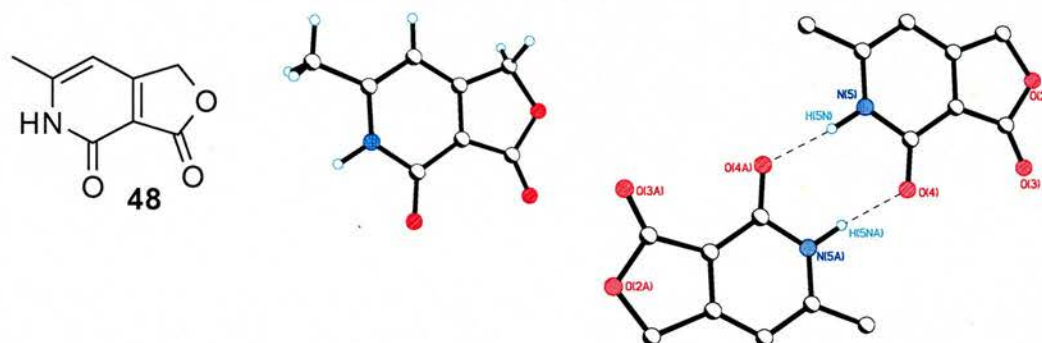


Figure 27. X-ray crystal structure of **48**. The unit cell consists of two molecules hydrogen bonding between the pyridone NH and CO groups. CCDC No. 292089.

The X-ray crystal structure of **48** shows the C(4)O(4) bond length is 1.24Å and the N(5)C(4) bond is 1.38Å. The approximate length of a C=O bond is 1.22Å and a N-C bond is 1.47Å,¹⁶⁰ this shows that **48** exist predominantly in the tautomeric form **56** in the solid phase although it cannot be ruled out that a small contribution is made from the tautomeric form **57**.

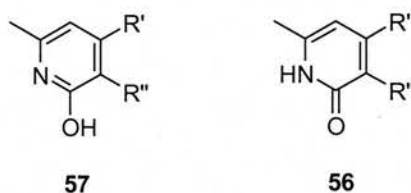


Figure 28. The two possible configurations of the pyridone ring.

The balance of these two tautomeric forms in pyridones has been widely studied in different phases.^{161,162,163} In the gas phase, the hydroxypyridine form, **57** dominates,¹⁶¹ whereas in an aqueous solvent the pyridone form, **56** is most prevalent.¹⁶² Dimers consisting of either two hydroxypyridines **57**, or two pyridones **56**, linked through hydrogen bonds, have been compared computationally and the

latter were predicted to be more stable.¹⁶² This means that whilst on initial inspection **57** would appear to be more stable tautomer due to the fully aromatised ring, in fact **56** is the more stable as it forms the most stable dimer in solution. Solvent molecules have been shown to affect the tautomeric equilibrium and so in solution there may still be a contribution from **57**.^{162,163}

The X-ray crystal structure of **48** clearly shows a hydrogen bonding network between the amide group of the pyridone rings. Additionally the lattice structure shows that these compounds π -stack with each other, as expected.¹⁶⁴ This hydrogen bonding and π -stacking network presumably accounts experimentally for the observed insolubility of **48**.

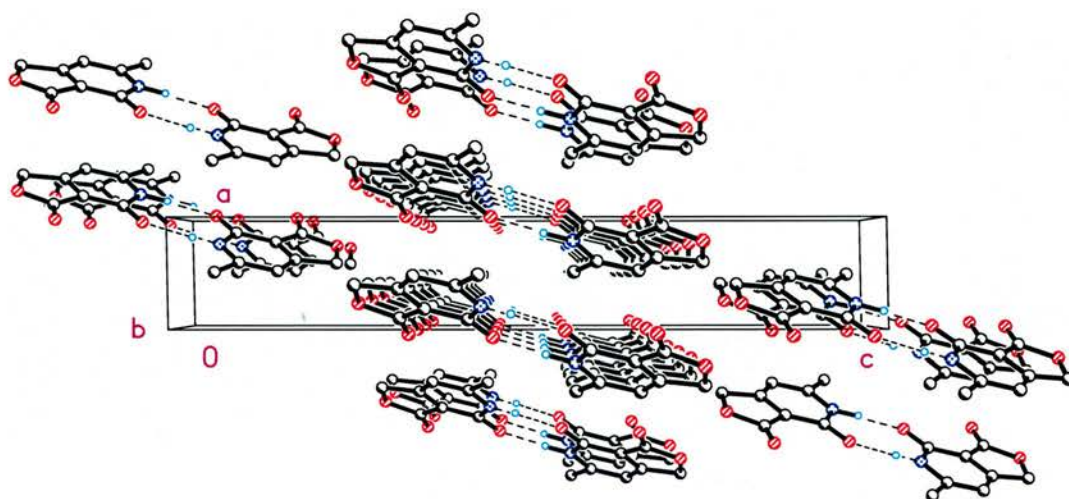


Figure 29. The packing structure of **48** from the X-ray crystal structure.

Synthesis of **44** was achieved in good yield from reaction of **48** and thiophene carboxaldehyde (**53**). The geometry of the exocyclic double bond has been drawn as *Z* in agreement with the report by Arustamova *et al.*¹⁵⁴ Studies using nOe NMR techniques on **44** showed that in solution the geometry of the exocyclic double bond was *Z*. This was confirmed by a nOe observed between H7 and H8 (see Fig. 30), this would only occur if **44** has *Z* geometry. A crystal structure was obtained which further confirmed the geometry as *Z* in the solid state (Fig. 30).

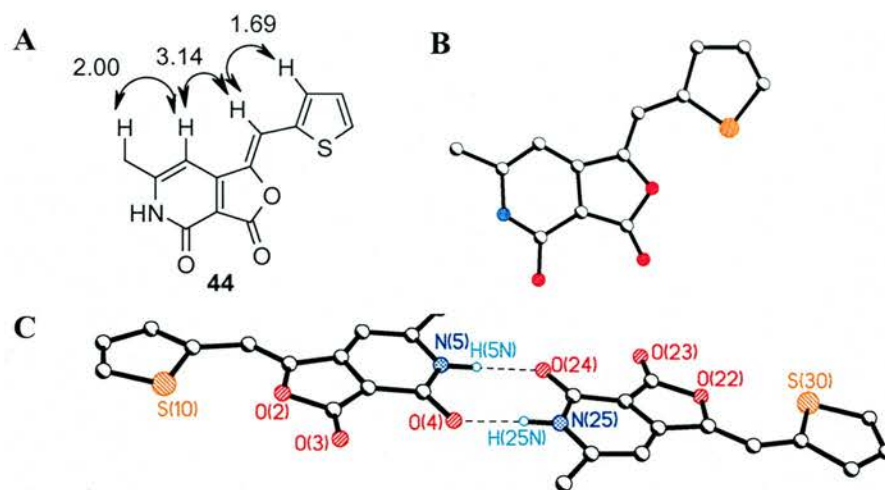


Figure 30. Crystal structure of **44** and the unit cell. **A** nOe interactions which confirm the *Z* geometry of **44**. The NOE experiment used was a selective gradient enhanced 1D NOESY experiment. **B** The geometry of the exocyclic double bond is *Z*. **C** The unit cell consists of two molecules with the existence of hydrogen bonds between the pyridone NH and CO groups. CCDC No. 292088.

Information gained from the crystal structure shows that for one of the pair of compounds which make up the unit cell the C(4)O(4) bond of the pyridone is 1.22 Å and the C(4)N(5) bond is 1.42 Å. These bond lengths are close to the length of a C=O bond at 1.22 Å and a N-C at 1.47 Å,¹⁶⁰ indicating that **44** exists in the solid phase as shown. Further evidence for compound **44** existing as tautomer **56** is the presence of a carbonyl stretch in the IR at 1673 cm⁻¹ in addition to the stretch at 1790 cm⁻¹ for the furanone carbonyl.

Complete characterisation of compound **44** required the assignment of the carbon spectrum. The large number of quaternary carbons, a number of them within 1-2 ppm of each other, does not allow for an assignment based purely on theoretical shifts using known substitution effects. 2D NMR studies using HSQC (Heteronuclear Single Quantum Coherence) and HMBC (Heteronuclear Multiple Bond Connectivity) spectra enabled the assignment of the carbons with the correlation between the 6-methyl and the pyridone carbonyl carbon being key (see chapter 3, section 3.6, for an example).

Having synthesised a pure sample of **44** and confirmed its structure, the next stage was to determine its biological activity. This was carried out at the University of Vermont in Professor Gary Ward's laboratory using the same assay procedure used in the high throughput screen (Chapter 1, Section 1.8).¹³⁷ Our batch of **44** was shown

to enhance host cell invasion by *T. gondii* at a minimum concentration of 3.1 μM, similar to that observed during the screen (later adjusted to 1.69 μM, section 2.2.1). As no degradation of the compound was observed (as judged by LCMS) this led to the conclusion that the active component identified from the screen was indeed **44**.

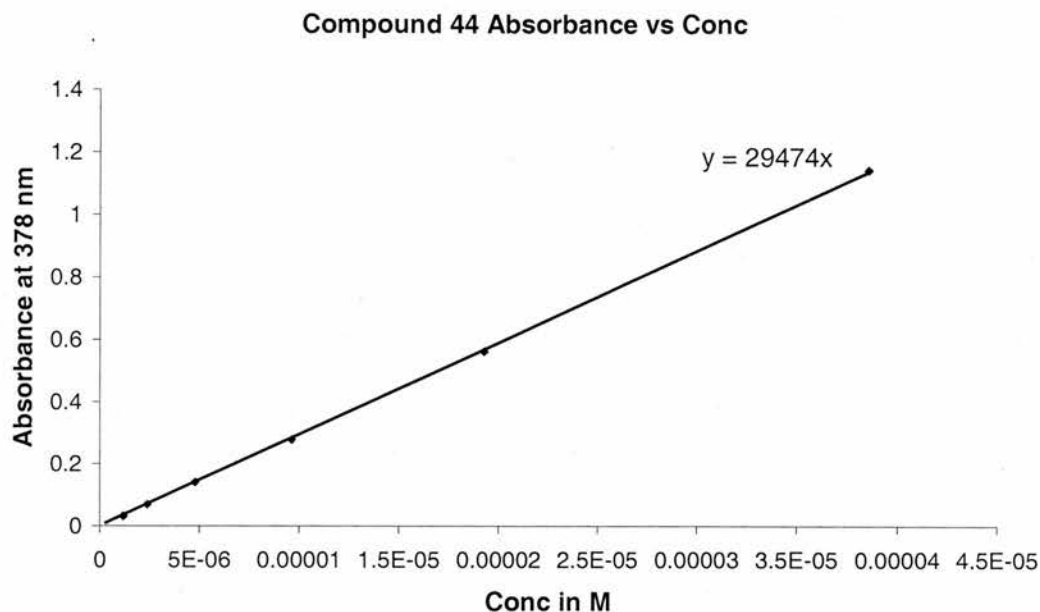
Although the assay procedure used is analogous to that used in the screen, the method of analysing the assay is different. During the screen the data for each compound were obtained using analysis software combined with an inverted fluorescence microscope, however for subsequent assays the level of enhancement was determined by eye. The enhancement is graded in terms of strong, medium and weak enhancement. Quoted values of lowest active concentration are the lowest concentration at which medium enhancement was observed (LAC).

2.1.1 Investigating the solubility of **44**.

Compound **44** is insoluble in many organic solvents, with DMF and DMSO being the only solvents **44** will dissolve in to any great extent. This insolubility can be explained by the hydrogen bonding between pyridone units and the π -stacking.

Insolubility of bioactive compounds can cause discrepancies in the observed biological activities. If the compounds are not completely dissolved, then the lowest active concentration measured will not be accurate. Therefore the solubility of **44** was investigated as follows:

Compound **44** is UV active and given that the concentration of a compound is directly related to the UV absorbance through the Beer Lambert equation: $\lambda = \epsilon cl$, the UV absorbance of a given solution could be used to determine its concentration. Therefore standard solutions of known concentrations of **44** were used to generate the extinction coefficient (ϵ). Great care was taken in preparing these solutions, in order to insure that all **44** was dissolved. UV spectra were taken of five or more solutions in DMSO at different known concentrations. The absorbance values obtained for these solutions were plotted against the concentrations (see Graph 1), giving the graph the formula $y = mx + c$ as $\lambda = \epsilon cl + 0$, where $y = \lambda$, $x = c$ and $m = \epsilon l$ (where $l = 1$ cm). Compound **44** has an extinction coefficient in DMSO of 29475 dm³mol⁻¹cm⁻¹, (as taken from the gradient of the line, Graph 1).



Graph 1. A graph plotting concentration verses absorbance for compound **44**.

Having defined the relationship between absorbance and concentration, the concentration of any given solution of **44** (in DMSO), can now be determined.

The biological assays used DMSO stock solutions which were made up to a concentration of 40 mM (theoretically), (solution **A**, see Fig. 31). These were further diluted into media to give a 168 μM solution (solution **B**, see Fig. 31). From this solution a dilution series was prepared. Compounds were then added to the cells by taking 15 μl of the 168 μM solution. The volume of media in the wells of the plate is 10 μl , causing a final concentration of 100 μM for the compound (solution **C**, see Fig. 31).

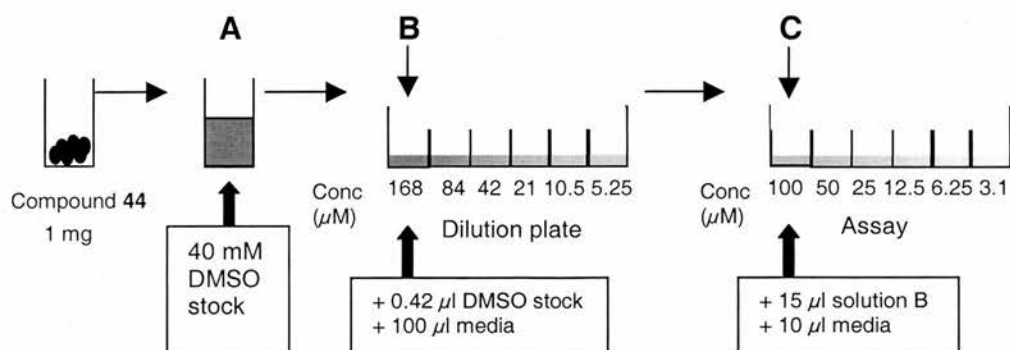


Figure 31. Schematic showing the dilution under gone by **44** under the assay conditions. **44** was dissolved in DMSO to give a 40 mM stock solution (theoretically), **A**. From that stock solution 0.42 μl was taken and added to 100 μl of media to give solution, **B**. A dilution series was prepared from

solution **B**. 15 μl of solution **B** was added to the assay which contained 10 μl of media. Giving an assay concentration of 100 μM for the highest concentration, **C**.

The actual concentration of a “40 mM” DMSO stock solution, (solution **A'**, mimicking solution **A**), was determined. To a 1 mg sample of **44** (weighed as accurately as possible using a 7 figure balance) the quantity of DMSO which would theoretically yield a 40 mM solution was added (solution **A'**). Centrifugation was used to pellet any undissolved solid. The supernatant was removed and diluted to a volume and concentration range suitable for UV analysis. Two readings were taken for the sample and the average of the two used to determine the concentration of the sample using the already determined extinction coefficient. This solution (**A'**) had a concentration of 38 mM, indicating that although **44** is not completely soluble it can be solubilised in DMSO to a large extent.

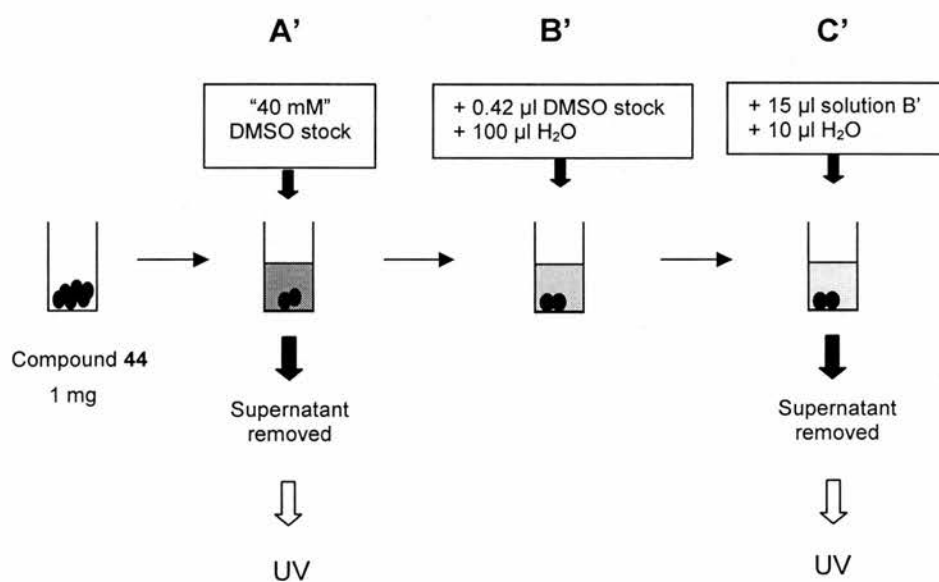


Figure 32. Schematic showing the solutions measured by UV. The quantity of DMSO required to achieve a 40 mM was added to **44**, giving solution **A'** (mimicking **A**). Any suspended solid was separated through gentle centrifugation. The supernatant was diluted by a known amount and its UV spectra taken. 0.42 μl of solution **A'** was added to 100 μl of water, giving solution **B'** (mimicking **B**). 15 μl of solution **B'** was added to 10 μl of water, giving solution **C'** (mimicking **C**). Any suspended solid was separated through gentle centrifugation. The supernatant was diluted by a known amount and its UV spectra taken.

The concentration of **44** under the assay conditions (solution **C**) is designed to be 100 μM . The DMSO solution of **44**, **A'**, was diluted into water using the same volumes as used in the assay (Fig. 31 and Fig. 32). 0.42 μl of **A'** was added to 100 μl of water (instead of media), giving solution **B'** (mimicking **B**, see Fig. 31). 15 μl of **B'** was added to 10 μl of water, giving solution **C'** (mimicking **C**). Centrifugation was used to separate any suspended solid in **C'** and the supernatant (20 μl) removed and diluted into DMSO (1ml). The UV absorbance of **C'** was determined and the concentration calculated to be 54 μM . The activity of **44** is quoted as the lowest active concentration determined using the dilution series. Given that the initial concentration 100 μM is in fact 54 μM the lowest active concentration (LAC) has been altered accordingly (see table 2). The LAC of **44** is therefore 1.69 μM (as opposed to the initially supposed 3.1 μM).

Theoretical concentration C (μM)	100	50	25	12.5	6.25	3.1
Actual concentration, C' (μM)	54	27	13.5	6.75	3.38	1.69

Table 2. Adjusted concentrations of **44** across the dilution series.

This observed difference in concentration means that **44** is more potent than previously supposed. The LAC for analogues of **44** reported have been adjusted according to their experimentally determined concentration throughout the rest of this dissertation unless otherwise stated.

2.1.2 The fluorescence properties of **44**.

In addition to contributing to the insolubility of **44**, the conjugated system resulting from the presence of the exocyclic double bond results in **44** being yellow in colour and fluorescent. Fluorescence measurements were taken of **44** (see Fig. 33). The emission maximum of **44** is 449 nm and the excitation maxima is 377 nm.

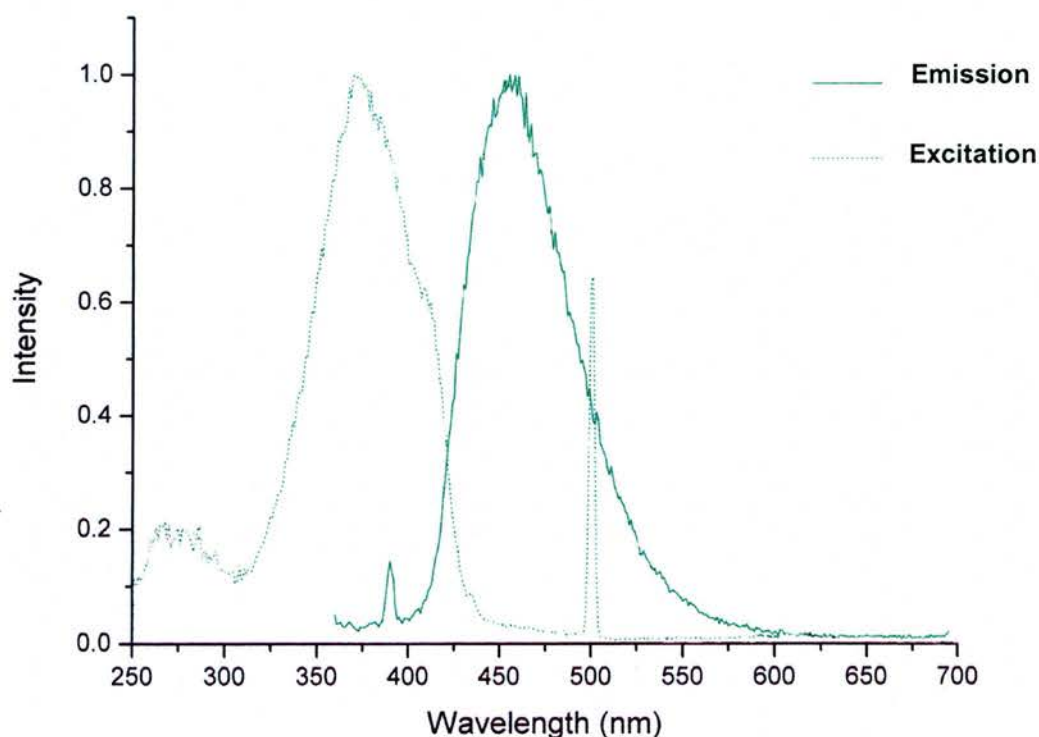


Figure 33. The fluorescence spectra of **44**.

The difference between the emission and excitation is known as the Stokes shift and describes the difference in energy between the light absorbed and the light emitted by the compound. The overlap between the excitation and emission spectra and the closeness of the two maxima means that **44**, although clearly fluorescent, **44** does not efficiently convert all the absorbed light into emitted light. This could be reached if there was no overlap between emission and excitation wavelengths, so the compound could not cause excitation of itself (meaning that not all of the emitted fluorescence is released to the environment). A more detailed discussion of the fluorescence properties of analogues of **44** is given in section 3.10.

2.2 Preliminary structure activity relationship data.

Unlike many small molecules isolated in forward chemical genetic screens, **44** appears to be relatively potent giving a robust and reproducible phenotype at low micromolar concentrations. It was therefore decided that initial SAR investigations should focus on which positions could be modified in order to incorporate a linker unit that is compatible with affinity chromatography for the purposes of target ID

studies. This required the synthesis of derivatives of **44** containing single structural modifications and then subsequent biological testing to determine if the modification had a detrimental effect on the biological activity.

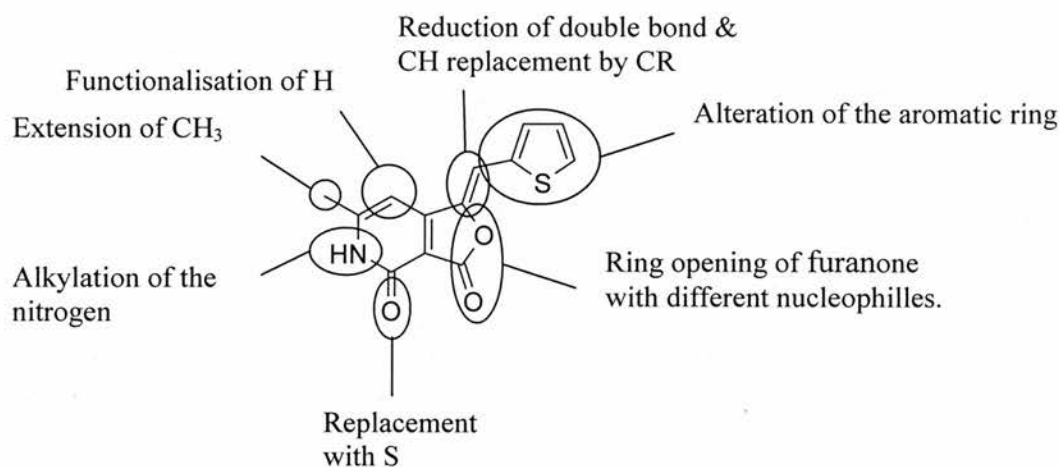


Figure 34. Areas of possible modification of compound **44**.

By examining the structure of **44** (Fig. 34) we identified at least eight structural modifications of interest: (i) alteration of the aromatic ring, (ii) alkylation of the pyridone nitrogen, (iii) reduction of the exocyclic double bond, (iv) extension of the CH₃ group, (v) replacement of the pyridone oxygen for a sulfur, (vi) functionalisation of the hydrogen in position 7, (vii) replacement of the hydrogen in position 8, (viii) alteration of the furanone through ring opening with a nucleophile (Fig. 35).

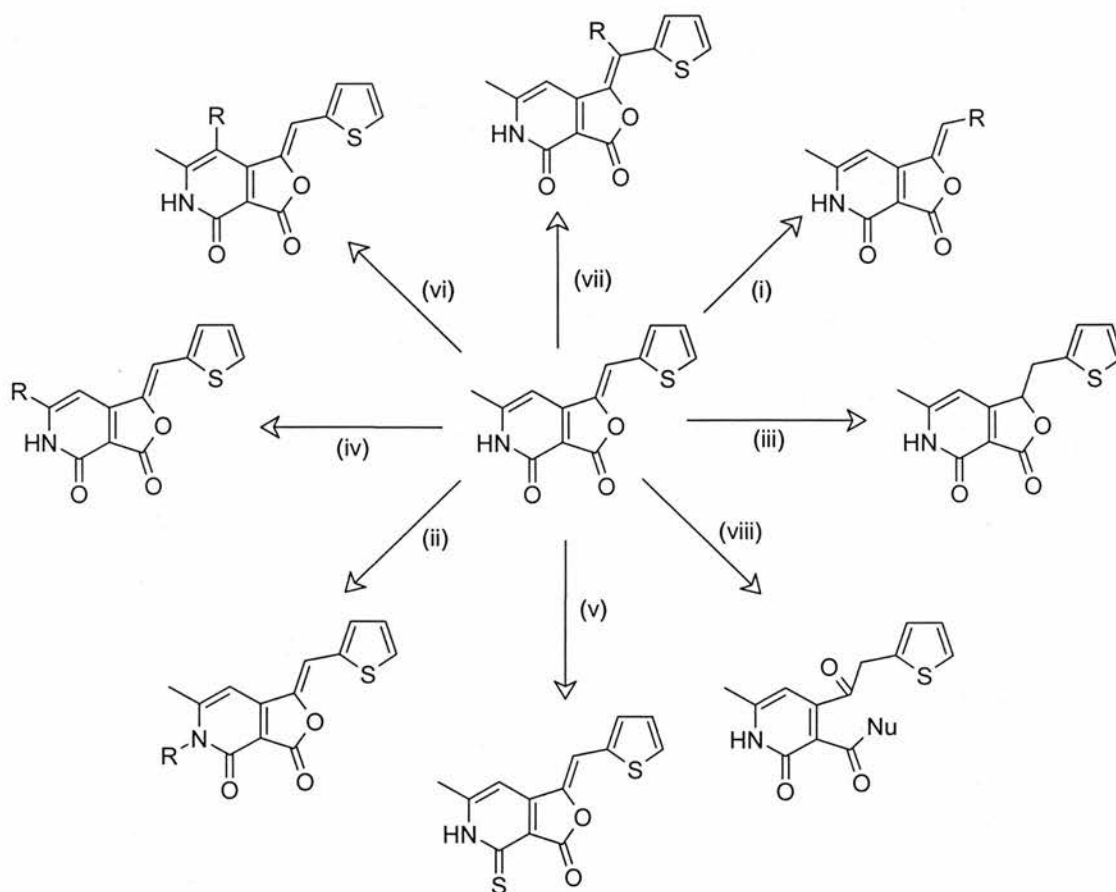


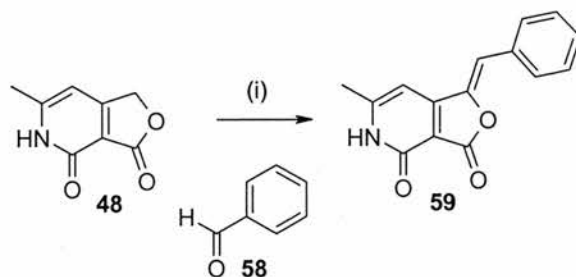
Figure 35. Possible derivitisation transformations.

These areas of modification lead to the proposal of the transformations as outlined in Fig. 35 and are discussed in more detail in the following sections.

2.2.1 Changing the aromatic group

The final step in the synthesis of **44** is a Knöevenagel-type condensation between furanone **48** and thiophene carboxaldehyde (**53**) (scheme 4). Modification of this procedure through the use of different aldehydes would be expected to lead rapidly to compounds with altered aromatic rings. Benzaldehyde (**58**) was selected as an appropriate initial aldehyde with which to investigate such alterations. It was reasoned that if compound **59**, the product of the reaction of **48** with benzaldehyde (**58**) (scheme 4), retained activity then this would open up the possibility of synthesising a series of compounds with different substituents on the benzene ring. Due to the large number of substituted benzaldehyde derivatives that are

commercially available, the chances of easily preparing an analogue of **44** with similar potency to **44** and that contains a linker unit would be increased in this series.



Scheme 4. The synthesis of 1-benzylidene-6-methylfuro[3,4-c]pyridine-3,4(*1H,5H*)-dione (**59**). (i) piperidine, methanol, reflux, 58%.

The yield of the condensation reaction to form **59** was comparable to that for the formation of **44**. Precipitation of the product **59** from the reaction mixture occurred in an analogous manner, therefore clean product was obtained through filtration and washing. The purity was confirmed by elemental analysis. This insolubility of **59** is mostly likely due to the retention of the pyridone unit and the ability of **59** to π -stack in the solid phase. Ease of isolation and purification is important in the preparation of derivatives and it was felt that this observation could influence the method chosen to prepare a small library of analogues of **44** (see section 3.3).

Compound **59** was tested in the invasion assay (described in section 1.8) and pleasingly was found to enhance invasion. The LAC of **59** was lower than that of **44**, 6.4 μM versus 1.69 μM . However, the activity was of the same order of magnitude and was therefore considered encouraging. It was concluded that the synthesis of further derivatives with modifications in this position could yield potent and useful analogues (see chapter 3).

Having identified a position which tolerated modification the primary aim of investigations into the SAR of **44** had already been achieved. In fact without any further SAR related investigations it would be possible to initiate the synthesis of a small library, to improve potency, and efforts towards a compound for target ID. However compounds which have been rendered inactive by their structural modifications have a use in control experiments (see section 4.2.2). Therefore the SAR investigations were continued.

2.2.2 *N*-alkylation

Often the alkylation of a nitrogen in a bioactive compound results in loss of activity, due to the disruption of essential hydrogen bonds.³⁷ Alkylation of the nitrogen of the pyridone ring in **44** and testing of the resulting derivative **60** would therefore indicate whether or not the NH group is essential for biological activity. With amide groups such as the one in the pyridone ring of **44** alkylation can occur in one of two possible positions: either the nitrogen will be alkylated to yield compounds of the general structure **61** or the oxygen will be alkylated to yield compounds of general structure **62**. The selectivity of the alkylation will depend on the relative hardness and softness of the nucleophile and electrophile that are used and on factors that include the nature of the nature of a base and reaction solvent.^{160,165,166}

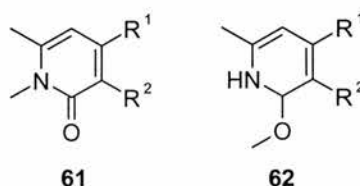


Figure 36. Possible positions for the alkylation of pyridones.

A nitrogen-based anion is typically viewed as a softer nucleophile than an oxygen-based anion, so to ensure *N*-alkylation a soft electrophile should be employed. Methyl iodide is a soft electrophile, so would be expected to allow selective alkylation of the nitrogen.

2.2.2.1 *N*-methylation

Methylation is a method used when determining the importance of a potential hydrogen bond to a heteroatom on the biological activity of a compound. Methyl groups are the least bulky alkylating group so assessing the effect of *N*-methylation on the activity of **44** would more directly correlate to the lack of a proton on the nitrogen (and hence ability to act as a H-bond donor) as opposed to the addition of a bulky group causing steric hinderance.

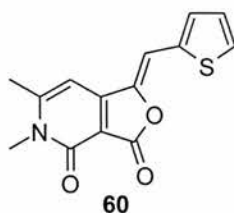
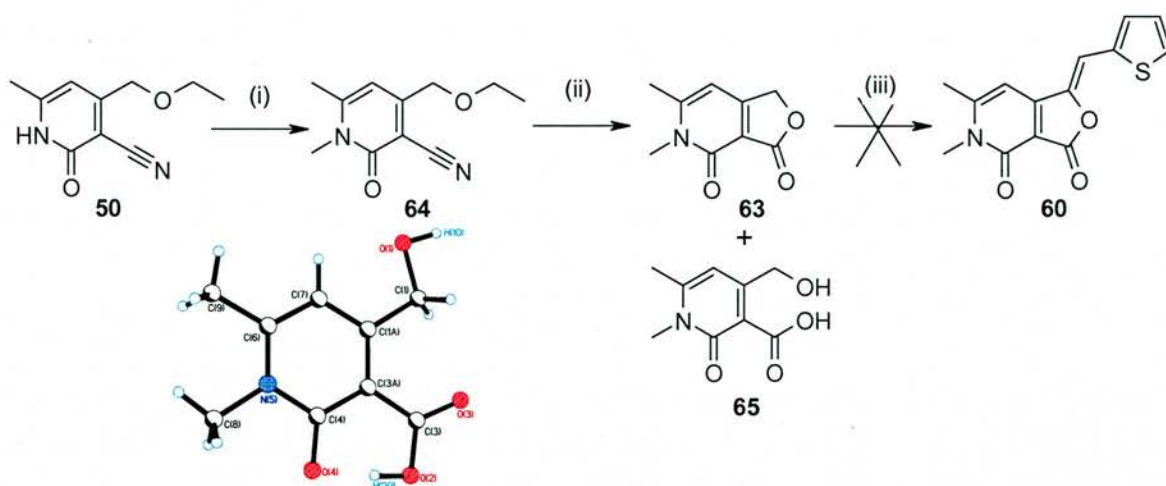


Figure 37. *N*-methylated derivative of **44**.

Two approaches to synthesising a *N*-methylated version of **44**, were adopted. First, an attempt was made to prepare a *N*-methylated precursor to **60**, the *N*-methylated furanone **63** (Scheme 5). Second an attempt was made to methylate **44** directly. The advantage of the first approach is that once **63** has been synthesised it can be used in conjunction with other derivatisation approaches (for an example see section 2.2.1). Initial experiments therefore centred on the synthesis of **63**.

2.2.2.1.1 A route to an *N*-methylated derivative of **44** via **63**.

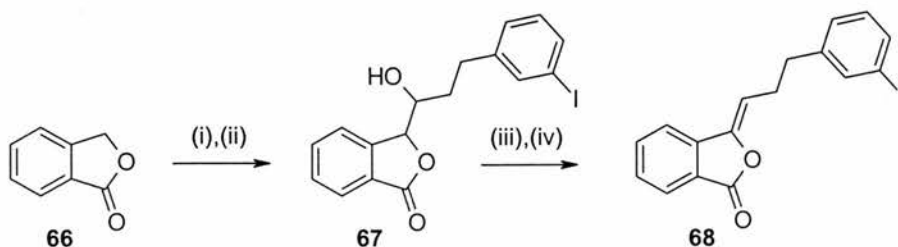
The proposed synthetic route to **60** through *N*-methylated precursors involves synthesis of the *N*-methylated pyridone **64** followed by ring closure to yield the furanone **63**, under analogous conditions to those to prepare **60**. The condensation of **63** with thiophene carboxyaldehyde (**53**) was expected to yield **60**. The synthesis of the *N*-methylated pyridone **64** was achieved in good yield using a literature procedure involving reaction with iodomethane and potassium hydroxide as the base.¹⁶⁷ The presence of a signal assigned to the pyridone carbonyl group at 1653cm^{-1} in the IR spectrum of **64** confirms that methylation has occurred on the nitrogen as opposed to the oxygen of the pyridone, as expected.



Scheme 5. i) CH_3I , KOH , ethanol, water, reflux, 93%, (ii) 50% aq. H_2SO_4 , 125°C , **63** (5% isolated yield), **65** (34% isolated yield), (iii) piperidine, ethanol, reflux,. CCDC No. 292090

Ring closure to yield the furanone **63** however proved difficult. Although **63** was produced, **65** was the major product of the reaction. Confirmation that *N*-methylation (as opposed to *O*-methylation) had occurred was obtained through an X-ray crystallographic analysis of **65**. Compound **65** however could be converted to **63** with tosic acid and tetrahydrofuran heating over molecular sieves. One possible reason for this difference in ability to form the furanone compound could be due to the change in the electronics of the pyridone. The electron donating property of the methyl group may, due to the delocalisation around the pyridone ring, cause the carbon of the nitrile group to be less electrophilic thus decreasing the likelihood of nucleophilic attack onto the carbon to form eventually the furanone. Alternatively the differences could result from an increased ease of facilitation of ring opening of **63** compared to **48**, by altering the solubility, thus making **63** more susceptible to attack by a nucleophile. The IR stretch assigned to the furanone carbonyl of **63** is at 1707 cm^{-1} , comparing that with stretch assigned to the same carbonyl for **48** which is at 1761 cm^{-1} indicates that the furanone of **63** is less reactive towards nucleophilic attack. This perhaps argues against the ring opening of the furanone causing the formation of **65**.

Condensation of **63** with thiophene carboxaldehyde (**53**) was expected to yield **60**. However, the increased solubility in organic solvents of **60** compared to **44** appears to affect the Knöevenagel-type condensation. None of the desired product precipitated from the solution and attempts to isolate any product or related material proved unsuccessful.



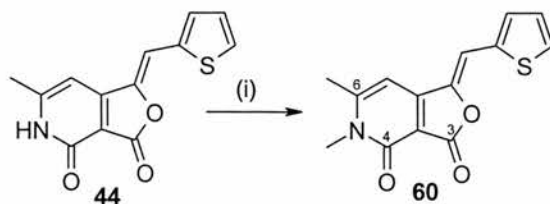
Scheme 6. Two step condensation reaction involving phthalide and LDA.¹⁶⁸ (i) LDA, THF, -50 °C, (ii) -78 °C, 3-(1-(2-iodobenzyl)propyl)aldehyde, 91%, (iii) MsCl, TEA, DCM, 0 °C, 98%, (iv) DBU, THF, 100%.

Alternative condensation procedures involving the use of lithium diisopropylamide (LDA) at low temperatures also proved unsuccessful. These procedures have been used in the literature to effect the condensation of an aldehyde with **66** (Scheme 6).¹⁶⁸ The intermediate compound **67** was isolated and then the condensation is affected through the conversion of the OH to a good leaving group (mesyl).

In summary this approach to **60** was unsuccessful due to difficulties with the condensation reaction and direct methylation of **44** was therefore pursued as an alternative approach.

2.2.2.1.2 Direct methylation of compound **44**.

Methylation of **44** was achieved through the standard procedure of MeI, K₂CO₃ in DMF (scheme 7). This yielded the *N*-methylated compound **60** in good yield. The procedure used for the methylation was expected to result in *N*-methylation rather than *O*-methylation. The IR spectrum of **60** obtained by this route is consistent with *N*-methylation due to the presence of a shift at 1651cm⁻¹ corresponding to the carbonyl functional group (a signal at 1778 cm⁻¹ assigned to the carbonyl group of the furanone was also present).



Scheme 7. *N*-Methylation of compound **44**. (i) MeI, K₂CO₃, DMF, 59%.

HMBC analysis was expected to be able to confirm the position of the methyl group, however, the carbon signal for C6 overlaps with the signal for C3a so a correlation between these carbons and the protons of the CH₃ group cannot provide any information. The strong ¹H-¹⁵N HMBC correlation between the nitrogen and the CH₃ at 3.48 ppm suggested that **44** was alkylated on nitrogen. If **44** was *O*-alkylated, a correlation would still be expected but it would not be expected to be as strong (as this would be 4 bond coupling as opposed to 2 bond coupling). Although this information in itself is not conclusive, used in conjunction with other information it helps confirm the structure of **60**.

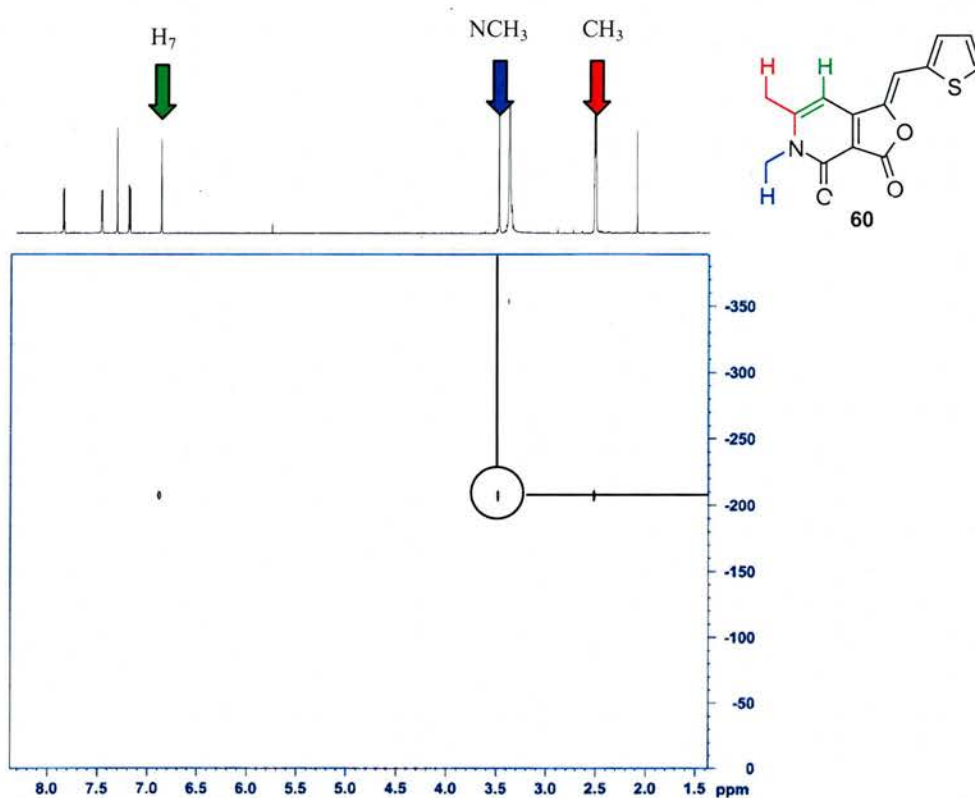


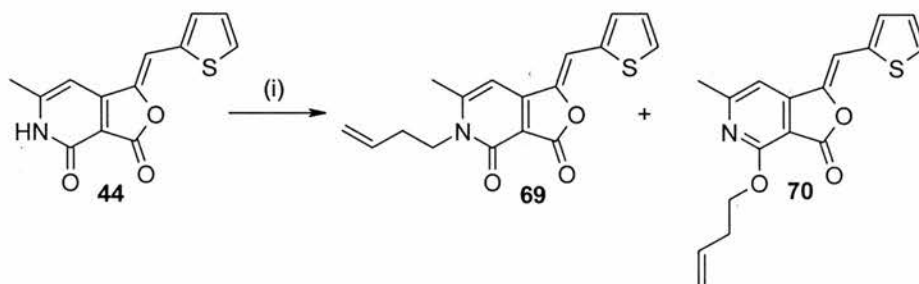
Figure 38. ¹H-¹⁵N HMBC of compound **60**. The HMBC shows the presence of through bond correlation between the nitrogen and both CH₃ groups and H₇. Given the correlation to the CH₃ group at 3.48 ppm it is likely that the compound is *N*-alkylated. However correlation could still be observed in the case of *O*-alkylation although it would be a 4 bond coupling as opposed to a 2 bond coupling.

The activity of **60** against *T. gondii* invasion was determined. Unsurprisingly compound **60** showed a significant drop in potency compared with **44** (25 μM compared with 1.69 μM). Interestingly the activity was not lost entirely. This was

interpreted as an indication that a hydrogen bond to the nitrogen aids binding but is not essential for biological activity.

2.2.3.2 *N*-Butenylation

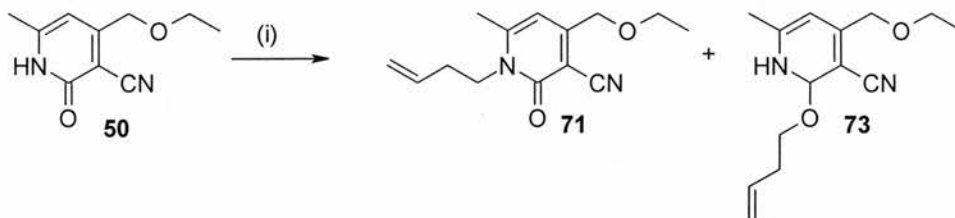
N-Methylation of **44** yields the more soluble compound **60**. The increased solubility could be advantageous for the derivatisation reactions if solvents other than DMF need to be used, (e.g. if extremely water sensitive reactions or reactions with halogenating reagents etc. were required). Additionally, alkylation of the nitrogen would also aid purification as the increased solubility would facilitate chromatographic separations. Removing the alkylating group after the purification or sensitive reaction would then be expected to yield the desired and probably insoluble pyridone compound which could be isolated by filtration followed by washing. The Westwood group has recently shown that the butenyl protecting group has an extended synthetic range compared to that originally reported by Barrett.¹⁶⁹ In this context we therefore attempted *N*-butenylation of **44** to yield **69**.



Scheme 8. Butenylation of compound **44**. (i) 4-Bromobut-1-ene, K_2CO_3 , DMF, **69** (15% isolated yield), **70** (4% isolated yield).

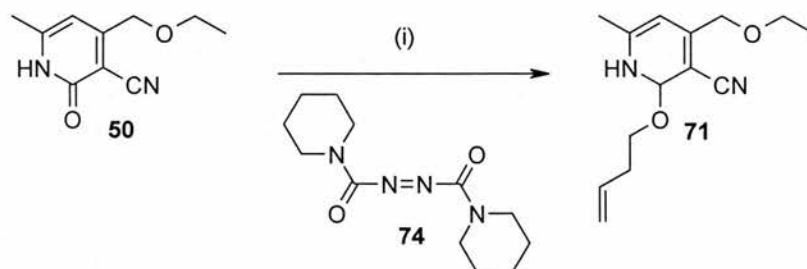
The reaction conditions for butenylation of **44** include the use of a bromide as opposed to an iodide. 4-Bromobut-1-ene is considered to be borderline soft/hard electrophile and therefore a reduction in *N/O*-alkylation selectivity was expected. This proved to be the case as the butenylation reaction of **44** gave both *N* and *O* alkylated products, **69** and **70** respectively. The yield of this reaction was low. Attempts to improve the yield through the addition of KI did not significantly improve the yield.

As discussed in section 2.2.2.1.1, another method of accessing **69** could be through the corresponding *N*-butenylated pyridone **71**. Butenylation of **50** using 4-bromobut-1-ene (**72**) under analogous conditions to those used for **44** also led to both *N* and *O*-alkylation.



Scheme 9. *N* and *O*-alkylation of pyridone **50**. (i) 4-Bromobut-1-ene, KOH, CH₃CN, **71** (38% isolated yield), **73** (33% isolated yield).

Selective butenylation of **50** proved to be possible through use of a modification of the Mitsunobu reaction using the diazadicarboxylate ADDP (**74**).¹⁷⁰ This was intended to produce the *N*-butenylated derivative of **50**, **71**. However, in this case only the *O*-alkylated compound **73** was observed. Selected *O*-alkylation has been observed in the literature.¹⁷¹

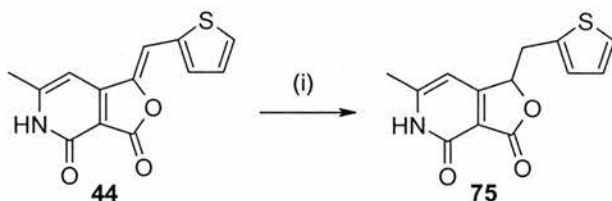


Scheme 10. Selective *N*-alkylation of pyridone **50**. (i) 3-Buten-1-ol, ADDP (**74**), TBP, THF, 62%.

The Mitsunobu reaction was investigated as a possible method of selective alkylation of **44**. However due to the relatively low solubility of **44** in THF the reaction was attempted in DMF. There is limited literature precedent for this reaction in DMF¹⁷² and unfortunately the analogous reaction conditions in DMF were unsuccessful.

2.2.3 Reduction of the exocyclic double bond in 44.

It was envisaged that the exocyclic double bond of **44** contributes significantly to the chemical reactivity and physical properties of **44**. This double bond is in conjugation with pyridone ring and therefore provides a link between the pyridone ring and the thiophene unit. This extended conjugation is responsible for the observed colour of **44**, (furanone **48** is a white solid and **44** is a yellow solid) as well as its fluorescent property. Conjugation also causes **44** to be planar, a fact that may be crucial to the interaction of the compound with its cellular target. Finally, the exocyclic double bond forms part of a strained furanone, the increased strain in the furanone and the modification of the oxygen atom into a better leaving group meaning that it is more susceptible to ring opening with nucleophiles. Given the dramatic effect this double bond was thought to enforce on the chemistry of the compound, reducing the double bond to give **75** and assaying **75** was expected to provide interesting biological and chemical information. It was proposed that reduction of the double bond could be achieved through hydrogenation in the presence of palladium on carbon.¹⁷³



Scheme 11. Reduction of the exocyclic double bond. (i) H₂, Pd/C, DMF, 56%.

Compound **75** was successfully synthesised and was found to be a white solid. Although there appears to be a slight increase in solubility, the effect is not dramatic, suggesting that the insolubility of these types of compounds is largely a factor of the pyridone unit.

Compound **75** was assayed against *T. gondii* invasion and was found to be completely *inactive* at 100 μM. This lack of activity can be rationalised by at least one of two possible reasons; reducing the double bond has (i) destroyed the planarity of the compound and (ii) reduced the reactivity of the furanone towards modification by protein nucleophiles. Disrupting the planarity of the compound could stop the compound from binding in a tight pocket. Additionally a degree of flexibility, due to

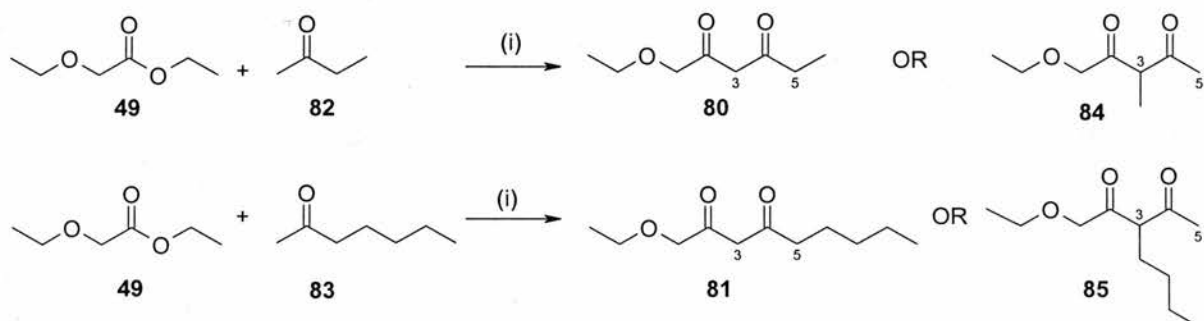
the now free rotation around the CHCH₂ bond, has been added to the compound so causing a possible additional entropic cost to binding. A chemist examining the structure of **44** in terms of possible mechanisms of biological activity, would suggest that the strained furanone could be opened by a nucleophilic residue on a protein or conjugate addition could occur to C8. This would lead to covalent modification of the protein. Whilst this would appear to be inconsistent with the observed reversibility of the action of **44** it may be possible to have a covalent protein modifier with a reversible biological activity.¹⁷⁴ However, the inactivity caused by the reduction of the double bond would add support to the argument that the double bond attached to the furanone in **44** is essential for its biological activity. The possibility that **44** exhibits its biological activity through ring opening of the furanone will be discussed in more detail in chapter 5.

Compound **75** is racemic, as a stereogenic centre has been formed in the reduction reaction and no reagents were used to apply selectivity towards a particular enantiomer. Separating enantiomers of compounds becomes particularly important with biologically active compounds, as different enantiomers may have different activities. This can be achieved through chemical separations (eg. addition of homochiral groups to create diastereoisomers that can be easily separated), or the use of chromatography (eg. chiral HPLC). However as **75** was inactive, obtaining pure samples of each enantiomer of **75** was not pursued.

2.2.4 Extension of CH₃

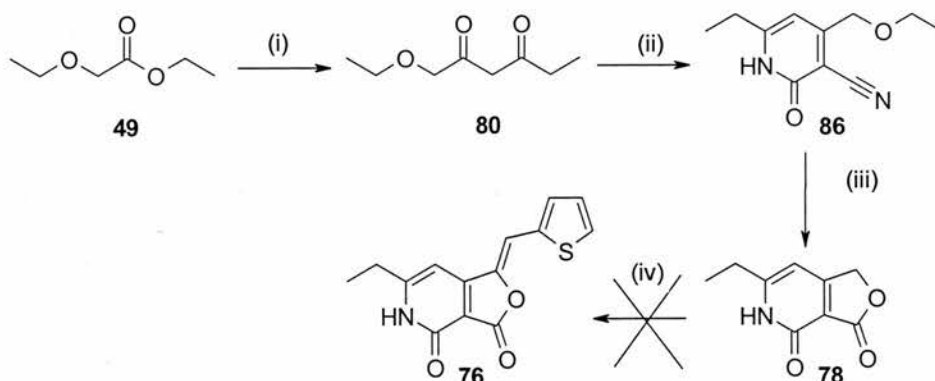
In order to investigate the amount of space available around the methyl group in compound **44**, the two compounds **76** and **77** (scheme 13 and 14), were selected as synthetic targets. The approach was taken of accessing modified furanones **78** and **79** in order to allow for the synthesis of multiple analogues, by also altering the aromatic group (see section 2.2.1). Additionally modifications of **44** are limited by the solubility of the compounds as the only viable reaction solvents are DMF and DMSO. It was proposed that incorporation of a more lipophilic substituent at the C6 position may counter this problem.

Obtaining the furanones **78** and **79** required the synthesis of the diketones **80** and **81** respectively. These diketones can be synthesised by the reaction of ethylethoxyacetate (**49**) with butan-2-one (**82**) and heptan-2-one (**83**) respectively. There are two possible products of the reaction of ethyl ethylethoxyacetate (**49**) with these diketones. The extended alkyl chain can either be in position 5 (e.g. **80** and **81**) or position 3 (e.g. **84** and **85**). Formation of the sodium enolate of the ketones **82** and **83** controls the product of the reaction.



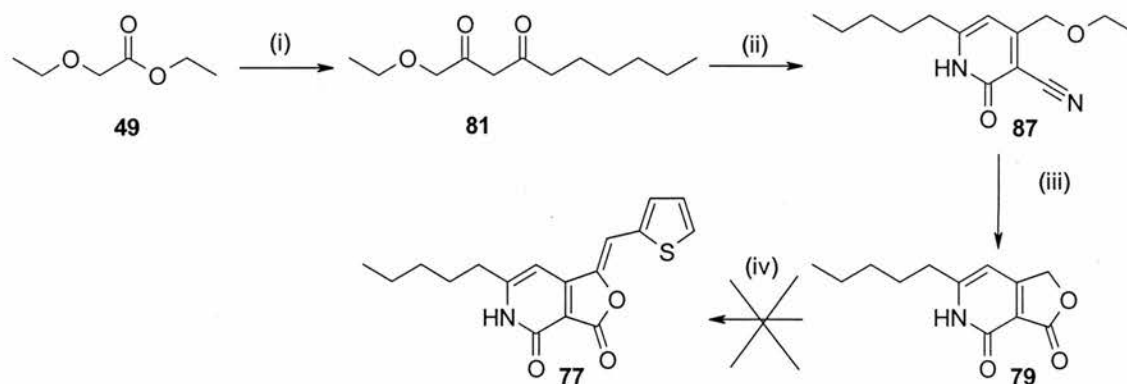
Scheme 12. The possible products of the reaction of enolates of **82** and **83** with **49**. (i) NaH, toluene, -10°C , 34% (**80**), 58% (**81**).

Although the enolates formed may not be of only one character, the formation of one product indicates that the enolate with the terminal double bond (position 1-2) is the enolate the compounds react through. Only the formation of **80** and **81** via the kinetic enolate was observed. Although the relatively low yields observed means that other products could be formed, no formation of **84** and **85** was observed.¹⁷⁵ The formation of **80** is consistent with the literature synthesis of this compound. ^1H NMR analysis confirmed the absence of an additional CH_3 signal in position 1 which would be expected for compounds **84** and **85**.



Scheme 13. Synthesis towards the 6-ethyl derivative of **44**. (i) NaH, toluene, 34%, (ii) cyanoacetamide, piperidine, 60°C , 34%, (iii) 50% H_2SO_4 , 125°C , 20%, (iv) piperidine, ethanol, reflux.

The yields for the synthesis of the diketones **80** and **81**, were comparable to the yield for the synthesis of **51**. The synthesis of the pyridone derivatives **86** and **87** were successful with the expected increase in solubility of **86** and **87** requiring that the reaction solutions be concentrated slightly to facilitate precipitation of the desired products. However the yields for the formation of **86** and **87** were poor compared to that for **50** due to the added difficulty in product isolation.



Scheme 14. Synthesis towards the 6-pentyl derivative of **44** (i) NaH, Toluene, 58%, (ii) cyanoacetamide, piperidine, 60°C, 22%, (iii) 50% H₂SO₄, 125°C, 22% (iv) piperidine, ethanol, reflux.

The increased solubility associated with **86** and **87** was also observed in the reaction to form the furanones **78** and **79** and meant that precipitation of the desired products did not occur and the furanones had to be isolated from the acidic solution through extraction. The yield of these transformations were low.

Unfortunately the condensation reaction did not work with furanones **78** and **79** when analogous conditions for the conversion of **48** to **44** were used. This is probably due to the increase in solubility of **78** and **79** which appears to be prohibitive to the condensation reaction, as found in the case of the *N*-methylated derivative (section 2.2.2).

Given that compounds that retained potency and compounds that were inactive had already been identified further attempts to prepare **76** and **77** were considered to be unnecessary at this point.

2.2.5 Replacement of the pyridone oxygen for a sulfur

A literature search for known biological activities of structures related to **44** identified derivatives of **44** with a sulfur atom in position 4. Consequently compounds **88**, **89** and **90** were purchased and assayed against *T. gondii* invasion in order to determine if these compounds were active in our biological system.

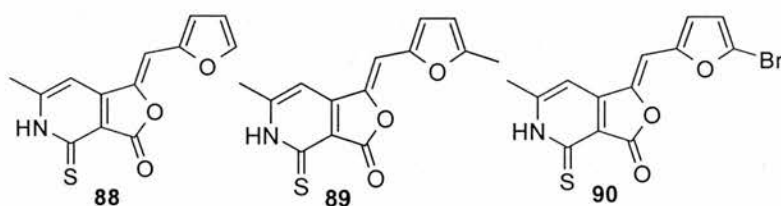


Figure 39. Compounds related to **44** with known biological activity.

The LACs of **88**, **89** and **90** are 25 μM , 3.1 μM and 6.25 μM respectively (values not solubility corrected). This retention in activity could suggest that they share the same cellular target (see chapter 1 for discussion surrounding this). Additionally the mechanism of action of these compounds could be the same for both systems (see chapter 5 for a discussion surrounding mechanism).

2.2.8 SAR Conclusions.

The studies into the structure activity profile of **44** have shown that the aromatic ring can be substituted without a loss of activity. Alkylating the pyridone nitrogen reduced the biological activity possibly implicating the involvement of the NH in a form of non-essential hydrogen bonding. Reduction of the exocyclic double bond caused a complete loss of activity, suggesting covalent modification of **44** by a protein(s) as a possible mechanism of action. Analogues with sulfur position 4 also retained activity, providing a possible link with compounds with an established biological target.

Having identified a position suitable for modification for the purposes of optimisation and target identification studies and also an inactive analogue of **44**, sufficient information had been gathered to allow the synthesis of a small library of compounds and to direct synthesis towards target identification studies. It was decided to pursue both of these avenues in parallel as it was believed that SAR data

generated through the synthesis and testing of a compound library would be of great value once potential protein targets had been identified and protein target validation studies were initiated.

2.3 Additional chemical reactivity associated with structures of type 44.

While investigating the chemical properties of **44** several additional reactions were identified that were judged to be of interest. One reaction in particular has the potential to impact on discussions surrounding the mechanism of action of **44** (Chapter 5).

2.3.1 Reaction of the furanone moiety with amines

The susceptibility of the strained furanone in **44** to nucleophilic attack has been discussed previously (section 2.2.3). Structural modifications of **44** through ring opening with a nucleophile were therefore examined. Butylamine was reacted with **44** and it was expected that ring opened compounds such as **91** or the lactam derivative **92** would be formed.

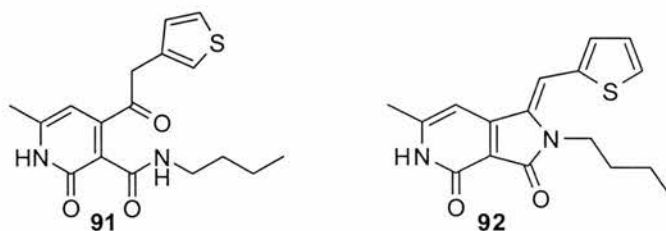
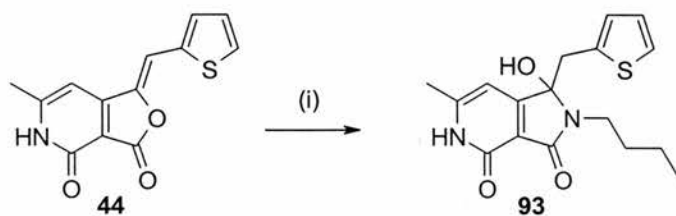


Figure 40. Two possible products which could be formed by the reaction of **44** with butylamine.

However ^1H NMR analysis of the single product formed on reaction of **44** with butylamine clearly showed the presence of diastereotopic protons which indicated the presence of a chiral centre. For example the signals corresponding to the NCH_2 appeared as two multiplets (3.43-3.10 ppm) due to the effect of the chiral centre.¹⁷⁶ Structures **91** or **92** do not contain a chiral centre and were therefore not consistent with the analytical data.

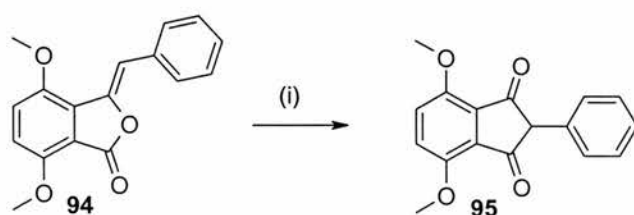


Scheme 15. The ring opening of **44** with butylamine. (i) Butylamine, DMF, room temperature, 28%.

The IR spectrum showed a carbonyl stretch at 1707 cm^{-1} , compared with 1790 cm^{-1} for **44**, this is indicative of the alteration of the furanone to a lactam functionality. In addition, mass spectral analysis showed that the molecular weight of the product was 332, ($355\text{ (M + Na}^+)$), ruling out **91** as a possible structure. Structure **93** is consistent with all the analytical data available at present.

2.3.2 Reaction of furanone in **44** with alkoxides

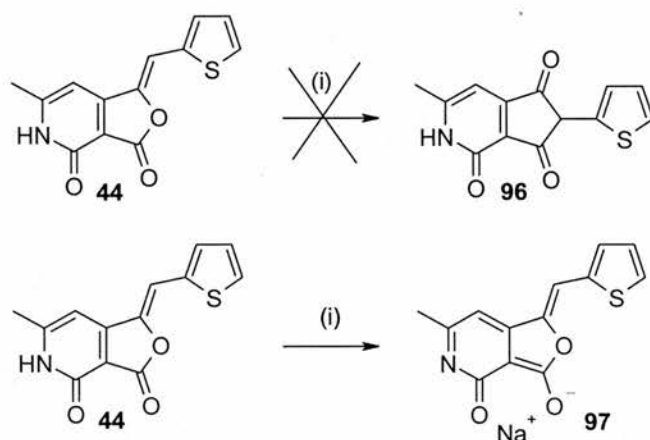
Literature precedent exists for the rearrangement of compounds containing strained furanones, (e.g. **94**), to yield diketone compounds such as **95**.¹⁷⁷ This rearrangement is achieved by refluxing **94** with sodium methoxide in methanol and presumably occurs via attack of the methoxide onto the furanone carbonyl (C3) causing ring opening of the furanone. Ring closing is then initiated by attack of the resulting enolate through the enolate carbon in position 8 leading to **95**.



Scheme 16. Literature precedent for rearrangement of strained furanone compounds to yield a diketone. (i) NaOMe, methanol, 90%.¹⁷⁷

In an attempt to carry out an analogous reaction, **44** was refluxed in methanol in the presence of NaOMe. Due to the insoluble nature of **44**, a heterogenous reaction mixture was formed. After refluxing for 2 hours the solid precipitate was filtered. Unfortunately, on analysis it was shown that the filtered solid was not **96** as intended, as indicated by the lack of two carbon signals corresponding to lower field carbonyl

carbons in the ^{13}C NMR spectrum. However there appeared to be an alteration in both the carbon and proton spectra suggesting that the starting material **44** was no longer present. When compared to the ^1H NMR spectrum of **44**, upfield shifts in the region of 0.3-0.5 ppm were observed for all of the signals of the product. The carbon spectrum also showed a shift of 5 ppm downfield for the majority of the signals. However larger shifts were observed for C6 and C3, in the region of 10 ppm. ^{15}N NMR of compounds **44** and **97** showed a shift in the signals assigned to the nitrogen atom from -203.6 to -108.1. This observed change was consistent with a change from a secondary amide nitrogen to a tertiary nitrogen and was interpreted as being due to the formation of **97** in this reaction.



Scheme 17. Reaction of compound **44** with sodium methoxide. (i) NaOMe, methanol, reflux, 2 hrs 35%.

Interestingly the change in carbon shift of the signal corresponding to the furanone carbonyl carbon (C3) in **44** in comparison to the lack of change in the shift for the signal corresponding to the pyridone carbonyl carbon (C4) indicates that the deprotonation does not only effect the pyridone but the entire conjugated system. **97** appears to exist as drawn as opposed to enolate of the pyridone carbonyl (C(4)O(4)). However, it is possible that this compound does not exist exclusively in one form and that a time-averaged view of the situation is being observed by ^1H NMR analysis.

In summary **44** has been shown to react with butylamine. This reaction has possible implications for the mechanism of action of these compounds (see chapter 5). Reacting **44** with methoxide did not result in the expected rearrangement reaction but led to formation of the sodium salt of **44**. This indicates that there may be some potential importance of the pyridone remaining protonated in order for **44** to react

successfully with nucleophiles, again this could have implications within the biological system. The typical pKa of a pyridone nitrogen is 11.7 this means that it can be deprotonated by arginine protein residues.

2.4 Chapter Summary

Compound **44** which was identified as active against *T. gondii* invasion has been synthesised and confirmed as the active component of the screen. Derivatives of **44** have been prepared and information obtained about structure activity relationships. This information will guide the selection of the compounds to be synthesised for the purposes of target identification and direct the studies towards compounds with increased potency.

Chapter 3: Parallel synthesis of derivatives

3.1 Library synthesis

The optimisation of compounds is a key component of the chemical genetics approach. In particular, compounds identified in forward chemical genetics screens initially have no protein target assigned to them. The “rational design” of compounds with improved potency is therefore not possible. Without the use of co-crystal structures or molecular modelling, the biological properties of compounds can best be improved through the synthesis and evaluation of targeted compound libraries. The previous chapter described the synthesis of individual compounds with the goal of defining a preliminary structure activity relationship for **44**. One main position was identified in **44** where modification could be achieved whilst retaining activity, the aryl group. This chapter focuses on the use of this information to prepare a series of analogues of **44** substituted at this position.

3.2 Available technology

Technological advancements in the field of high throughput chemistry allow for accelerated synthesis of large numbers of compounds.¹⁷⁸ This can shift potentially the slowest step of a forward chemical genetic programme away from the traditionally labour intensive process of analogue synthesis. There are several new technologies available to facilitate the majority of the steps involved in library synthesis, these range from: 1) enabling the synthesis of large numbers of compounds by: parallel synthesis, microwave synthesis, solid phase synthesis (the use of split and pool); 2) facilitating the dispensing of liquids or solutions using robotics; 3) reduction in the time consuming process of compound isolation through the use of high throughput purification techniques; 4) rapid analysis of samples through the use of automated analytical equipment.

3.2.1 Parallel Synthesis

Parallel synthesis techniques enable large numbers of compounds to be synthesised rapidly. Various systems are available ranging from manual to fully automated.¹⁷⁸ Reaction blocks, (**A**, Fig. 41), are manual systems, that allow for the synthesis of up to 96 compounds in one round. They are often supplied with filtration and separation devices, enabling the use of both solid and solution phase chemistry. Reaction blocks can be coupled with robotic equipment to form workstations with different degrees of automation, (**B**, Fig. 41). These systems can often carry out air and moisture sensitive reactions. High and low temperature reaction blocks are available as well as those which allow each separate reaction within a block to be carried out at different temperatures, so that within the same block one reaction can be held at $-30\text{ }^{\circ}\text{C}$ and another at $150\text{ }^{\circ}\text{C}$, allowing rapid optimisation of reaction temperatures. Devices of various volumes allow for reactions on the scale of anywhere from 5 mg to 5 g. The increasingly diverse range of equipment means that the technology can be selected to suit the chemistry and not the reverse.



Figure 41. Technology available to facilitate the synthesis of compound libraries. Images A-D illustrate examples of the technology used in library synthesis.

A web based search identified the following companies as manufacturers of technology required for library synthesis. The letters in brackets correspond to the type of instrumentation they supply. A = manual reaction blocks, B = automated systems, C = Microwaves, D = LCMS systems, E = HPLC systems, F = ELSD's, G = NMR systems. Accelab (B), Advanced ChemTech (B), Agilent Technologies (D, E), Alltech Associates (F), Applied Biosystems/MDS SCIEX (D), Argonaut Technologies (B, E), Biotage (E), Brucker Biospin (G), Buchi (A, E), CEM Corporation (C), Charbdis Technologies (A, B, E), Chemglass (A), Chemspeed (A, B), CSPS Pharmaceuticals (A), FlexChem

(A), GBC Scientific Equipment PTY Ltd (E), Genevac (A), Gilson Inc. (B, E), H&P (A), Hitachi (E), Heidolph (A), HEL Inc (B), Innolabtech (B), Isco Inc (E), J-Kem Scientific (A), Mettler Toledo (Bohdan) (A, B), Milestone Inc. (C), Orochem Technologies (A), Personal Chemistry/Biotage (C), Polymer Laboratories (F), Prolab Instruments (B, E), Radleys (A, E), Rapp Polymere (A, B), Robosynthon (A), Sedere (F), Shimadzu (D, E), Stem corporation (A), Spyder Instruments (A), Tecan (B), Torviq (A), Varian Inc. (D, E, G), Waters (D, E) and Zinsser Analytic (A, B).

3.2.2 Microwaves

Microwave technology is of increasing importance in chemistry. Microwaves have been shown to increase the rate of reactions dramatically reducing reaction times from hours to minutes.¹⁷⁹ There remain questions over the nature of the improvement in the reaction rate but there appears to be no doubt of its effectiveness.¹⁸⁰ Machines which enable individual microwave reactions (e.g. C Fig. 41) are available and when coupled with a robotic arm, they can function in a high throughput capacity. When coupled with decreased reaction times this enables large numbers of compounds to be synthesised in a shorter time than would otherwise be possible. Microwave systems that allow for 24 reactions in parallel are also available, thus combining the parallel synthesis and microwave technologies.

3.2.3 Solid phase synthesis

Parallel synthesis is one method of accessing large numbers of compounds rapidly. The number of different compounds which can be accessed however is dictated by the number of reactions steps. Solid phase synthesis is the application of synthetic chemistry on substrates which have been immobilised on a synthesis resin. This technique allows for the use of excess reagents to encourage complete reaction, as washing the resins will remove any unwanted material. Cleavage of the compounds from the resin yields the final product without the need for any further purification, provided the chemistry is sufficiently well optimised. Split and pool synthesis is a technique developed to maximise the productiveness of solid phase chemistry (Fig. 42). This technique removes the one-to-one relationship that exists in parallel synthesis where the number of reactions set up and number of products achieved is equal and thus split and pool synthesis can increase the overall productivity.

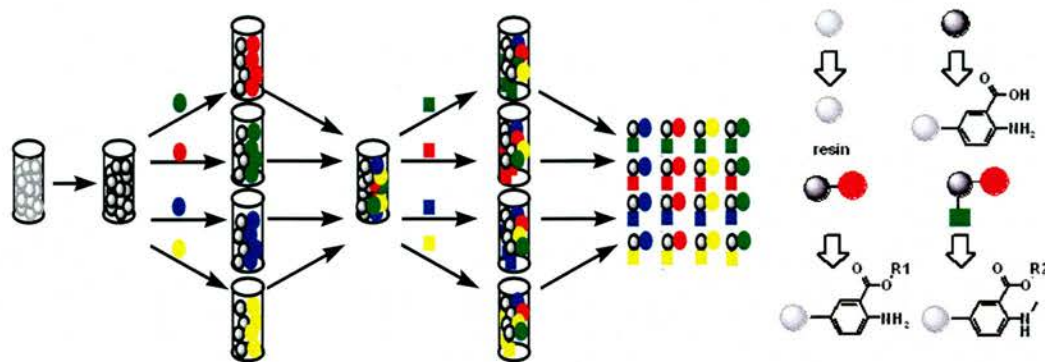


Figure 42. The technique of split and pool synthesis. Split and pool allows for the synthesis of large numbers of different compounds (in smaller quantities) in a minimal number of steps. Beads are loaded with a core structure, they are then divided into small pools, e.g. 4. In one example, each pool is loaded with a different R1 group using analogous chemistry. The beads are then recombined. The beads are then split again, into 4 in this case, and a different R2 group added to each of the four pools of beads. The combination of these two events gives $4 \times 4 = 16$ different compounds in only 8 chemical reactions (+ loading and cleavage). This process is repeated to give more combinations.

Two leading companies in solid phase synthesis technology are IRORI (Discovery Partner International) and Mimotopes. These companies have developed methods of pooling and tracking batches of resin, to facilitate their use in the context of split and pool synthesis.²²

3.2.4 Automated purification

Unlike most occasions when solid phase synthesis is used, solution phase parallel synthesis typically involves a component of purification after the synthesis of the compound. Synthesis of large numbers of compounds rapidly is less feasible when these compounds are to be purified individually. To take the full advantage of parallel synthesis systems, automated purification devices are essential.¹⁸¹ Automated injection and collection systems can be coupled to HPLC systems. The technology now exists to allow purification by chromatography of up to 10 compounds in parallel. Reaction solutions can therefore be removed from the reaction vessels and injected directly into a HPLC through the use of robotics yielding clean products rapidly. Purification can be directed by UV, mass spectrometry or ELS. Reactions

where the product is insoluble under the reaction conditions can lead to isolation of the product by filtration using available filtration units.

3.2.5 High Throughput compound analysis

Analysis of compounds is a crucial component in organic synthesis. Technologies, such as liquid chromatography-mass spectrometry (LCMS, Fig. 41, **D**) provide along with a chromatogram, mass spectra, and when coupled to a diode-array detector, UV spectra. These automated machines can inject samples 24 hours a day and enable compound purity and molecular weight to be assessed. LCMS is therefore a technique commonly used to analyse large compound libraries, although the restricted information (no information with regards to atom connectivity) has the potential to lead to ill-informed conclusions (see section 3.7).¹⁸²

Automation of NMR is made possible through the use of sample changers. New technology which is particularly relevant to library synthesis is flow NMR, which allows for samples to flow in and out of an analysis cell within the NMR probe. These machines are automated and allow NMR spectra to be generated directly from a library of compounds contained in plates. MS/MS systems can also be combined with this system. This means that samples can be analysed both by NMR and mass spectrometry, generating information as to the molecular weight and structure of the compound, as well as an indication of purity, all in an automated fashion.

In an ideal world, the advances in technology described here shift the rate determining step in analogue synthesis away from practical issues and back to fundamental questions of chemical design and feasibility. Whilst this technology enables the synthesis of an increasing large number of compounds, it is often necessary to begin a library synthesis programme by preparing small focused libraries. This is a good method of determining both the chemical and biological viability of the library in question and also allows for the identification of potential problems which may arise ahead of a large scale library synthesis.

3.3 Selection of the reaction for library synthesis

The structure activity relationship data associated with **44** (Chapter 2) identified derivatives which retained activity, including **59** which was shown to have activity approximately equivalent to **44**. This flexibility in the aromatic group can be exploited through the use of different aldehydes to explore this region of the compound and potentially access compounds with increased potency. There are a large numbers of commercially available aromatic aldehydes that should allow the synthesis of libraries of various sizes.



Figure 43. The structures of compounds **44**, **59** and **48**.

Use of the Knöevenagel condensation reaction of **48** with an aromatic aldehyde also has an advantage from a synthetic perspective. Products from the condensation reaction are insoluble under the reaction conditions, suggesting that purification could be possible with filtration and washing alone. The ease of this purification method was attractive, particularly given the automated synthesis equipment available in the Westwood laboratory at the time of this research. The Büchi syncore (section 3.2.1, Fig. 41 A) enables twenty four reactions to be set up in parallel. Filtration equipment is also available with the syncore to facilitate the removal of the reaction solution from each reaction vessel, allowing any precipitate formed to be isolated and washed relatively easily. It was therefore decided to prepare a small focused library of analogues of **44** using this technology. The size of the reaction vessels in the Büchi syncore enabled the reactions to be performed on a scale large enough to allow for further chemistry or analysis to be made on all 24 products.

3.4 Adaptation of the reaction protocol for use in parallel synthesis.

Synthetic procedures can rarely be taken directly from individual reactions in solution and used in parallel synthesis equipment. Although the technology is available to allow the reactions to be performed under inert gases and in dry conditions not all equipment facilitates this. The selection of the equipment is usually governed by other criteria such as scale, purification methods and availability. The Büchi syncore was selected based on availability, the chosen purification method, the desired scale of the reactions and the size of the library produced. In general, this equipment is less amenable to running reactions that require dry conditions. Therefore the water sensitivity of the reaction had to be investigated. Additionally the practical issues surrounding the addition of different aldehydes to each reaction vessel led to the question: does the order of reagent addition dramatically affect the outcome of the reaction?

Establishing that the reaction is not significantly water sensitive was achieved by running the reaction without dry solvents. This is important as the dispensing of the reaction solvent through the Buchi filtration unit, a labour and time saving measure, is not amenable to the solvents remaining dry.

When the reactions were carried out using standard glassware, methanol was added to **48** in the flask. **48** did not begin to dissolve in the methanol until the reaction was heated to 60 °C, (even then **48** was not fully dissolved). The aldehyde and piperidine were then added at 60 °C. The difficulty of adding all 24 aldehydes at the same time in order to ensure the 24 reactions have the same reaction times, required a modification of this part of the procedure. Experiments showed that the reactions were still successful when **48** and the corresponding aldehydes were added to the reaction vessel first, followed by the methanol and the piperidine, and the reaction was then heated.

Once the reaction conditions required to enable the parallel use of the reaction were established, the next stage was to optimise both the scale of the reaction and the use of the equipment through the use of individual reactions in the syncore. The reaction of **48** with thiophene carboxaldehyde (**53**) was used to optimise the conditions. Reactions on a 0.9 mmol scale were shown to give amounts of material in the region of 100 mg. This scale provided sufficient material for further reactions to

be performed on the 24 different products, as well as full characterisation of all the compounds.

One of the problems encountered with the syncore was that the reactions often “bumped”. This resulted in the solid precipitate from the reaction collecting in the lid of the syncore, presenting the possibility of contamination of adjacent reactions. This problem was addressed in two ways. Although the syncore was programmed to shake the reaction vessels, insufficient agitation was proposed as the route of the problem. Glass rods or balls were placed in the reaction vessels. This appeared to aid mixing of the reaction mixtures and reduced the bumping.

The design of the syncore is such that the condenser unit is only in contact with ~18% of the reaction tube, whereas the heating element is in contact with ~35%. This is the inverse to the situation found with most conventional reflux set ups, where the surface area cooled by the condenser is over twice the area heated by the oil bath. Problems arose due to this low contact area when the vessels were heated at reflux, as in only a small fraction of the reactions did the condenser work effectively. This led to solvent loss and was believed to contribute to increased “bumping”. Tap water was initially used in the condenser, however, when this was substituted for silicon oil at 0 °C using a Hüber mini chiller, no solvent loss or bumping was observed.

3.5 The selected aldehydes.

Once all the practical problems had been resolved, the next task was selection of the aldehydes to be used. Previous experiments had established that only aromatic aldehydes could be used in this reaction, as the reaction of **48** with butyraldehyde proved unsuccessful. There are over 6000 aromatic aldehydes commercially available. Although no strict rationale was used to select the aldehydes the **24**, **53**, **58**, **98-119**, were chosen to include examples of: single aromatic rings with *ortho*, *para* and *meta* substituents, fused ring systems, heterocyclic aromatic rings and both electron withdrawing and donating substituents. The aldehydes used are shown in Fig. 44.

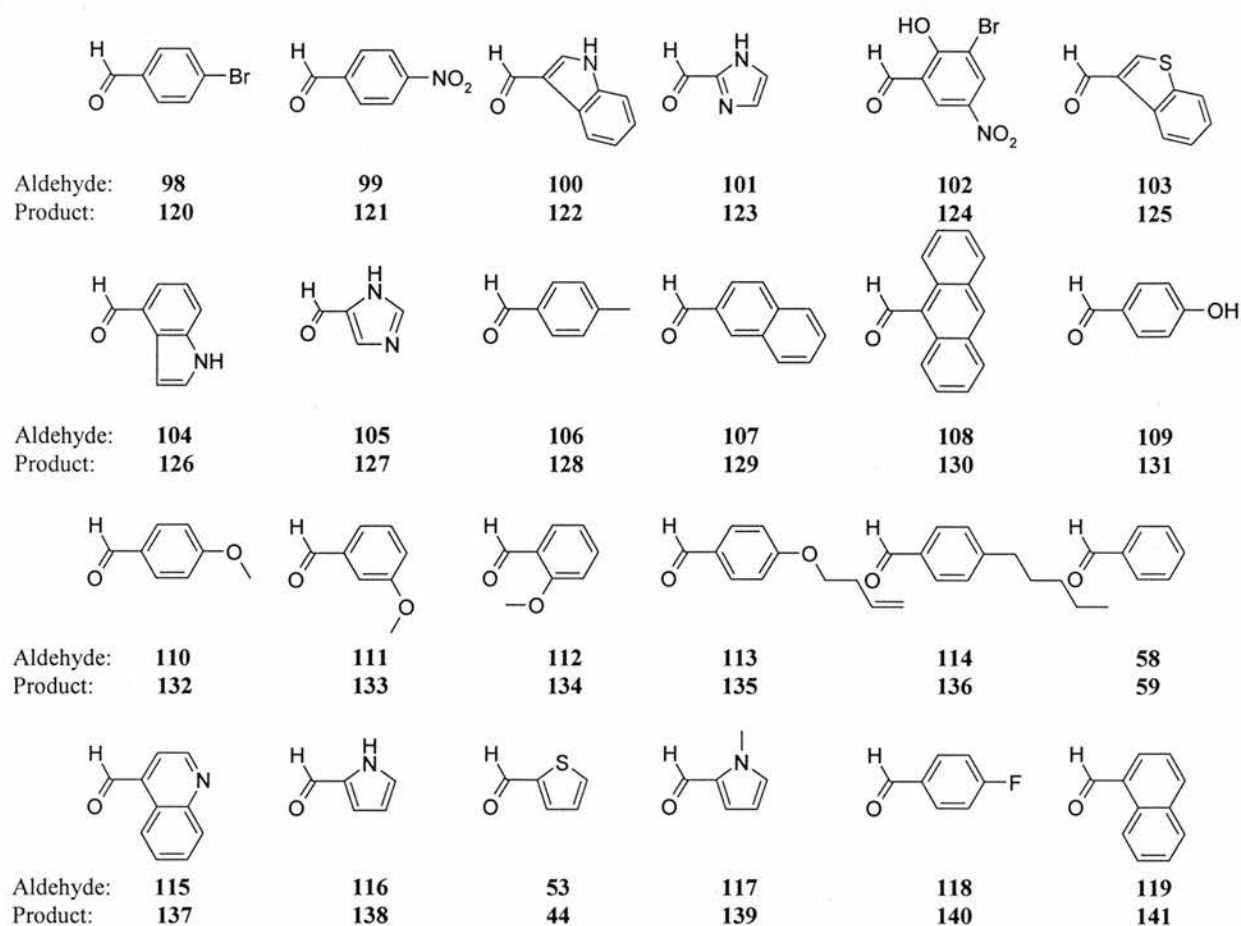


Figure 44. The aldehydes selected for use in the parallel synthesis.

Two of the compounds prepared in the parallel synthesis were **44** and **59** (using aldehydes **53** and **58** respectively). These compounds had been synthesised in solution previously and they were therefore going to provide a marker of the successes of the synthesis (yield, purity) compared to reactions carried out individually in standard chemical glassware.

3.6 Success of the parallel synthesis.

Of the 24 aldehydes used, only two of the reactions with **48**, were unsuccessful, those with aldehydes **99** and **115** (see section 3.7 for further discussion). Initial analysis of the reactions used ^1H NMR and LCMS. ^1H NMR provided information as to the structure of the compound. LCMS provided information as to the molecular weight of the compound and a UV spectrum. Additionally LCMS

enables an impression of the purity of the compounds to be gained (see Appendix). Once the success of the reactions had been established more detailed analysis was used to confirm the structure in greater detail.

One of the significant structural questions surrounding the products of these reactions was the geometry of the exocyclic double bond. There are two methods by which the geometry of the double bond can be assessed: (i) examining X-ray crystallographic data, (ii) using nOe techniques. Both of these techniques were previously used to confirm the geometry of **44** (section 2.1).

Only one of the derivatives synthesised was of sufficient crystallinity to enable X-ray crystallographic analysis, compound **132**. The structure of **132** confirmed the geometry of the double bond as Z (see Fig. 45).

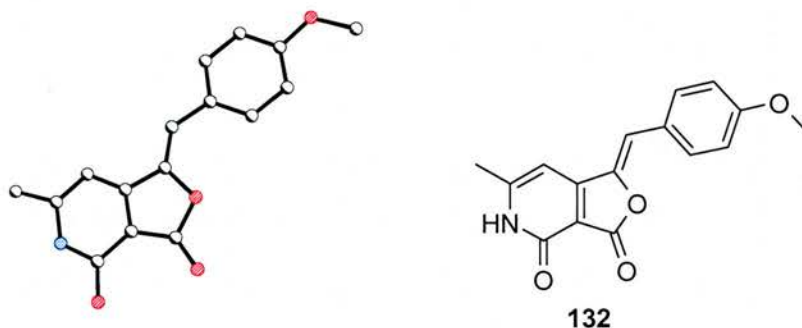


Figure 45. Crystal structure of compound **132**, a product from the parallel synthesis. CCDC No. 292091.

Compound **132** was also subjected to nOe experiments. Irradiating the signal corresponding to H7 appeared to produce a negative nOe at the frequency corresponding to the signal for H8 (see Fig. 46).

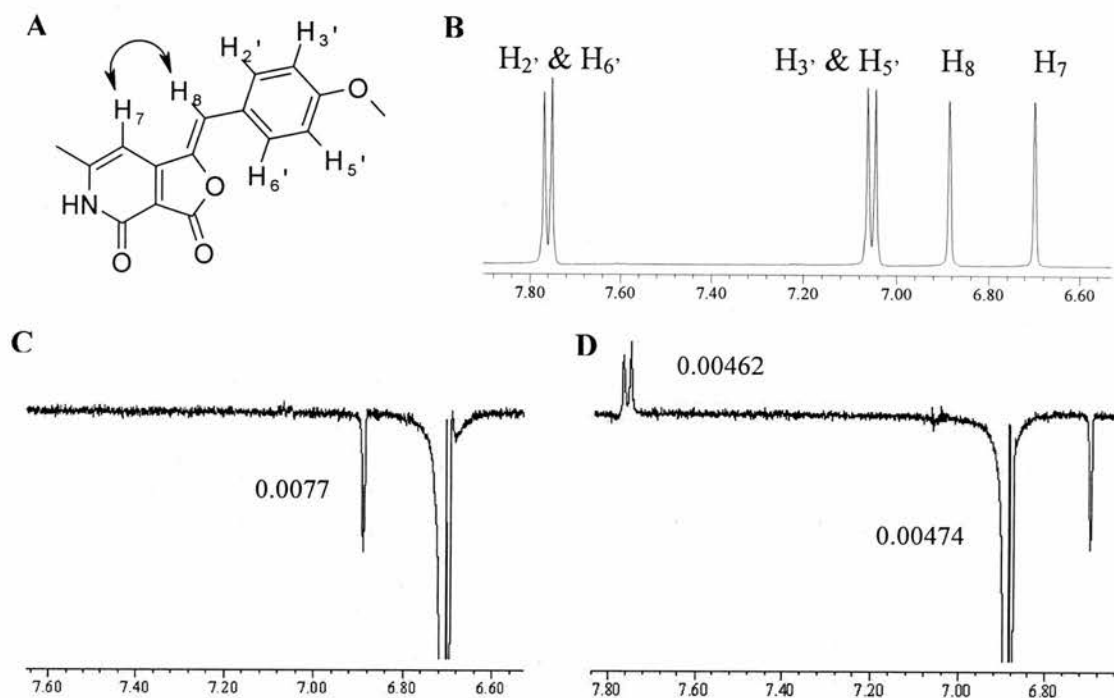
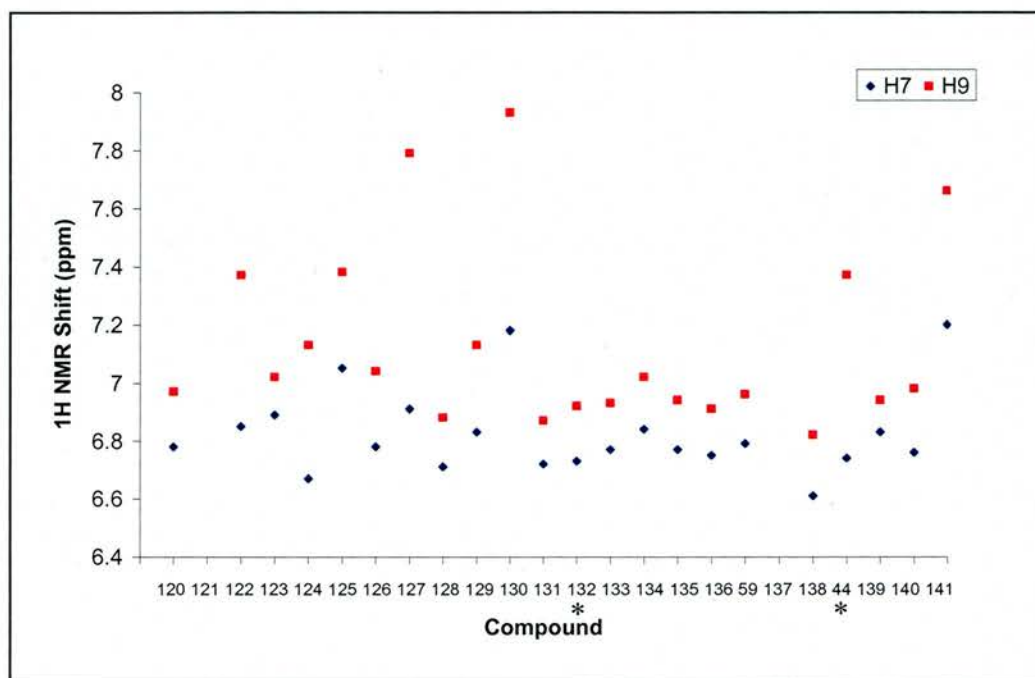


Figure 46. **A** Compound **132**. **B** An expansion of the aromatic region of the ^1H NMR spectrum of **132**. **A**. **C** nOe produced by irradiation of the pyridone proton, H7. **D** nOe produced by irradiation of the proton of the exocyclic double bond, H8. The nOe experiment used was a selective gradient enhanced 1D NOESY experiment.

Negative nOes are generally caused by large molecules (>2000 Daltons) and small molecules (<400 Daltons) produce positive nOes.^{183,184} This is inconsistent with the observed effect on the signal for H8 upon irradiation of H7. One possible explanation results from the fact that the signals observed for H7 and H8 are within 0.2 ppm of each other. When irradiating a proton there is a small region outside of the point of maximum irradiation which also experiences irradiation. The closeness of the signals corresponding to H7 and H8 means that it is possible that the nuclei cannot be irradiated independently of one another. This may result in the signal appearing as a negative nOe and would not allow any true nOe signals to be observed. The observation of a positive nOe for H2' and H6', on irradiation of H8, while a negative nOe is observed for H7 is consistent with this interpretation of the results.

The close proximity of the signals corresponding to H7 and H8 in many types of the analogues of **1** that were prepared (Fig. 47) was unfortunate and examination of the shifts for the signals of the H7 and H8 protons across the 22 compounds, (Fig. 47), showed that for 14 of the 22 compounds these signals were within 0.2 ppm of each

other. Therefore nOe analysis was determined not to be a viable technique for establishing the geometry of the double bond of these 14 derivatives. One of these derivatives **132** has already had the geometry assigned.



Compound	H7	H8	Compound	H7	H8	Compound	H7	H8
120	6.78	6.97	128	6.71	6.88	136	6.75	6.91
† 121			129	6.83	7.13	59	6.79	6.96
122	6.85	7.37	130	7.18	7.93	† 137		
123	6.89	7.02	131	6.72	6.87	138	6.61	6.82
124	6.67	7.13	* 132	6.73	6.92	* 44	6.74	7.37
125	7.05	7.38	133	6.77	6.93	139	6.83	6.94
126	6.78	7.04	134	6.84	7.02	140	6.76	6.98
127	6.91	7.79	135	6.77	6.94	141	7.20	7.66

Figure 47. Comparison of the ^1H NMR shifts corresponding to H7 and H8. * compounds for which X-ray crystallographic and nOe data exist showing Z geometry. † compounds which did not produce the desired product.

As the other eight compounds have shifts for H7 and H8 which are further apart these were examined by nOe. The geometry of **125**, **126**, **130** and **141** was confirmed to be Z. Difficulties with the nOe's of **122**, **124**, **127**, and **129** meant confirmation of the geometry of these compounds was not possible but the indication was that they are likely to have Z geometry.

The remaining 13 compounds are still unassigned. It was proposed that by comparing the shifts corresponding to H7 and H8 for each compound it might be possible to gain some evidence as to whether the *Z* geometry was in a particular compound. Altering the geometry of the exocyclic double bond would be expected to alter the ^1H NMR shifts of the pyridone proton H7 due to deshielding of H7 by the aromatic ring (Fig. 48). Additionally, the *E* geometry could be expected to twist the compound, causing it to no longer be planar. This could effect the conjugation of the compound and therefore the signals in the ^1H NMR.

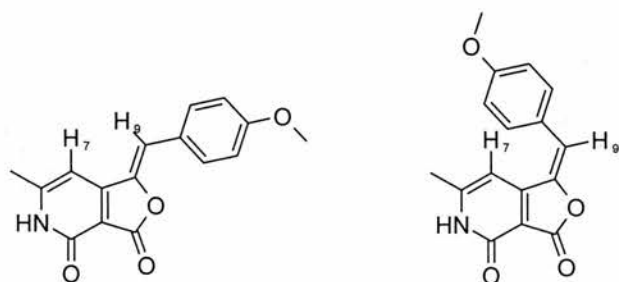


Figure 48. The two possible geometric isomers of **132**.

The majority of these compounds have signals corresponding to H7 and H8 in similar regions (see Fig. 47). This, combined with the already confirmed geometry of **44** and **132**, suggests that at least 14 of the 22 compounds are likely to have *Z* geometry. This analysis method is not ideal but was required under the circumstances. It is worth noting at this point that the parallel synthesis only yielded one compound for each reaction, indicating that only one isomer had been formed from the reaction.

Examining the X-ray crystal structure of **132** also provides information as to which tautomer of the pyridone unit is predominant in the solid state. The C(4)O(4) and C(4)N(5) bond lengths of the pyridone ring, for the four molecules which make up the unit cell in the crystal structure, were between 1.23-1.25 for the C(4)O(4) bond and 1.30-1.40 for the C(4)N(5) adding further evidence that this compound exists in the tautomeric form drawn.

In addition to analysing all the compounds by ^1H NMR they have been characterised. As discussed previously, one challenge was the assignment of the ^{13}C spectra. HSQC and HMBC data were gathered on all the compounds to aid assignment. Identification of a correlation between the CH_3 of the pyridone ring and the pyridone carbonyl (Fig. 49) enabled the assignment of the furanone carbonyl by

default. Further H-C correlations enabled the assignment of the other quaternary carbons.

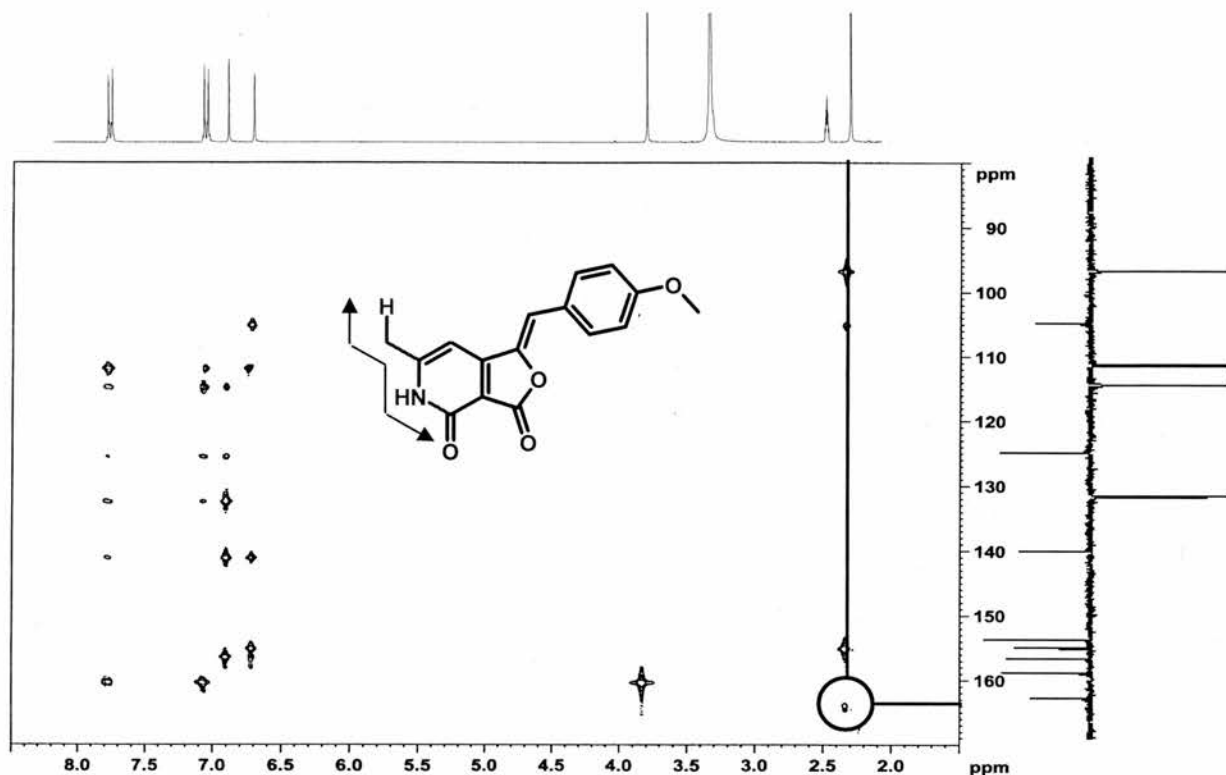
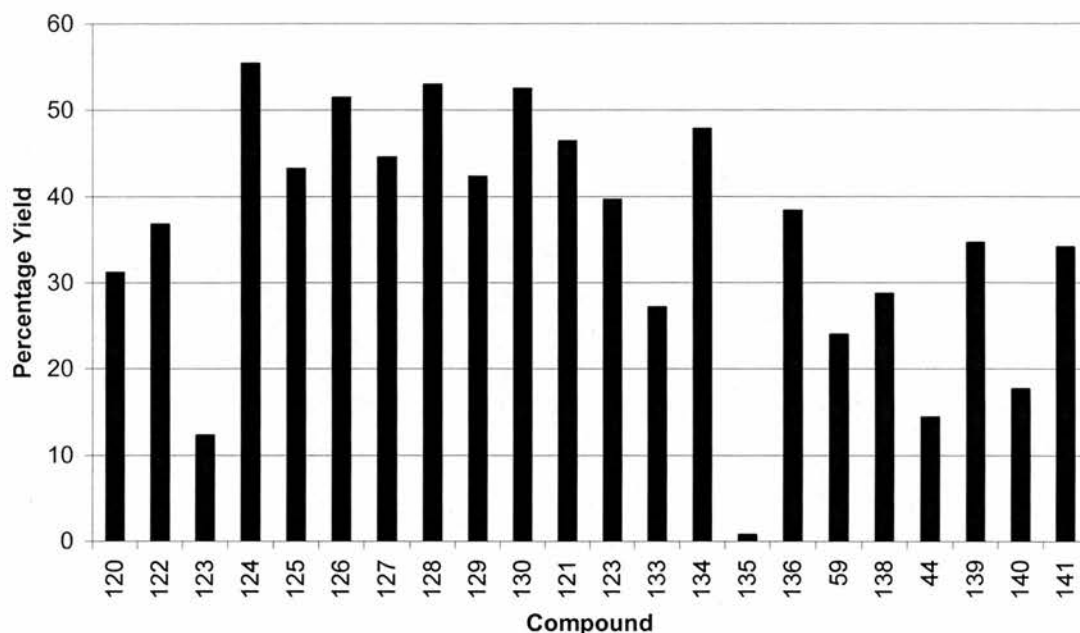


Figure 49. ^1H , ^{13}C and HMBC spectrum of compound **132**. HMBC correlations were observable between the CH_3 and a carbon with a ^{13}C shift of 164 ppm. The carbon was therefore assigned as the carbonyl carbon of the pyridone. NMR in $\text{DMSO}-d_6$, the ^{13}C spectrum shown highlights the quaternary signals, the signal corresponding to CH_3 has been omitted to facilitate the clarity of the image.

In general, the yields for the parallel synthesis were lower than those observed when the reactions were done individually. Reduction in the yield for **44** was encountered from 72% to 14%, the yield for compound **132** went from 58% to 24%. Overall however two thirds of the compounds had yields of greater than 30%, with one third of the compounds with yields over 40%. These reduced yields are not entirely surprising. As discussed, some of the experimental parameters were altered to allow for the use of the reaction in a parallel format. Additionally the nature of the filtration device, although significantly reducing the time required to isolate the products, meant that if any of the products were present as fine solids, significant amounts of solid passed through the frits and were lost. In addition, for compounds

with increased solubility in the reaction solvent the amount of isolated solid was reduced and the yield of the reaction appeared lower than the actual yield.



Graph 2. The yields obtained for the different compounds in the parallel synthesis.

Compound **135** gave a particularly low yield, which may in part be due to the alteration in time of aldehyde addition. The individual synthesis of **135** was achieved in 58% yield. Ideally comparisons of yields would be made across the 22 compounds, leading to conclusions as to the effect of the aldehyde structure on the yield. However the significance of the yields is reduced by the factors discussed. This was evident when comparing the relative yields for **44** and **59**, in solution the yield for **44** is higher by 14%, however with the parallel synthesis the yield for **59** is 10% higher than **44**.

3.7 Reactions which were unsuccessful in the parallel synthesis format

The compounds discussed so far are those which gave the desired product. However, two reactions were not successful. Compounds **121** and **137** were the intended products of these unsuccessful reactions. Noticeably both of these compounds have electron-withdrawing substituents positioned such that they would be expected to facilitate ring opening of the furanone by a nucleophile to yield intermediates of type **142** (Fig. 51).

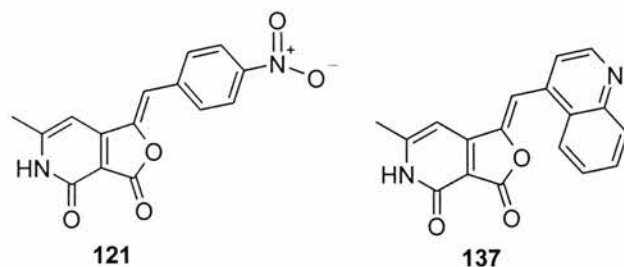


Figure 50. Compounds that were not successful in the parallel synthesis.

The ^1H NMR of the product of 4-nitrobenzaldehyde (**99**) is not very different from the ^1H NMR which would be expected for **121**. The most obvious difference was the lack of a singlet in the 6-7 ppm region for a signal corresponding to H8. The only singlet observable in that region can be assigned to the pyridone proton, H7. The importance of the ^{13}C NMR has already been alluded to and this was one incidence where it proved very important. Unlike **1**, two of the quaternary signals for the product were much further downfield in the region of 180-190 ppm. These types of shifts are commonly observed for carbonyl peaks which have carbons α to them, in comparison to the higher field shifts observed for carbonyls with a heteroatom in the α -position. This is therefore consistent with ketone carbonyls rather than esters or amides.

As electron-withdrawing substituents, (i.e. NO_2), would be expected to facilitate ring opening, this was postulated as a possible route to the observed product. Once **121** has undergone ring opening there are two mechanisms by which **142** can close again to form a five-membered ring. Either the 1-position oxygen can attack the carbonyl carbon to reform the furanone (compound **121**) or the carbon α to the 1-position carbonyl can attack the carbonyl carbon to give the rearranged cyclopentane derivative **143** (see section 2.3 for a more detailed discussion). The likelihood is that **143** would exist mostly as tautomer **144**, due to additional stabilisation possibly brought about by the internal hydrogen bonding between the enol hydrogen and the amide carbonyl. It is also possible to draw a resonance structure of **144**, such that the electrons are delocalised on to the electron withdrawing nitro group, **145**. The increased stabilisation of **144** would explain the formation of this product over the reversion to **121**. The structure of **144** is consistent with the observed analytical data including the ^{13}C spectrum.

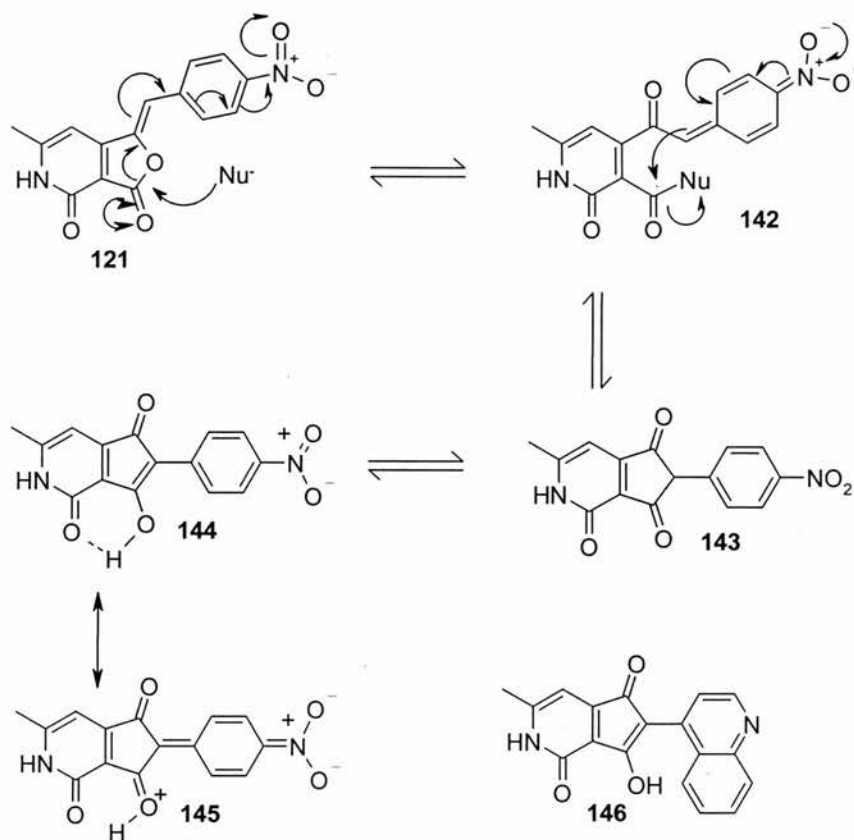


Figure 51. The formation of alternative products from the parallel synthesis.

The product of the reaction of quinoline-4-carboxaldehyde (**115**) with **48** gave spectroscopic data analogous to that of **144** leading to the conclusion that the product was **146**.

Assignment of the structure of these compounds was only possible through the use of NMR and IR techniques. LCMS data, which are typically used as the only point of analysis for large libraries of compounds, would not have suggested that the compounds formed were anything other than **121** and **137**. The same molecular weight would be observed for **121** and **137** as for **143** and **146** respectively. Significant differences were observed in the UV absorbance maxima across the collection of compounds (as exemplified by the differences in fluorescence maxima (section 3.10)). Therefore, it was not possible to use differences in UV absorbance to determine that an alternative compound had been synthesised.

3.8 Determination of solubility

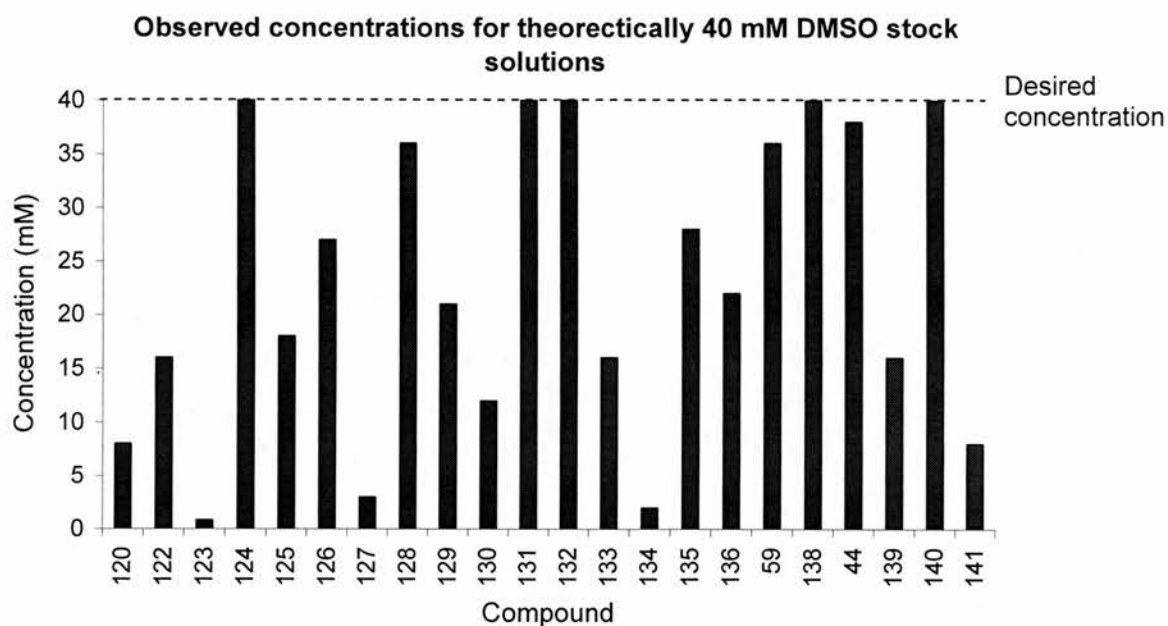
Compounds **44**, **59**, **120**, **122-136**, **138-141**, (the products of the 22 successful reactions) were observed to have low solubility in both organic and aqueous solutions. The degree of solubility of the compounds is significant when determining the biological properties of these compounds (as discussed in Chapter 2, section 2.2.1). Therefore, an analogous approach to the one used to determine the solubility of **44** was used to determine the solubility of the 22 compounds.

A detailed discussion of how the solubility levels were determined has been provided previously (chapter 2, section 2.2.1). However, in order to coincide directly with the procedure used in the biological assay, the initially formed DMSO stock solutions were not centrifuged prior to addition to water (see point 6). The solubilities of the different compounds vary significantly across the series (graphs 3 and 4). Differences that were observed in the solubility in DMSO include:

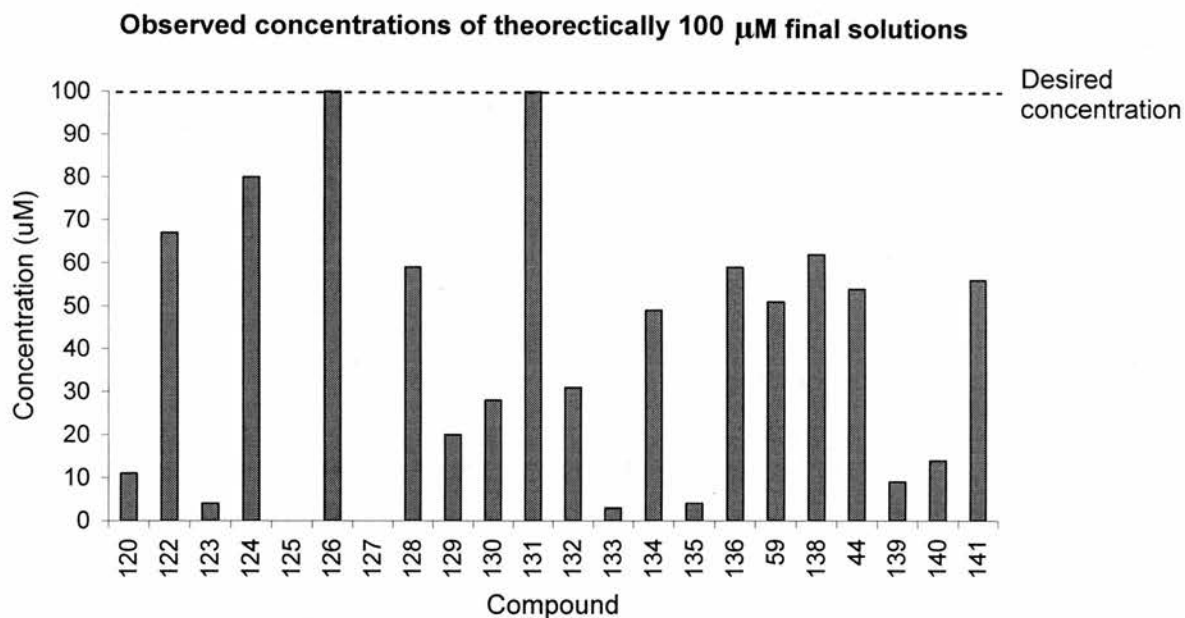
1. Compounds were identified which approach the desired concentration in DMSO of 40 mM: **124**, **131**, **132**, **138** and **140**.
2. Compounds **123**, **127** and **134** are particularly insoluble in DMSO with the concentrations of the stock solutions being less than 10 times the concentration desired (graph 3). Both **122** and **123** contain imidiazole ring systems. This structural link indicates that the imidiazole ring is likely to be reducing the level of solubility of these compounds, possibly due to the formation of extra hydrogen bonding leading to more favourable compound-compound interactions. There is no obvious explanation for the insolubility of **134** in DMSO and the steric hindrance, caused by the *ortho* methoxy group, might have been expected to disrupt crystal packing hence increasing solubility.

Differences in solubility were not only observed in DMSO, the solubilities under aqueous conditions were shown to vary greatly.

3. Highly soluble compounds in aqueous conditions are: **122**, **124**, **126** and **131**. Two of the compounds **124** and **131** also displayed high levels of solubility in DMSO. Both of these compounds contain hydroxyl groups which possibly provide an explanation for the observed solubility. **122** and **126** both contain an indole ring.
4. Compounds that showed low solubility under aqueous conditions were: **120**, **123**, **125**, **127**, **133**, **135**, **139** and **140**. The observed solubility of these compounds has important ramifications for any reported biological activity (section 2.10).



Graph 3. Observed concentrations of solutions intended to be at 40 mM in DMSO.



Graph 4. Observed concentrations of solutions intended to be at 100 μ M in Water/DMSO. Although the solubility of the compounds in the 100 μ M solution has been described as the aqueous solubility of these compounds it is worth recalling that these solutions contain \sim 0.25% DMSO.

5. Compounds were identified with significant differences in solubility in DMSO and water. **135** and **140** showed good solubility in DMSO however they had poor

solubility in water. These compounds do not contain hydrophobic aromatic residues which makes this result surprising.

6. Due to the experimental procedure used, some compounds (**122**, **126**, **134**, **141**), achieved higher concentrations in water than would appear possible given the reported solubility in DMSO. This could occur due to particulates in the DMSO solution dissolving once added to the water, implying that these compounds have significant solubility in aqueous conditions. Two of them, **122** and **126**, are indole structures which could form hydrogen bonds with water potentially resulting in increased water solubility. There is no obvious rationale for the differential DMSO/water solubility of **134** and **141**.

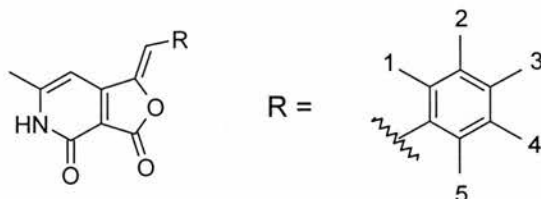
3.9 Biological activities

The biological activities of these compounds were determined using the assay described in Chapter 1 (Section 1.8). In brief, potency levels for these compounds are described as the lowest active concentration (LAC, the lowest concentration at which medium enhancement of invasion was observed, Section 2.1.1). The LAC was determined by comparing the effect of the compounds over a series of 2-fold dilutions in comparison with control wells (DMSO only). The difference observed between the “test” and “control” wells led to the classification of the effect as strong, medium or weak. This was achieved manually, meaning this assay is qualitative although in all cases the experiment was repeated several times and the assay scoring carried out independently by at least two members of the Ward laboratory. All the LAC values discussed here have been corrected for compound solubility (or lack thereof).

Given the qualitative nature of the assay and the significant impact of compound solubility, the values are only quoted to 1 significant figure. In trying to interpret the data, it was also decided that the results were most reliable when dealing with differences between compounds of a factor of 4. Taking this into account clear trends were still observable across the series of compounds. The biological evaluation of the compounds is presented in three tables based on the type of ring systems where R is: a substituted benzene ring; a heterocyclic 5 membered ring; greater than one ring system.

3.9.1 Benzaldehyde derivatives

Table 3 indicates that substitution of the thiophene ring of **1** for a benzene ring was tolerated as previously discussed (Section 2.2.1) and that altering the substituents on the benzene ring can have a significant effect on the activity of the compound.



1	2	3	4	5	lowest active concentration (μM)	Compound number
H	H	CH ₃	H	H	7	128
H	H	CH ₂ CH ₂ CH ₂ CH ₂ CH ₃	H	H	60 (weak)	136
OCH ₃	H	H	H	H	50 (weak)	132
H	OCH ₃	H	H	H	0.09	133
H	H	OCH ₃	H	H	4	134
H	H	OCH ₂ CH ₂ CHCH ₂	H	H	0.5	135
H	H	OH	H	H	6	131
H	H	H	H	H	6	59
H	H	Br	H	H	0.2	120
H	H	F	H	H	2	140

Table 3. Biological activities of derivatives where R = a substituted benzene ring. These activities are corrected for solubility and determined by the lowest concentration medium enhancement is observed at. Those compounds with (weak) written next to the figure do not cause medium enhancement and the figure quoted is the concentration at which they cause weak enhancement.

1. The original hit compound **44** has a LAC of 2 μM. Studying table 3 shows that in the benzene ring series three compounds **120**, **133** and **135** have been identified which are more potent than **44** with LACs of 0.2 μM, 0.09 μM and 0.5 μM respectively.
2. Additionally table 3 shows two compounds **132** and **136** which are only weakly active against parasite invasion. **124** also contains a substituted benzene ring and displays the same weak activity against invasion (see section 3.5 for structure). **124**, **132** and **136** exhibit good to reasonable solubility and this lack of activity

therefore is a result of either a genuine structural mismatch with the target protein(s) (eg. due to *ortho* substituent) or due to factors relating to cell permeability.

3. a) Comparing compounds with substituents in the *para* position, the effects of electronics on the biological activity shows that compounds with electron-withdrawing groups are more potent compared with hydrogen, e.g. **120** versus **59**. The difference in potency between **120** and **140** compared with **59** does not correlate with their relative electron-withdrawing capabilities.

b) Electron-donating groups do not appear to have an obvious trend, with **136** showing weak activity **128** and **134** showing similar activity to **59** and **135** having increased potency.

Substitutions in the *para* position which are electron-withdrawing have already been observed to effect the ease of ring opening of the furanone (section 3.7). The observed increase in potency of **120** may be of importance in assessing the mechanism of action of these compounds (see Chapter 5 for a further discussion).

4. Comparison of **132**, **133** and **134** illustrates the significant impact the position of a given substituent on an aromatic ring can have on the biological activity of the compound. The activities of these three compounds range from only weak levels of activity shown for **132**, to a greater potency than **44** for **133** (0.09 μ M). This difference suggests that substituents in the *ortho* position are not tolerated, which may be due to a structural restriction imposed by the target(s) or an alteration in the physical properties of the compound due to a possible disruption in the planarity of the conjugated π -system.

5. Increasing the steric bulk in the *para* position did not decrease compound potency. Compounds **135** and **120** have increased bulk in this position and are in fact more potent than the original compound **44**. This can be interpreted as a substituent in this position may fill a hydrophobic pocket and improve the binding of the compound. However, the pentyl derivative **136** showed weak activity. This was surprising and it is apparently inconsistent with the results for **135** and **120**.

6. Comparing **131** and **134** shows that it is not possible to improve the potency by increasing the number of hydrogen bond donors and so potentially picking up additional hydrogen bonding through the *para* position of the benzene ring.

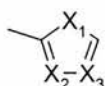
In summary, electronegative substituents in the *para* position appear to increase potency; substituents are not tolerated in the *ortho* position; bulky

substituents can be tolerated in the *para* position; compound potency can not be increased by picking up additional hydrogen bonding at the *para* position.

3.9.2 5-Membered heterocyclic rings

Compound **44** contains a thiophene ring, so the comparison of **44** with compounds containing other 5-membered heterocyclic rings (table 4) was particularly interesting.

1. Substitution of the sulfur in **44** for a nitrogen, **138**, led to no loss in activity (4 μM compared with 2 μM respectively).
2. Comparison of **138** and its *N*-methyl derivative **139** indicates that it was not possible to pick up extra hydrogen bonding interactions, as **139** and **138** have comparable LAC values.



X ₁	X ₂	X ₃	lowest active concentration (μM)	Compound number
S	C	C	2	44
NH	C	C	4	138
NCH ₃	C	C	1	139
NH	C	N	No compound	127
NH	N	C	4 (weak/inactive)	123

Table 4. Biological activities of 5-membered heterocyclic ring containing compounds.

3. Two compounds, **123** and **127**, displayed little or not activity. This dramatic drop in activity can be explained by their insolubility. Compound **127** is completely insoluble under the assay conditions meaning that essentially no compound is present in the assay. Compound **123** was shown to be very weakly soluble which would again explain its very weak activity.

3.9.3 Fused aromatic rings

Increasing the size of the R group to accommodate more than one aromatic ring did not have a detrimental effect on the LAC of the compounds (table 5) except in the case of **125**, which is completely insoluble. This is interesting as the position of the second aromatic ring in most of these cases could be viewed as being akin to an *ortho* substituent, **122**, **125**, **126**, **130** and **141** (discussed in section 3.9.1 point 4).

1. There are three compounds which have two fused aromatic rings which are of greater potency than **44**. **122**, **126** and **129** have with LACs of 1 μM , 0.8 μM and 0.6 μM respectively.
2. The indole-containing **126** was considerably more potent than the naphthyl-containing **141**. This could be explained by the generation of new hydrogen bonding interactions.
3. The 2-naphthalene derivative **129** was more potent than **141** (LAC = 0.6 compared with 7 μM). This observed potency of **129** indicates that the potency of the indole compounds **122** and **126** is unlikely to be due to the generation of additional hydrogen bonding.
4. Compound **130** shows that bulky groups are tolerated in this region of the compound. Again it is surprising that a ring system which would appear to be akin to a di-*ortho* substituted ring systems shows this level of potency (2 μM) given the inactivity of **132**. Also surprising is that **130** is more potent than **141**.

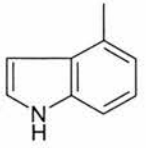
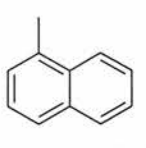
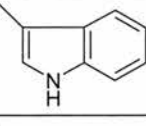
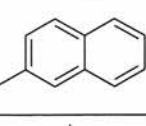
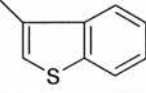
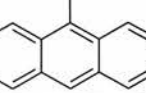
Compound	lowest active concentration (μM)	Compound number	Compound	lowest active concentration (μM)	Compound number
	0.8	126		7	141
	1	122		0.6	129
	No compound	125		2	130

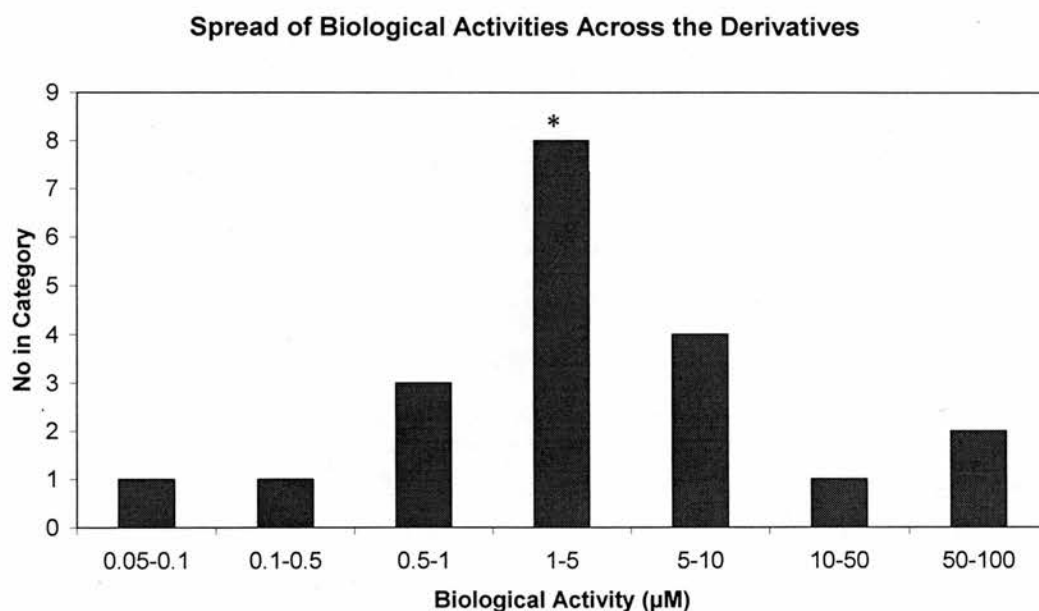
Table 5. Biological activities of compounds with greater than 1 ring system.

5. **125** was significantly less active than compound **44**. **141** is also less active than **44** although the difference is not as significant, so may fall within the experimental errors. The observed decrease in activity for **125** can be explained by the insolubility of the compound. There is however no obvious rationale for the decreased activity of **141**.

3.9.4 Summary

Altering the R group of these compounds has been shown to have a dramatic effect on their activity. Of all the compounds made, five were more potent than **44** and three were inactive due to insolubility.

Graph 5 shows the spread of the LAC values across the compounds. The activities have been grouped into seven categories. This provides an overview of the activities of these compounds given that the nature of the assay (section 1.8) does not readily lend itself to precise quantitative analysis.



Graph 5. The biological activities of the compounds synthesised by parallel synthesis. The compounds are grouped into categories to allow ease of visualisation. * Compound **44** is in this category.

Graph 5 demonstrates that there is a good spread of activities across the series, which would have been impossible to design (without knowing the biological target). This spread of activities indicates that there is likely to be some form of recognition event occurring during the action of these compounds. This therefore points to the possible existence of a specific cellular target of the compounds. After isolation of a potential cellular target, validation of this target could be facilitated by determining if the same activity profile across the compounds is observed with the isolated target.

3.10 Fluorescent properties

Fluorescence is an interesting and useful property. Compound libraries have been generated specifically to identify compounds with different fluorescent properties. M.-S. Schiedel *et al.* synthesised a library of fluorescent dyes of the coumarin type.¹⁸⁵ These compounds are of interest for their pharmacological activity and their applications as laser dyes, fluorescent labels, emission layers in organic light-emitting diodes and as optical brighteners. Of the 151 compounds which were synthesised 34 compounds were identified as “hit” compounds with intense fluorescence over different emission ranges. Q. Zhu *et al.* synthesised a library of 140 Dapoxyl™ dyes (2-pyridyl-5-aryl-oxazoles) analogues.¹⁸⁶ Dapoxyl dyes have been used widely as pH indicators and molecular probes.¹⁸⁶ Screening of the fluorescent properties of the library identified three new fluorescent dyes with novel properties. Rosania *et al.* used a combinatorial approach to synthesise a library of styryl dyes.¹⁸⁷ Styryl dyes are known to label mitochondria. A collection of aldehydes were reacted with pyridinium iodide compounds in a microwave (Fig. 52). The library produced was screened for both the emission wavelength and the localisation of the compounds. In addition to identifying compounds with different emission maxima, compounds were also identified that labelled different cellular compartments, thus providing new chemical tools for use in biology.

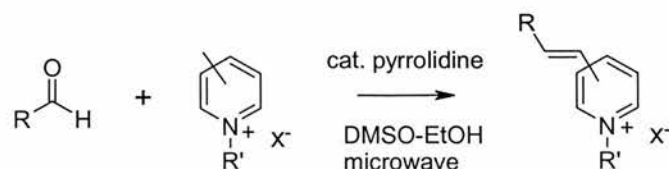


Figure 52. Combinatorial approach to a fluorescent library.¹⁸⁷

The compounds prepared in this chapter have been shown to have very varied physical and biological properties. Another physical property that these compounds possess was that they are fluorescent. When electrons in a compound are excited to a higher energy state the energy which is released on their return to the ground state can be emitted in various forms: fluorescence, phosphorescence, intersystem crossing and internal conversion. These forms depend on the spin states they are excited from and to and the length of time of the emission (Fig. 53). Fluorescence occurs through emission from an excited state to a ground state of the same multiplicity. Excitation of an electron to higher energy levels such as S_2 is followed by an emission of energy through internal conversion to S_1 , from which the electron reaches the ground state through fluorescence emission.

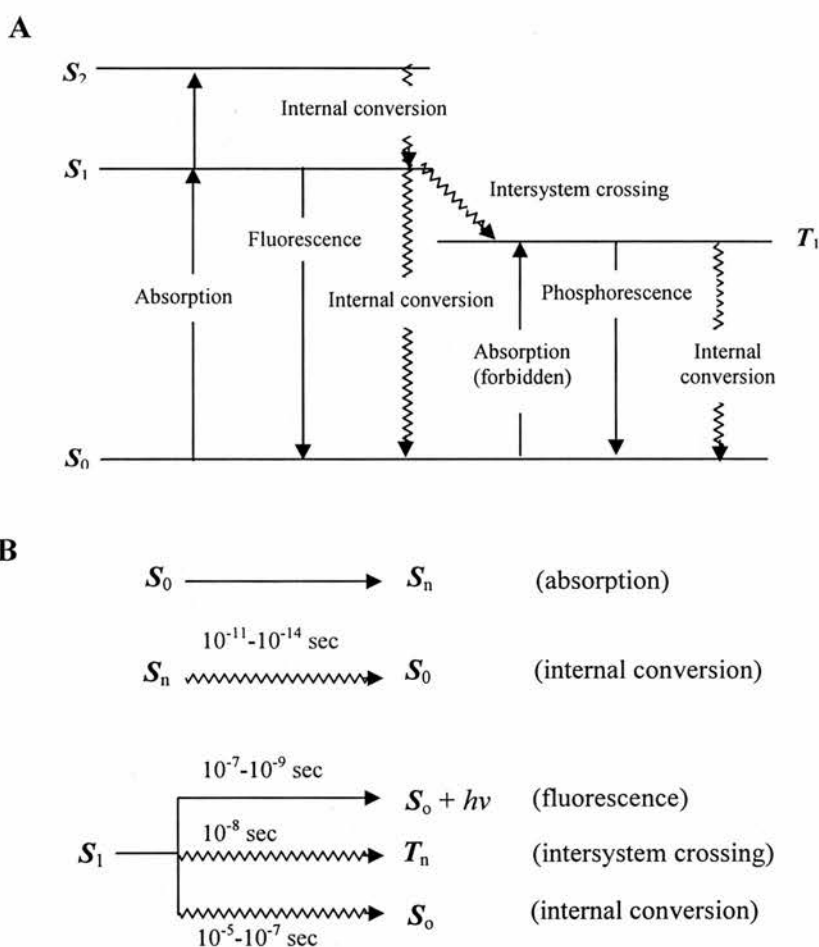


Figure 53. Schematic energy level diagram. **A** This diagram shows the different forms of energy emission and absorption. S is singlet, T is triplet. The S_0 state is the ground state, with the subscript numbers indicate the individual states. **B** An outline of the lifetimes of the different processes.

Adapted from figure 1 in Theory and interpretation of fluorescence and phosphorescence, R. S. Becker.¹⁸⁸

The fluorescent shift observed for a compound is affected by the number of π -electrons present in the compound, the larger the number of π -bonds the lower energy required for excitation and therefore the lower the energy of emission. Lower energy transmissions appear at longer wavelengths, further into the red region of visible spectrum. Therefore in systems with extended conjugation, where the energy gap between the ground state and excited state is reduced, lower frequencies of fluorescence are observed. Substituents on aromatic ring systems affect the observed fluorescence. Those substituents which increase the level of conjugation effect the fluorescence in the manner described above. Other substituent effects do exist but the effect is minimal in comparison to extending the conjugation. One such effect is the inductive effect where groups which feed in electron density (eg. CH_3) produce a shift towards the red region (higher wavelengths) and groups which withdraw electron density (e.g. Cl) produce a shift towards the blue region (shorter wavelengths).

Fluorescence emission and excitation maxima of compounds are determined using a fluorimeter. There are other routes by which the energy gained can be lost, (Fig. 53) in order to determine the extent to which these other pathways are in use the quantum yield of a compound can be obtained. Quantum yields are a measure of the efficiency of the fluorescence of a compound, i.e. how much of the energy absorbed is emitted as fluorescence. The quantum yield of an emission is defined as the number of quanta emitted per exciting quanta absorbed. If the fluorescence is efficient then the quantum yield is 1 as all the absorbed energy is emitted in the form of fluorescence. Comparing quantum yields of derivatives requires that the compounds be excited at the same wavelength, so that their efficiency can be compared.¹⁸⁹ This is not always possible when the difference in excitation between the compounds is too large. Unfortunately, this proved to be the case with this collection of derivatives.

Fluorescence is also of interest from a chemical genetics perspective. Various biological experiments use fluorescently labelled proteins or organisms, for example the *T. gondii* screen assay.¹³⁷ Compounds with inherent fluorescence properties may interfere with such experiments, so derivatives can be selected which do not fluoresce at the wavelength being studied.¹⁹ Additionally the fluorescence property can be used

to select a derivative which is visible using a fluorescence microscope to enable the location of the compound in the cell to be investigated.¹⁸⁶

Fluorescence emission and excitation spectra were taken of all the compounds (table 6). The emission maxima vary significantly across the compounds with values as low as 450 nm and as high as 590 nm observed for **120** and **124** respectively. Similar variation can be observed within the excitation maxima with values ranging from 270 nm to 500 nm for **130** and **124** respectively.

Compound number	λ_{em} (nm)	λ_{ex} (nm)	Compound number	λ_{em} (nm)	λ_{ex} (nm)
120	447	359	132	468	372
122	508	432	133	443	360
123	485	307, 375	134	476	375
124	590	429, 503	135	476	372
125	481	382	136	437	362
126	496	399	59	443	354
127	470	378	138	495	429
128	453	361	44	449	378
129	455	374	139	502	432
130	557	389	140	447	354
131	479	377	141	474	379

Table 6. The fluorescence emission and excitation maxima of the 22 compounds.

Comparison of the *ortho*, *meta* and *para* methoxy analogues **132**, **133** and **134**, has already highlighted the effect substituent position can have on the biological properties. Variation of substituent position was also observed to affect the fluorescent properties of the compounds (Fig. 54).

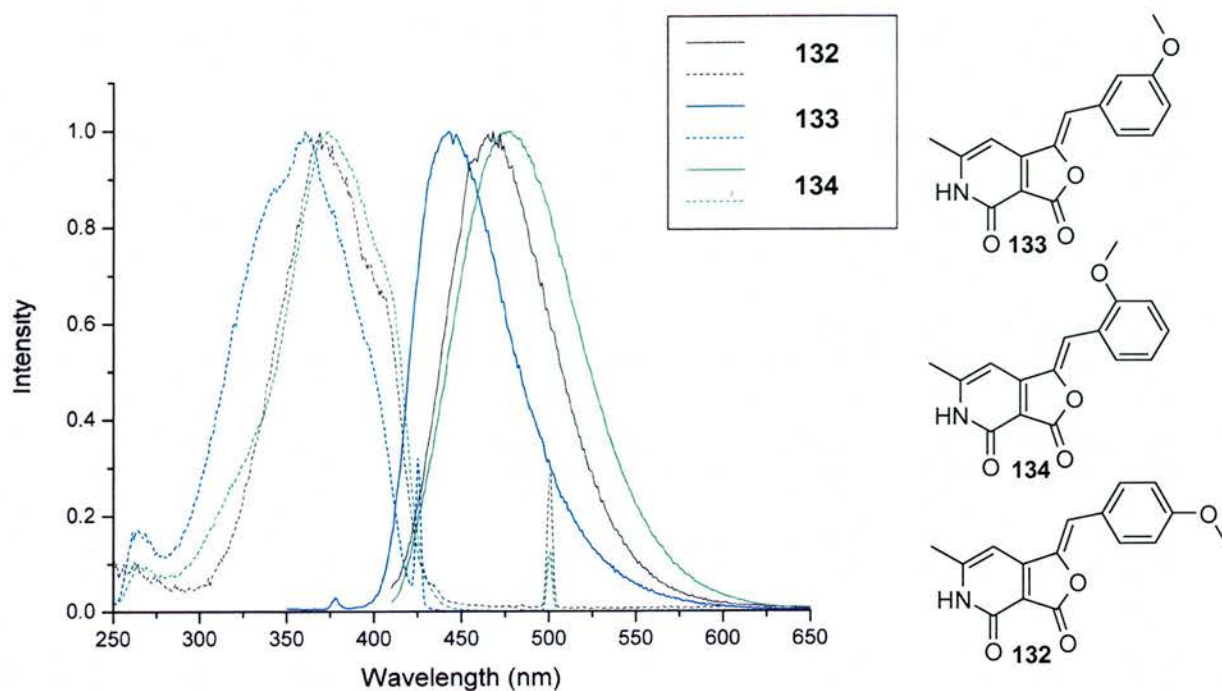


Figure 54. The fluorescence spectra of the methoxy analogues **132**, **133** and **134**. — = emission, - - - = excitation.

The excitation and emission maxima of the *meta* substituted analogue **133** are blue shifted (lower wavelengths) compared to those observed for **132** and **134**. This difference can be explained by the ability of the methoxy groups in **132** and **134** to extend the delocalisation of the electrons in the aromatic ring (and hence the whole compound). The methoxy group of **133** is unable to facilitate the delocalisation of electrons.

As expected, **128**, an analogue with a *p*CH₃ functional group, caused a red shift in the emission (453 nm) compared to **120** (447 nm) which contains an electron withdrawing Br substituent.

The emission maxima of these compounds was used to determine if the compounds could be observed in a cellular environment using a fluorescent microscope. Experiments were performed by Prof. G. E. Ward at the University of Vermont in which the compounds were incubated with the host cells and parasites and a fluorescent microscope was used to observe the location of the compounds within the host cells and the parasite. In general, no localisation was observed but compounds **124** and **127**, did stain the hosts cells and in ~10% of the host cells very bright vacuoles were visible. Some of the compounds appeared to weakly stain the parasites, but in these cases the fluorescent signal was distributed diffusely throughout

the parasite. Assuming that these compounds all bind to the same target(s), the lack of coherent fluorescence staining illustrated that it is unlikely that the target(s) of these compounds will be found through fluorescence co-localisation studies

3.11 Summary

The parallel synthesis of 22 analogues of **44** has given an insight into both the physical and biological properties of these types of compounds. Primary objectives of forward chemical genetic projects such as these involve the synthesis of compounds with improved physical and biological properties (section 1.6.1). This objective has successfully been achieved as compounds with increased potency have been identified. The information achieved from these studies will facilitate the target identification process and, once a target(s) has been identified, provide a valuable set of tools with characterised properties for use in various experiments.

Chapter 4: Developing reagents for affinity chromatography

This chapter focuses on the development of reagents for use in target identification and more specifically, affinity chromatography experiments. Initial results from the use of these reagents are discussed in chapter 5.

4.1 Target ID in protozoan parasites

Given the small number of FCG screens which have taken place in the area of parasitology, it is not surprising that there are only a few reported target identification investigations in this research field. The reported target identification studies generally involve compounds that have been identified through experimental methods other than screening. There are however, successful examples of target identification in parasitology using; affinity chromatography, radiolabelling and photoaffinity labelling techniques.

4.1.1 Affinity chromatography

Affinity chromatography is the most commonly used target ID method in parasitology (see section 1.3.1.1 for a description of the technique).^{36,37,38,41,42,43} For example, experiments have sought to identify the binding partners of; quinolines **147-149** in *P. falciparum*,⁴³ gwnnpaullone (**150**) in *L. mexicana*³⁸ and polyamine **151** in *T. cruzi*.⁴¹

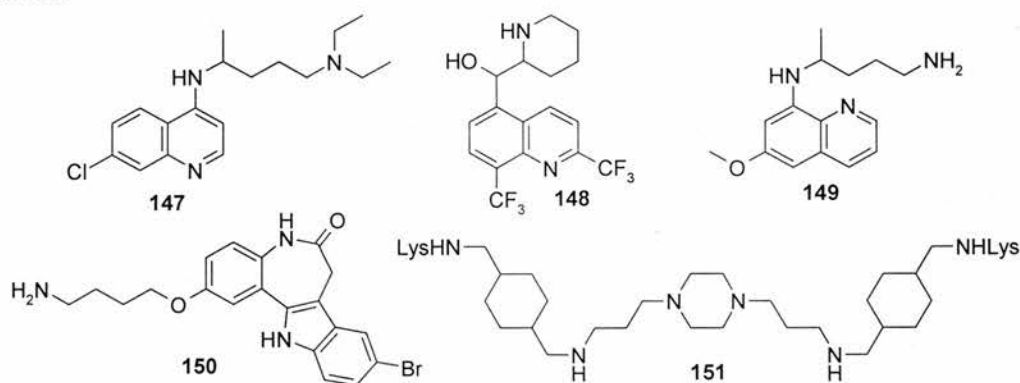


Figure 55. Compounds utilised in affinity chromatography experiments. Chloroquine (**147**), mefloquine (**148**), primaquine (**149**), gwnnpaullone (**150**) and a polyamine **151**. Where stereochemistry is not drawn the compound is racemic.

Knockaert *et al.* performed affinity chromatography experiments using the same resin with four different protozoan parasites: *P. falciparum*, *L. mexicana*, *T. cruzi* and *T. gondii*.³⁷ The compound which was loaded onto the resin was a derivative of purvalanol B (**152**) an established cyclin-dependent kinase (CDK) inhibitor. Using this affinity resin Knockaert *et al.* aimed to identify novel protozoan kinases that bound to purvalanol B (**152**). A control resin was also used in Knockaert's research to identify those proteins that were non-specifically bound to the resin (see section 4.2.2 for a further discussion on controls).

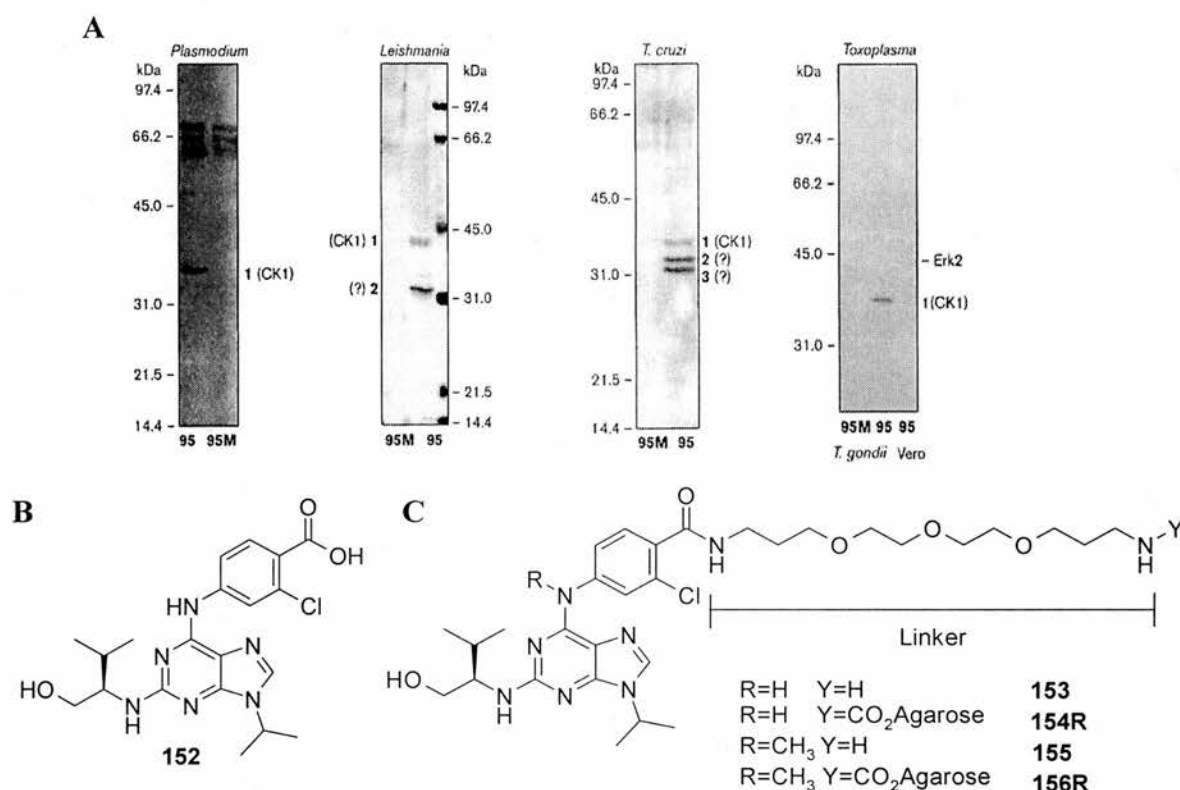


Figure 56. The use of affinity chromatography in parasitology. Affinity chromatography studies were carried out using crude protein extracts from four different protozoan parasites *P. falciparum*, *L. mexicana*, *T. cruzi* and *T. gondii* and resins loaded with either **153** or **155**, (see C). The bound proteins were eluted and analysed by SDS-PAGE/silver stain, (see A). The gels clearly show the isolation of one protein CK1 for *P. falciparum*, two proteins are identified with *L. mexicana*, CK1 and an unknown protein. CK1 was also pulled down from the *T. gondii* and *T. cruzi* extracts (with two other unidentified proteins in the latter case). Image reproduced from Knockaert *et al.*, *Chem. Biol.*, **2000**, *7*, 411-422.³⁷

The kinase CK1 was identified from all four of the parasitic lysates indicating the conservation of this kinase as a purvalanol B (**152**) binding partner across the parasites. Apicomplexan parasites *T. gondii* and *P. falciparum* did not appear to

contain any other kinases which would bind purvalanol B (**152**). However, incubation of the resin with extracts from the kinetoplasts *L. mexicana* and *T. cruzi* identified one and two additional unknown kinases, respectively. This series of experiments demonstrates the success possible through the use of affinity chromatography. Additionally, it also highlights the fact that once a reagent has been synthesised it can be used in experiments involving different species.

4.1.2 Radiolabelling

Radiolabelling has been used to identify the cellular targets of a known inhibitor of *Eimeria* growth, 4-[2-(4-fluorophenyl)-5-(1-methylpiperidine-4-yl)-1H-pyrrol-3-yl]pyridine (**157**).¹⁹⁰ The radiolabelled analogue **158** was synthesised and incubated with crude protein lysates from *E. tenella* (Fig. 57). Purification of the lysate through a series of columns yielded two fractions which were identified as containing **158**. Examining these fractions by 1D SDS page showed the presence of two major protein bands that were distinctive to these fractions. The 120 kDa protein, (Fig. 55B), was sequenced and identified as cGMP-dependent protein kinase (PKG).

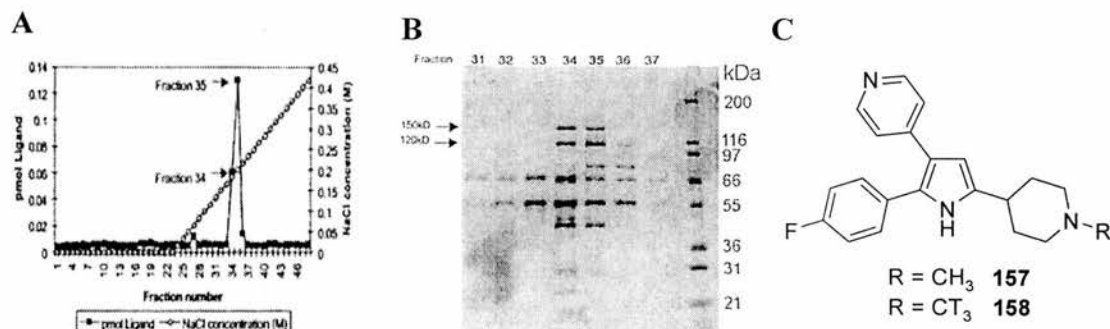


Figure 57. The use of radiolabelling in parasitology. **A** Chromatogram showing fractions (34 and 35) that contain the radiolabelled compound **158**. **B** Silver stained gel indicating two bands at 150 kD and 120 kD which were present as the major distinct bands in fractions 34 and 35. Image reproduced from Gurnett *et al.*, *J. Biol. Chem.*, **2002**, 277, 15913-15922.¹⁹⁰

Further experiments confirmed the identity of the target by inhibition of isolated PKG and by the generation of **158** resistant *Eimeria* and *T. gondii* parasites, which were subsequently shown to have mutations in the catalytic site of PKG. Compound **158** is also known to inhibit the growth of other apicomplexans: *T. gondii*, *N. caninum* and

B. jellisoni.¹⁹⁰ Following the success of the experiment with *Eimeria*, further investigations have sought to confirm PKG as the cellular target of **158** in *T. gondii*¹⁹¹ and to investigate the function of PKG in apicomplexan motility and invasion.¹⁹²

4.1.3 Photoaffinity labelling

Photoaffinity labelling experiments within parasitology have centred on the mechanism of action of the antimalarial drugs chloroquine (**147**) and mefloquine (**148**) with the goal of identifying their cellular targets in *P. falciparum*.^{55,56,193,194} A radioiodinated analogue of the chloroquine (**147**) derivative **159** was incubated with *P. falciparum* infected erythrocytes and identified two proteins which bind to **159** at 42 and 32 kDa.⁵⁵ The 32 kDa protein was later identified as lactate dehydrogenase (LDH).⁵⁶ Two proteins of 67 and 28 kDa were also identified from incubation of a radioiodinated analogue of **160**, a photoaffinity derivative of mefloquine (**148**) with *P. falciparum*, however, the identity of these proteins has not yet been reported.

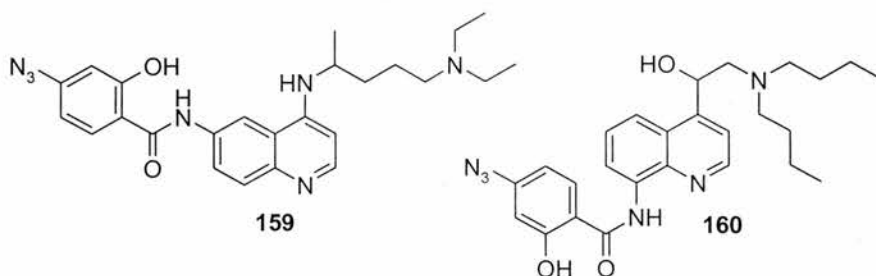


Figure 58. The use of photoaffinity labelling in parasitology. ASA-Q (**159**), a photoaffinity reagent based on chloroquine and ASA-MQb (**160**) a photoaffinity reagent based on mefloquine. **159** and **160** were radioiodinated using Na^{125}I and used in photoaffinity experiments as mixtures of regioisomers. **159** and **160** were used as racemic mixtures.

Radiolabelled versions of **159** and **160** were also incubated with serum from infected and uninfected erythrocytes. The identification of binding between **160** and a serum polypeptide apo-A1 suggested possible serum transportation of these drugs. Additionally, **160** was found to bind stomatin which is involved in cation transport indicating a possible route for uptake into erythrocytes.¹⁹³

4.2 Designing a reagent for affinity chromatography

Affinity chromatography was selected by us as the method for trying to establish the cellular target of compound **44**. The requirements to perform affinity chromatography experiments are: i) a derivative of **44** immobilised onto a resin, ii) a control resin based on **44** and iii) a protein extract.

4.2.1 Immobilisation of a derivative of **44** onto an affinity resin.

Derivatives synthesised for affinity chromatography contain a linker system which terminates in a functional group suitable for attachment to a resin. The derivative (or a protected analogue) must retain the biological activity observed with the original compound. The selection of the position of attachment of the linker system is therefore key to the success of any affinity chromatography reagent. The structure activity relationship data outlined in Chapter 2, enabled the selection of this position.

Whilst the position of a linker is important, equally important is the selection of the specific linker system to be attached. Linker systems typically contain large numbers of CH₂ units and the resulting derivatives are hydrophobic in character. These linker systems can undergo hydrophobic collapse when loaded onto a resin rendering the compound unable to interact selectively with any proteins presented to it. Additionally the large hydrophobic surface areas presented by linker units of this type could lead to an increase in the level of non-specific protein binding.

The hydrophobic nature of linker systems is often reduced through the addition of heteroatoms into the chain. Polyethyleneglycol (PEG) units are often used, these consist of multiple copies of OCH₂CH₂ units.¹⁹⁵ This reduces hydrophobic collapse but there are additional problems which are associated with linker units of this type, e.g. metal ion chelation. If the linker units are of an unsuitable length, (greater than four CH₂CH₂O units), then the linker is able to chelate metal ions such as calcium and magnesium also resulting in a collapse of the linker unit. Linker systems made of three CH₂CH₂O units have been used successfully, for example Knochart *et al.* used this linker unit to immobilise the Purvalanol B (**152**) derivative **153** onto an agarose resin to give resin **154R** (see section 4.1.1, Fig. 56).³⁷

Other linker systems use repeating units of 6-aminohexanoic acid ($\text{H}_2\text{N}(\text{CH}_2)_5\text{CO}_2\text{H}$) (**161**).¹⁹⁶ The amide units disrupt any intralinker hydrophobic interactions and they have the advantage of not having the metal chelation problems potentially experienced with PEG based linkers. This type of linker unit is commonly used with the biotin-streptavidin system. Videlock *et al.* tagged peptides with biotin (**7**) (Fig. 59) in order to create a probe Bio-pYZZZ (**162**) for use with a phase display library.¹⁹⁶ The biotinylated probe was attached to an avidin-coated 96 well plate through the biotin-avidin interaction and incubated with 11 different cDNA phage display libraries in order to identify a series of different proteins containing Src homology 2 (SH2) domains.

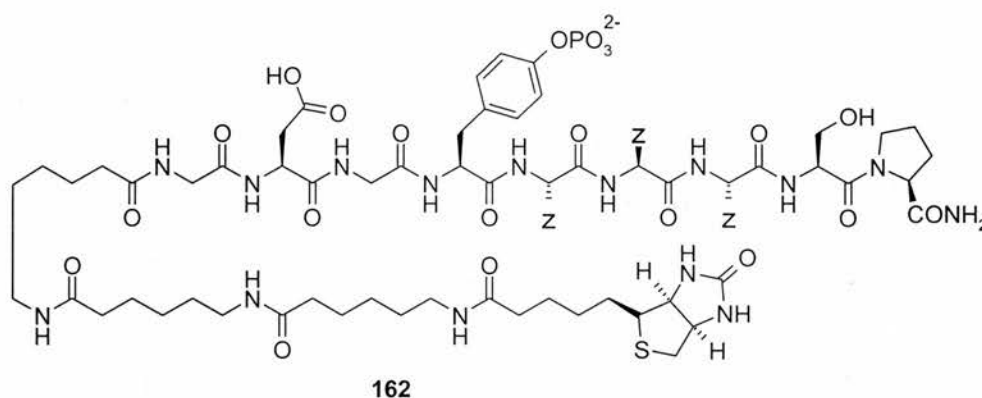


Figure 59. Reagent used in the screening for proteins containing SH2 domains. Bio-pYZZZ (**162**).

From the 464 screened phages, 50 phages were identified and of these, 40 encoded for proteins with SH2 domains.

4.2.2 Controls for affinity chromatography experiments

Once the compound has been successfully modified and loaded onto a resin the focus moves towards the preparation of a control. Controls are required in affinity chromatography experiments in order to identify proteins that bind non-specifically to the resin. There are two types of controls which can be used in affinity chromatography; i) competitive inhibitors/enhancers, ii) a control resin.

4.2.2.1 Competitive enhancers/inhibitors

Control experiments using competitive enhancers/inhibitors involve the preincubation of the protein extract with the compound prior to exposure to the affinity resin. In principle, the compound will bind to its target protein(s) in the extract and hence prevent it from binding to the affinity resin. Displacement of the compound by the resin bound analogue is unlikely as immobilisation of a compound via the attachment of a linker unit usually results in a decrease in potency.³⁴

Competitive inhibitors have successfully been used, for example Chen *et al.* synthesised three derivatives, **163**, **164** and **165** of myriocin (**166**) with different points for attachment onto a resin, in order to identify myriocin-binding proteins.³⁴ These derivatives were synthesised and tested as their benzylated derivatives **167**, **168** and **169**, to avoid potential problems of reactivity and stability of the thiol.

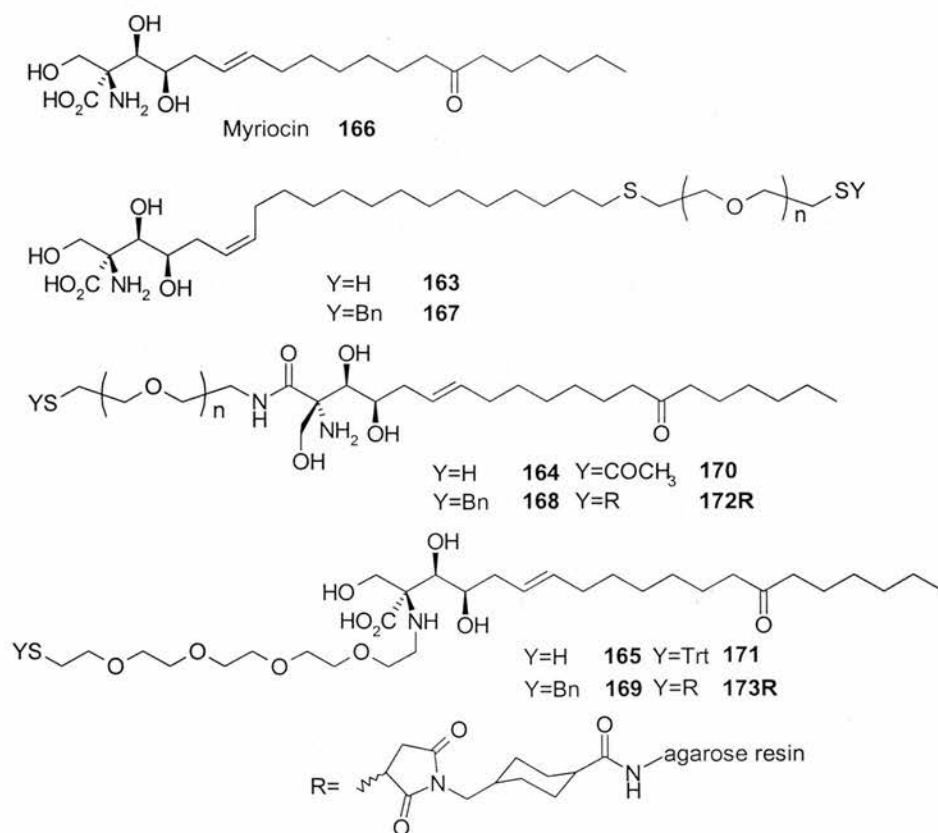


Figure 60. Myriocin (**166**) and derivatives with linker units with differing points of attachment.

Of the three different points of attachment: C20 **167**, C1 **168** and N **169**, **168** and **169** were less active than myriocin with IC₅₀s of 300 nM and 500 nM respectively (IC₅₀

10 nM for myriocin). **167** was completely inactive, illustrating the importance of the selection of the position for the linker. After an alteration in the protecting group of **168** to an acetyl group and **169** to a trityl thioether, deprotection of **170** and **171** was achieved. Subsequent loading of **164** and **165** onto Affi-Gel 102 resin yielded the reagents **172R** and **173R**. **3R** was incubated with CTLL-2 lysates which had been preincubated in the absence or presence of myriocin (**166**). Two myriocin binding proteins LCB1 and LCB2 were reported to bind to the resin. These proteins did not associate with the resin in the presence of myriocin (**166**).

4.2.2.2 Control resins

Control resins are prepared by attaching an inactive analogue which is as close in structure as possible to the active compound on the resin. Frequently the disruption of an essential hydrogen bond through the methylation of an NH group provides a suitable inactive analogue. Knockeart *et al.* synthesised a control resin for use in their affinity chromatography experiments (see section 4.1.1, Fig. 56).³⁷ The NH group in the 6-position had been identified as essential for Purvalanol B (**152**) activity. Therefore, immobilising a N6-methylated Purvalanol B (**152**) derivative **155**, to give resin **156R** generated a control resin (Fig. 56).

4.2.3 Protein extracts

Protein extracts in affinity chromatography experiments contain a large number of different proteins from which it is hoped that putative protein target(s) can be identified. Extracts are usually generated through the incubation of large numbers of the cell of interest with a detergent which breaks down the cell wall and solubilises the proteins. Selection of the detergent can have a dramatic effect on which proteins are solubilised (made evident in experiments outlined in section 4.4). Octylglucoside (**174**) and NP40 (**175**) are two detergents which can be used to generate cellular extracts.¹⁹⁷

Given the large number of proteins which can be isolated from a cell, it is sometimes necessary to reduce this number through various forms of prefractionation. Fractionation can involve the isolation of proteins from different cellular compartments, such as the proteins from the nucleus versus the cytoplasm. This can

be achieved through the careful selection of detergents and use of centrifugation.¹⁹⁷ Identifying binding partner(s) from fractionated protein samples can also provide information as to the probable cellular location of the identified proteins. Prefractionation can also be used to look for specific types of cellular targets. Graves *et al.*⁴³ wished to identify quinoline protein targets from the purine binding proteome of *P. falciparum*. One of the approaches they took involved using an ATP-Sepharose resin to select from crude protein lysates of both *P. falciparum* and human red blood cells, those proteins which bind ATP.⁴³ Elution of the ATP loaded columns with three quinolines successfully identified two proteins from human red blood cells which bind the quinolines, but unfortunately no proteins were identified from *Plasmodium falciparum*. This use of ATP-Sepharose resin has since been commercialised by Serenex®.

Prefractionation of the parasite proteins is less desirable in this case as the amount of protein that can be obtained from a reasonably sized parasite culture is low. Therefore, the protein extract used in experiments described in this thesis was a crude non-prefractionated extract prepared from *T. gondii* parasites.

4.3 Synthesis of an affinity chromatography reagent.

For ease of discussion in this chapter the cellular target of **44** will be referred to as a protein. While it is certainly true that the target of **44** is most likely to be a protein, it is important to acknowledge that **44** could be influencing the parasite through other cellular components. For example interacting with cell walls to allow increased permeability to ions (without rupturing the cell walls (as shown by the cytotoxicity assays)) or intercalating DNA. The cellular extracts used in affinity chromatography are formed using detergents and centrifugation. Centrifugation removes any large components of the parasite but essentially anything from within the parasite which is solubilised by the aqueous detergent is in the extract and can bind to the resin. It is often the case that membrane proteins are poorly represented within an extract, reducing their chances of being successfully identified if they are the target(s).

4.3.1 Synthesis of an analogue of 44 containing a linker unit.

4.3.1.1 Selection of a linker system and point of attachment.

Possible positions for incorporation of the linker system have already been highlighted (Chapter 2).

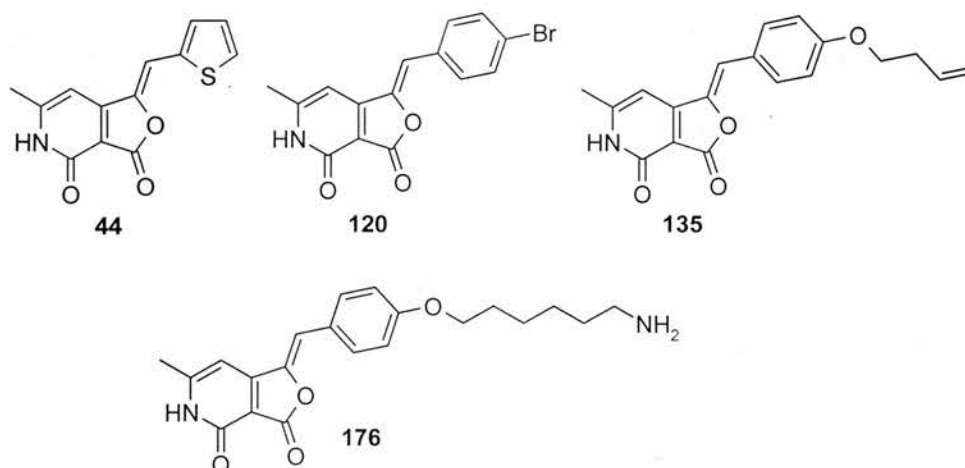
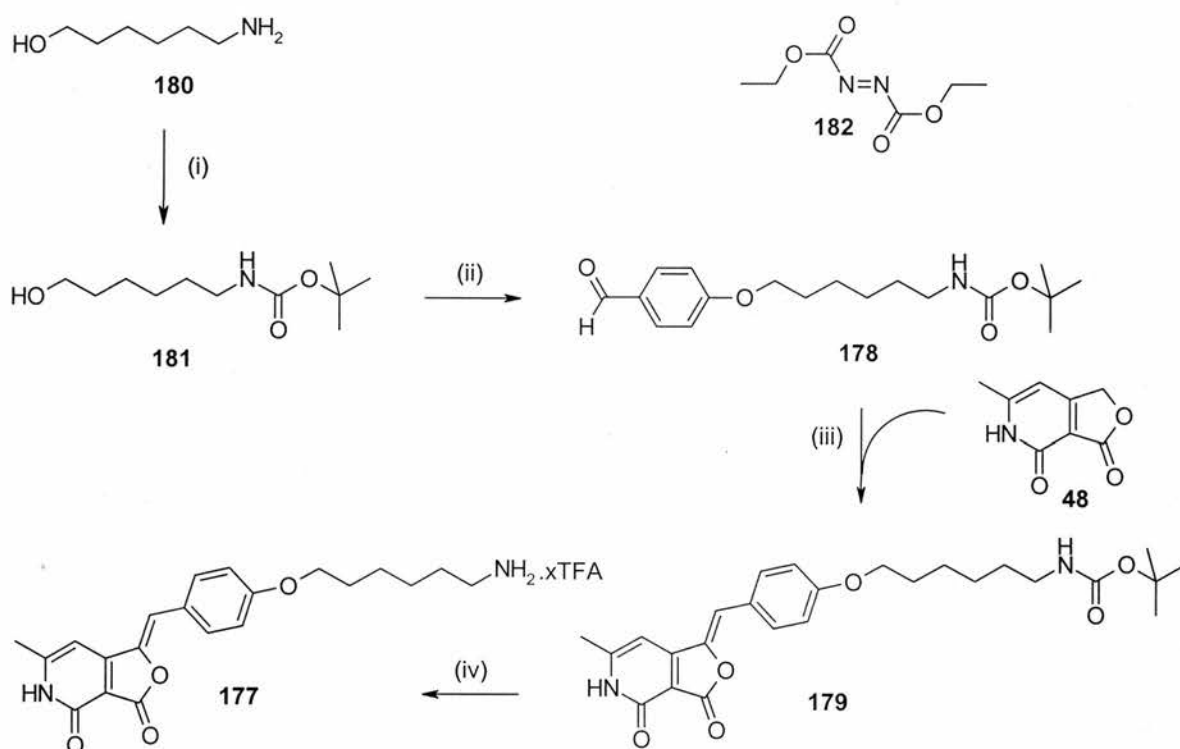


Figure 61. Compounds of relevance to designing an affinity chromatography resin.

Parallel synthesis of derivatives with alternative aromatic rings indicated that substitutions in this area were not only tolerated, but in some cases improved the activity in the cell-based assay. Synthesis and testing of the *para*-bromo **120** and the *para*-butenyl **135** derivatives showed that large substituents can be incorporated at the *para* position and the biological activity retained. This led to the identification of a target compound with a linker system attached *para* to the aromatic ring, compound **176**. The linker unit was selected as it contains only six CH₂ units which should minimise hydrophobic collapse. Collapse due to chelation with metal ions is not a problem with this alkyl linker system. The nucleophilic NH₂ group provides a possible point of attachment to an affinity resin. A large number of affinity resins are commercially available, a number of which are activated towards attachment of compounds through an NH₂ group (e.g. resin that contain *N*-hydroxysuccinimide ester functional groups).

4.3.1.2 Synthesis of a compound containing a short linker unit.

Compound **177** was synthesised by a four step synthetic sequence (Scheme 18). The flexibility of the Knoevenagel-type condensation has already been established so it was proposed that reaction of aldehyde **178** with the furanone **48** would yield the Boc protected precursor **179** of **177**. Synthesis of the aldehyde **178** was achieved by a Mitsunobu reaction between Boc protected 6-aminohexan-1-ol (**181**) and 4-hydroxybenzaldehyde (**109**) in good yield. The Mitsunobu reagent DEAD (**182**) was used in this case.



Scheme 18: (i) Di-*tert*-butyl dicarbonate, DCM, 0°C, 82% (ii) 4-Hydroxybenzaldehyde (**109**), DEAD (**182**), triphenylphosphine, toluene, 80°C, 98% (iii) 6-Methyl-1H,5H-furo[3,4-c]pyridine-3,4-dione (**48**), piperidine, ethanol, 95°C, 60% yield, (iv) TFA, RT, >95%.

Condensation of **48** with **178** proceeded with an acceptable yield. Although the alkyl chain in **179** resulted in its increased solubility in ethanol, **179** was observed to precipitate in small quantities while the reaction mixture was under reflux. Allowing the reaction to cool to room temperature resulted in the precipitation of additional **179**. Improved yields were observed when the reaction volume was kept to a minimum. Deprotection of **179** with trifluoroacetic acid gave **177** is an essentially

quantitative yield. **177** was isolated without the neutralisation of excess TFA, therefore **177** was isolated as its TFA salt, as confirmed by ^{19}F NMR.

Extensive analytical characterisation of both **179** and **177** was achieved. The signals in the NMR of **179** and **177** for H7 and H8 are with 0.05 ppm of each other. Therefore, as discussed previously nOe studies are not an appropriate method to determine the geometry of the exocyclic double bond. However, an X-ray crystal structure of **179** was obtained confirming the *Z* geometry, consistent with other analogues **179**.

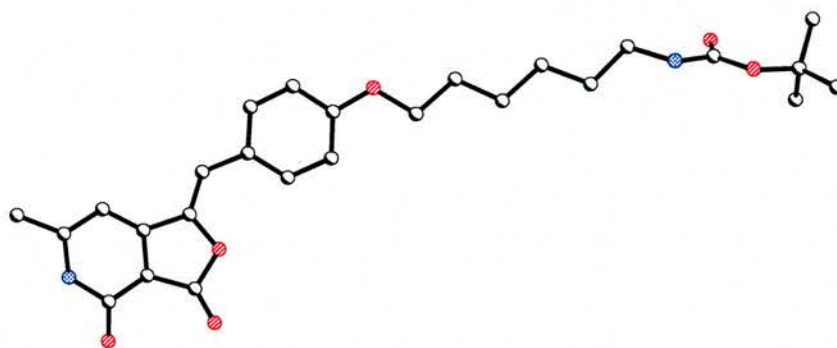


Figure 62. X-ray crystal structure of **179** confirming the *Z* geometry. CCDC No. 292092.

The LACs of **179** and **177** were found to be $25\ \mu\text{M}$. It was planned that **177** would be bound to the resin through an amide bond, therefore, **179** more closely resembles the species which would be found on the resin. Although the decrease in activity is undesirable, it was decided to continue along this route and prepare the affinity resin.

4.3.1.2.1 Light transformation

Compound **179** was observed to undergo a clean chemical transformation when dissolved in deuterated chloroform in the presence of sunlight. The transformation was observed through an alteration in the ^1H NMR spectrum (Fig. 63). The most dramatic change was in the signal corresponding to the olefinic proton H8 which shifted upfield from 6.37 to 4.76 ppm. Smaller shifts were also observed for the signals corresponding to the pyridone proton, the H2' and H6' protons of the aromatic ring and the CH_2 adjacent to the oxygen.

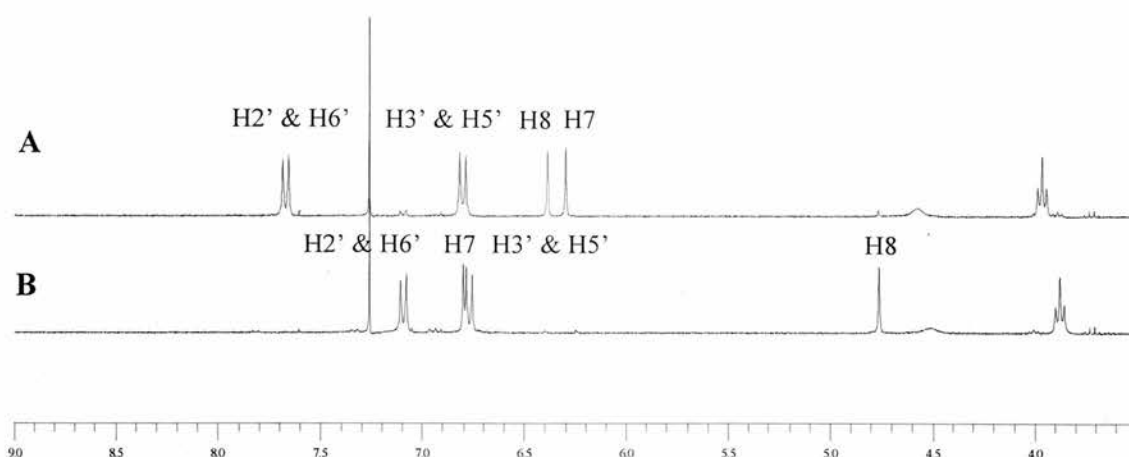


Figure 63. ^1H NMR spectra of the observed light transformation of **179**. **A** ^1H NMR of **179**. **B** ^1H NMR of the product of the light induced transformation of **179**.

Similar shifts were also observed for the signals corresponding to the respective carbons in the ^{13}C NMR spectra. The signal for C8 shifted from 113.0 to 51.5 ppm; C1 from 140.9 to 89.7 ppm. The shift in the signals for C8 and C1 indicated the absence of the double bond, however there was no change in multiplicity of these signals in the ^1H NMR.

The mass spectrum of the compound revealed the presence of a peak at 960, twice the mass of **179**. IR spectroscopy of **179** shows a stretch assigned to the furanone carbonyl at 1795 cm^{-1} . The light transformed product gives a stretch in the IR spectrum at 1785 cm^{-1} , possibly indicating that this structural feature is almost certainly retained but that a change on the furanone ring has occurred that modified the stretching frequency of the carbonyl group.

Investigations into the conditions required for this transformation involved the following: light, pH, oxygen, solvent. The transformation was also observed in $\text{DMSO-}d_6$. Dissolving the sample in chloroform and keeping the sample in the dark prevented any transformation. Samples in base washed or CDCl_3 which had been degassed again underwent the transformation in light, suggesting no acid or oxygen dependence.

Having established that the transformation was light induced, the structural features of the compound required for this transformation were investigated next. **183**

had been synthesised for the purposes of identifying a control compound (section 4.5.2.2). **183** was dissolved in CDCl_3 and $\text{DMSO-}d_6$ and monitored over the same time periods as used for **179**. Interestingly no chemical transformation was observed.

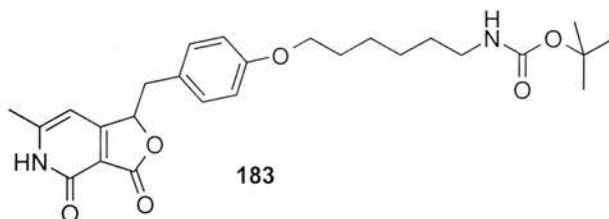


Figure 64. A compound which did not undergo a chemical transformation.

The most significant piece of information acquired which enabled the proposal of a structure for the new product was gained through HMBC spectra. Coupling was observed between C8 to C8' which was indicative of an AA' system. Given this and the observed molecular weight the cyclobutane ring containing structure **184** was proposed (as drawn in Fig. 65).

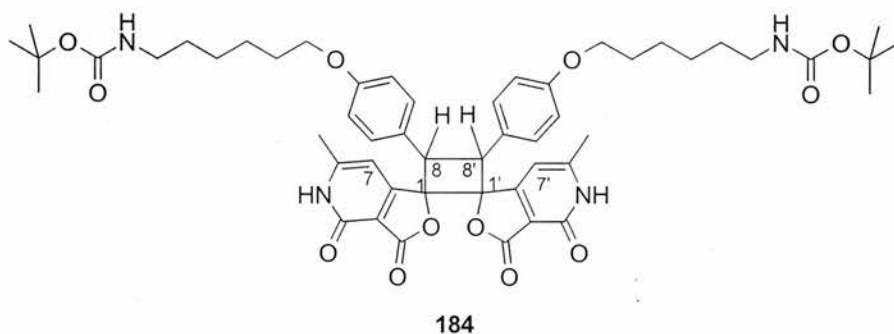
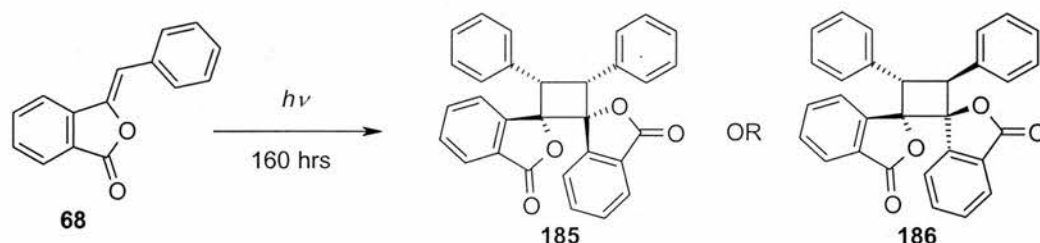


Figure 65. The proposed product of the transformation of **179** in light.

The symmetry of **184** and the new carbon-carbon bond between C8 and C8' explains the AA' coupling observed. As H8 is no longer an allylic proton an upfield shift for the signals corresponding to H8 and C8 in the ^1H and ^{13}C NMR spectrum would be expected and were observed. The chemical shift in the ^{13}C NMR spectra of the signal corresponding to C1 was also consistent with the assigned structure. The minimal changes observed in the ^1H and ^{13}C NMR signals corresponding to the pyridone ring supported retention of this structural component. The IR spectrum can be rationalised based on the continued presence of the furanone, but on the absence of the double bond (c.f. **179**).

Formation of **184** was probably via a [2+2] cycloaddition reaction. This also explained why transformation of **183** was not observed. [2+2] cycloadditions have been reported in the literature between two molecules of benzylidenephthalide (**68**), to yield either **185** or **186** (Scheme 19).¹⁹⁸ The authors were not able to assign the exact structure only to conclude that it was one of two symmetrical compounds, **185** or **186**, formed from *Z*-benzylidenephthalide.



Scheme 19. Literature [2+2] cycloaddition of benzylidenephthalide (**68**).¹⁹⁸

A [2+2] cycloaddition between two molecules of **179** could potentially lead to 6 different products (see Fig. 66). However, the ¹H NMR spectra of **184** indicates that it must be symmetrical. Therefore, the product of this reaction cannot have the structure **187** or **188**. **189** and **190** require the two starting monomers to have *E* geometry for the exocyclic double bond. **191** and **192** require *Z* geometry.

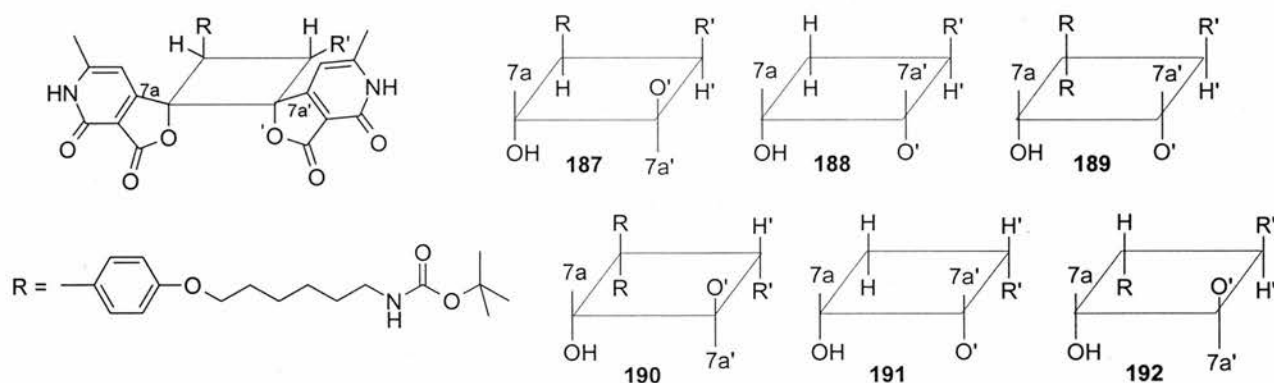


Figure 66. The possible products formed by a [2+2] cycloaddition of two molecules of **179**.

179 was confirmed as having *Z* geometry, which would imply that the structure of the new product is either **191** or **192**. This does not however rule out the possibility that **179** is undergoing isomerisation to the *E* isomer prior to the [2+2] cycloaddition on a

time scale shorter than that being observed. Experiments are on going in an attempt to further define the stereochemistry of **184**.

Light induced transformations were not observed with the majority of the analogues of **44**. The cycloaddition is likely to be concentration and irradiation wavelength dependent. **179** is the only analogue which was successfully dissolved in both CDCl_3 and $\text{DMSO-}d_6$. The difference in the excitation maxima of these compounds has already been highlighted (section 3.10). The implication of this transformation is that any chemical or biological experiments with this compound or its deprotected derivative **179**, must be performed in the absence of any light.

4.3.2 Loading of compounds onto affinity resins

4.3.2.1 Affinity resins

Commercially available affinity chromatography resins contain functional groups such as *N*-hydroxysuccinimide esters or imidazole carbamates that are used to attach the compound under study. These functional groups are chosen so as to be sufficiently stable to allow storage of the resin but also to enable compounds with amine functional groups, or proteins to form covalent links to the resin under mild conditions. The formation of a covalent linkage is important as the compound must stay bound to the resin under elution conditions that are sufficiently strong to displace proteins which are non-specifically bound to the resin.

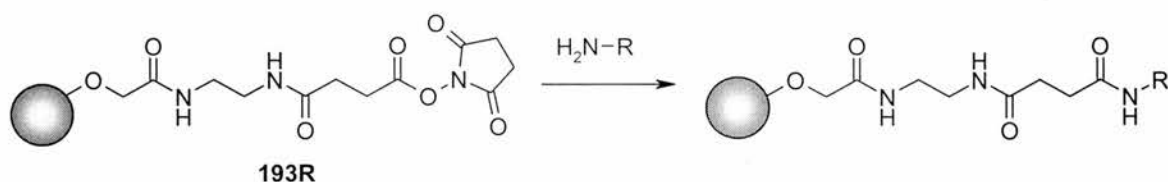


Figure 67. The loading of an amine onto a resin through the use of a *N*-hydroxysuccinimide ester group. This figure also shows the linker chain on Affi-Gel 10 (**193R**) which terminates in a *N*-hydroxysuccinimide ester and the final product which is produced after loading of an amine on the resin.

The resins used were 4% cross-linked agarose resins. Agarose resins are highly porous as they are composed of agarose fibres knitted together. Tolerance of agarose

resins to a range of temperatures, pHs and solvents mean they are favoured as they offer sufficient versatility to enable the resins to be loaded under conditions that are optimal for a variety of ligands.¹⁹⁹

4.3.2.2 The loading level of the affinity resin

Once the compound has been loaded onto the resin the loading level of the compound, i.e. the yield of the reaction, needs to be determined. This loading level defines how many of the loading sites on the resin have been modified with the compound and how many remain unreacted (e.g. remain as *N*-hydroxysuccinimide esters).

4.3.2.2.1 Significance of the loading level

Achieving the desired loading level is important as this can affect the availability of the compound for binding to its protein binding partner(s). The number of remaining loading sites which proteins themselves could potentially covalently bind to must also be considered. It can be argued that the ideal loading level would be 100%; if all of the resin was covered with compound then protein(s) interacting with the compound would be higher. In addition, there would be no unreacted loading sites to bind covalently to proteins. However, there is the question of how close the compounds are to each other in space. If the compounds are close enough to allow them to interact with each other, through hydrogen bonding or π -stacking (as these compounds have already been shown to do) this could prohibit them from binding to proteins. There also needs to be sufficient space around the compound for the protein to bind unhindered. For all these reasons, loading levels have been used for these types of experiments in the region of 10-20%.^{34,200}

4.3.2.2.2 Determining the loading level

Determining the loading level is not as straight forward as determining the yield of a reaction. Characterising the resins directly is difficult meaning that the most frequently used method to determine the loading level is to measure the amount of compound unsuccessfully loaded onto the resin. Once the resin has been incubated

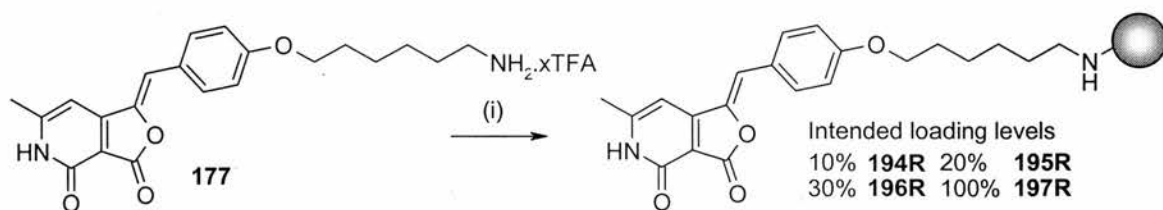
with the compound for the required amount of time, (usually 24hrs), the resin is carefully washed to remove any unbound compound. The washings are then combined and examined.

One method of determining how much compound has not been loaded onto the resin is the concentration of the washings *in vacuo* and the use of NMR to determine the quantity of compound present. This is achieved through the addition of a known amount of a standard to a solution prepared from the washings. The standard must have a characteristic and clear signal in the ^1H NMR which is significantly removed from the signals corresponding to the compound being loaded. Integration of a signal corresponding to the standard with respect to the compound allows for the concentration of the compound to be determined.

In our case, however, it has already been shown that these compounds absorb UV light and that this property can be used as a means of determining their concentration (see section 2.2.1). Therefore this was the method chosen to determine the amount of compound not loaded onto the resin and hence the loading level of the resin.

4.3.2.2.3 Loading 177 onto an affinity resin.

177 was loaded onto the resin as described in scheme 20. All the following procedures were carried out in the dark where possible. The deprotection of **179** with trifluoroacetic acid (TFA) gives compound **177** as the TFA salt. Therefore, loading **177** onto a resin required the addition of a base (triethylamine).



Scheme 20. Loading of compound **177** onto an affinity resin. (i) Affi-Gel 10 (**193R**), triethylamine, DMSO, 24 hrs.

Resins were loaded at the intended levels of 10 (**194R**), 20 (**195R**), 30 (**196R**) and 100 (**197R**) %. The varying levels were selected in order to enable optimisation of

the affinity chromatography protocols. After incubation of the resin with **177**, the resin was washed with DMSO. The volume of the washings was made up to 50 cm³. This was then diluted further and a UV spectrum taken of the sample. Standard solutions were used to determine the extinction coefficient of **177**. The loading levels of the resins were determined as: 9.4 (**194R**), 19.9 (**195R**), 29.1 (**196R**) and 97.8 (**197R**) % suggesting that in general the loading protocol was efficient.

4.3.2.3 Blocking of unreacted sites

The presence of any remaining loading sites on the resin was a potential problem due to the possibility that proteins may become covalently linked to the resin. Blocking of these sites with compounds such as ethanolamine is commonly used to avoid this problem.^{201,202} Ethanolamine reacts with the *N*-hydroxysuccinimide esters and leaves the surface of the beads covered with alcohol groups which the proteins cannot covalently bind to.

However, there is literature precedent for the use of resins without blocking the remaining loading sites (in cases where non-covalent complexes are formed with the protein).²⁰⁰ If the interaction between the protein target(s) and the compound is strong enough the protein should preferentially bind to the compound especially given that the compound aided by the linker unit will have been loaded onto the most accessible sites on the resin. However, if the compound covalently modifies its the protein target(s), then the presence of unreacted loading sites may be problematic.

In our case, ethanolamine should be used to block the resins. However, derivatives of **44** react with butylamine, and would also be expected to react with ethanolamine. Incubating the loaded resin with ethanolamine could therefore cause unwanted modification of the compound on the resin. Two approaches were taken to study the possibility of blocking the resin (i) examining the plausibility of using ethanolamine and (ii) examining the resin itself to determine the need for resin blocking.

4.3.2.3.1 NMR experiment to determine the potential success of blocking with ethanolamine

The relative reactivity of resin-bound **177** compared with the resin-bound *N*-hydroxysuccinimide (NHS) esters could potentially enable selective reaction of the ethanolamine with the loading sites only (when the appropriate amount of ethanolamine is added). To address this question, ^1H NMR spectroscopy was used to examine the relative reactivities of the two functional groups involved. Compounds **198** and **179** were dissolved in $\text{DMSO-}d_6$ in the ratios **198:179**, 80:20, in order to mimic 20% loaded resin.

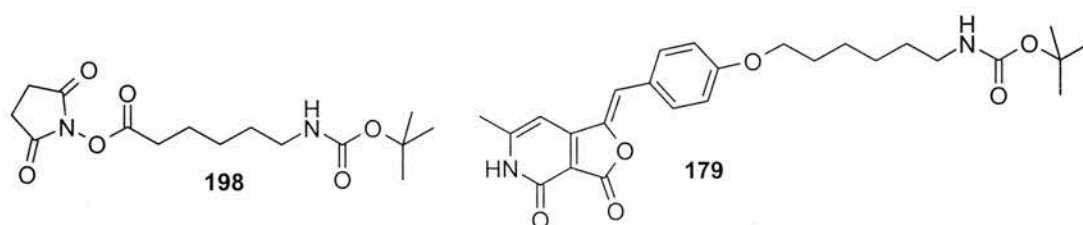


Figure 68. Compounds incubated with ethanolamine to determine their relative reactivity.

Addition of ethanolamine in the ratio ethanolamine:**198:179**, 100:80:20, caused a reaction with both **198** and **179** as expected. The ethanolamine in the ratio ethanolamine:**198:179**, 30:80:20 caused only the **198** to react with the ethanolamine. This experiment suggests that 80 equivalents of ethanolamine could be added to create selective blocking without altering the active compound on the resin.

However, **177** occupies the 20% most accessible sites on the resin. In the ^1H NMR experiment described above the two different compounds were equally accessible making this experiment not directly comparable to the situation on the resin. Given this difference in accessibility, one approach would be to incubate the resin with enough ethanolamine to block 30% of the resin, which would then only react with the NHS-esters, meaning overall 50% of the resin will have blocked sites (either with compound or ethanolamine). This would mean that the 50% of loading site still available for protein binding would at least be the 50% least accessible sites.

4.3.2.3.2 Attempts to analyse the resin

The above experiments (section 4.3.2.3.1) have assumed that all the sites on the resin not loaded with **177** are still present as NHS esters after loading of **177**. There is the possibility that the excess of loading sites may be hydrolysed over time or during the loading procedure. Resins which had been treated under three different conditions: unmodified **193R**, loaded with **177** (**195R**), hydrolysed with NaOH **199R**, were dried and IR spectrum were taken (Fig. 69).

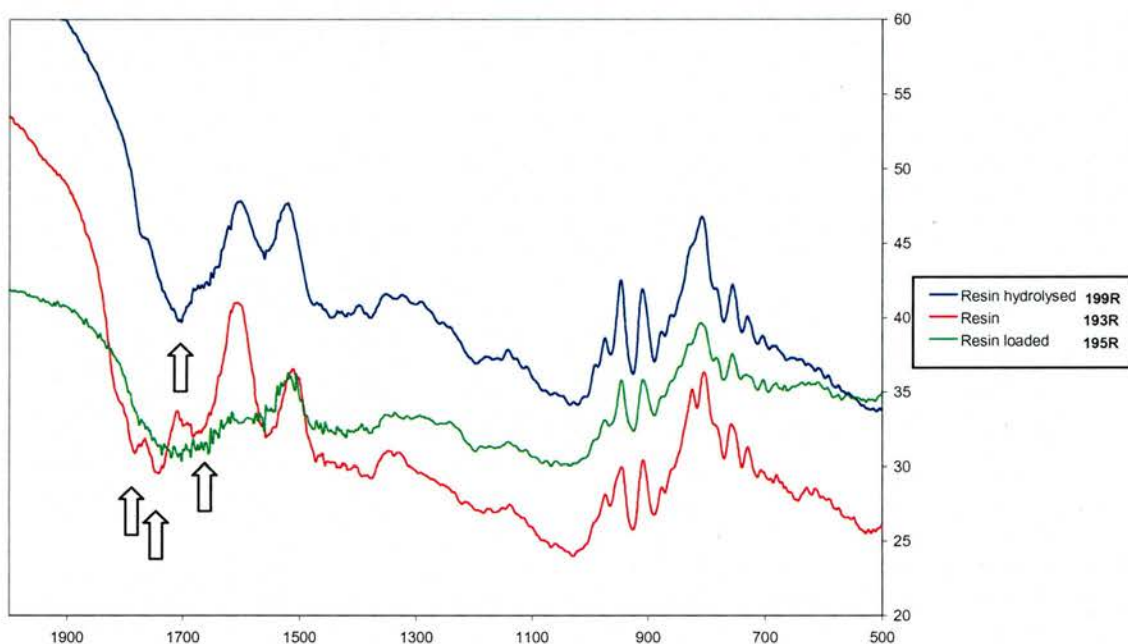


Figure 69. IR spectra of three resins. Blue line: resin hydrolysed by incubated with NaOH (**199R**). Red line: unmodified resin (**193R**). Green line: resin loaded with **177**, (**195R**).

There are definite differences observable between the unmodified resin (**193R**) and the hydrolysed resin (**199R**). The **193R** has two main peaks at 1800 and 1700 cm^{-1} which disappear on hydrolysis, (the peak indicated for the spectrum of the hydrolysed resins probably corresponds to the acid carbonyl). This indicates that incubation of the resin with NaOH is sufficient to hydrolyse the NHS-ester groups as expected and that this can be observed by IR. The signals between 1800 and 1700 cm^{-1} appear to have disappeared in **195R**. This could indicate that the remaining NHS groups have been hydrolysed under the loading conditions. However, the region of 1600-1800 in the IR spectrum of the loaded resin is unclear meaning this conclusion may be

incorrect. This lack of definition in the IR spectrum could be due to the addition of two extra carbonyls from compound **177** increasing the absorbance in this area. Unfortunately IR spectrometry has been unable to determine the extent to which the loading groups are still present.

Given the difficulties with resin blocking, both synthetic and analytical, all the affinity chromatography experiments reported here were performed with unblocked resins.

4.4 Initial affinity chromatography experiments performed in the laboratory of Prof. G. E. Ward at the University of Vermont.

The purpose of this was to initiate affinity chromatography experiments. There were 5 variables which were to be examined: (i) detergent selection, (ii) loading levels, (iii) elution conditions, (iv) controls and (v) cell lysate.

The basic procedure for the affinity chromatography experiments is outlined in Chapter 6 (section 6.3.3.3) and summarised in Figure 70. The initial experiments involved the use of blank resin (**193R**) as a control.

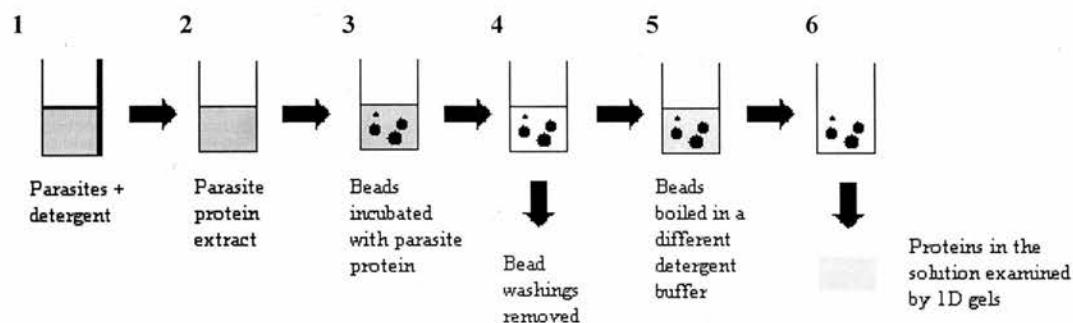


Figure 70. Schematic describing the general procedure for the affinity chromatography experiments.

(i) The first experiment examined the effect of the detergent used to generate the cell lysate. Two detergents, octylglucoside (**174**) and NP40 (**175**), were used to generate *T. gondii* extracts. These were incubated with both blank resin, **193R** and **195R**. Examining the 1D gels, showed that a few protein bands appeared to be eluted from **195R** that were not eluted from the control resin **193R**.

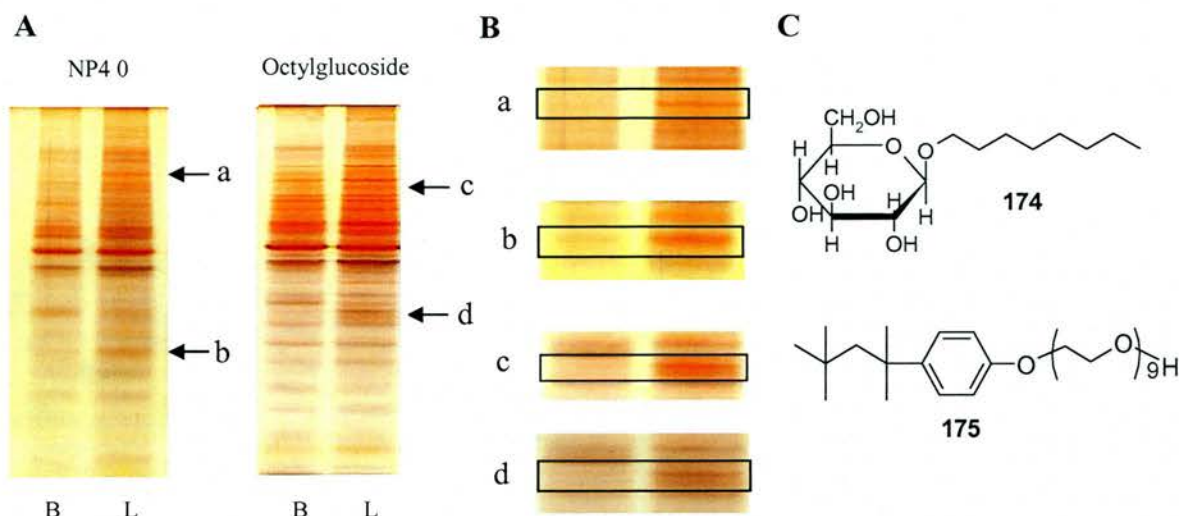


Figure 71. The different proteins extracted from the parasite with the two detergents octylglucoside (**174**) and NP40 (**175**). **A** 1D gels of the proteins extracted with octylglucoside (**174**) and NP40 (**175**), B = proteins identified from incubation with **193R**, L = proteins identified from incubation with **195R**. **B** Enlarged images of the bands observed from incubation with **195R**, which do not appear to be present on incubation with **193R**. **C** Structures of the two detergents used. Octylglucoside (**174**) and NP40 (**175**). The arrows indicate bands which appear to be present in the protein extract from the loaded resin but not the blank resin.

Large numbers of proteins bound to both **195R** and **193R**, consistent with a large amount of non-specific binding to the resin. The large amount of “back ground noise” generated by this non-specific binding is problematic and hinders affinity chromatography experiments. Identifying single proteins which are binding to **195R** and not **193R** is made more challenging the larger the number of proteins which have to be discounted. This difficulty can be observed qualitatively by examining the 1D-gels (Fig. 71). Those points highlighted on the 1D gels as appearing different between blank and loaded resin do not necessarily correspond to 1 protein. The large number of proteins on the gel means that the observed differences could be due to more than one protein and differences are almost certainly obscured.

Importantly, analysis of the final detergent wash (from step 4, Fig. 70), confirmed that no further proteins were being eluted. Assuming that **195R** does not actually have this many binding partners, these results suggest that the wash conditions themselves were not ideal. The elution conditions chosen must balance; i) the need to remove all the non-specifically binding proteins and ii) the avoidance of disrupting the compound-target protein(s) complexes. However, high levels of non-

specific protein binding have previously been reported in the literature in the context of successful affinity chromatography experiments.^{39,203}

Given the literature precedent, the observation of possible specific bands (Fig. 71), and the time constraints imposed on this section of the project, further preliminary investigations were attempted using these conditions. Any conclusions discussed from these experiments are made tentatively with the aim of highlighting areas of development for future affinity chromatography experiments.

The two different lysates generated using octylglucoside (**174**) and NP40 (**175**) indicate the significant impact that detergent selection can have. Therefore, future experiments will be conducted with both of these detergents in parallel.

(ii) Extracts were incubated with **194R**, **195R**, **196R** and **197R**. No discernable differences were observed between the proteins bound to the four resins. This would imply that loading levels may not have as dramatic effect as expected. However, given the large number of proteins eluted, these experiments were very difficult to interpret.

(iii) A series of different final elution (5, Fig. 70) conditions were investigated. Four of these are highlighted in figure 72. A clear difference in the number and quantity of proteins eluted from the resin was observed, reinforcing the importance of the selection of the final elution conditions.

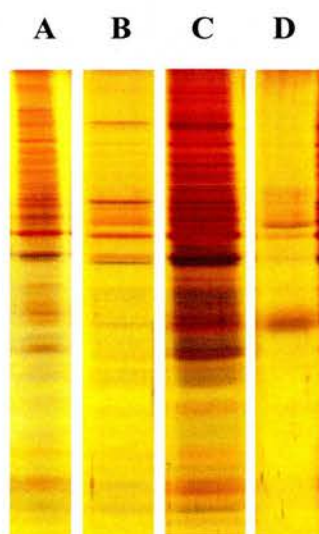


Figure 72. Varying the elution conditions. **195R** after incubation with *T. gondii* extract and washing with the buffered detergent solution was eluted with: **A** boiling in SDS, **B** acetonitrile, **C** triethylamine

and **D** glycine. The quantity of protein loaded on the gels for each elution condition was calculated to be equal (assuming the same amount of protein was eluted in each case).

(iv) **44**, **59**, **135** and **179** were used in two experiments: (i) preincubation of the *T. gondii* extract with each of the four compounds (ii) attempted specific elution of proteins from the extract bound to **195R** with solutions of the four compounds. Unfortunately the compounds were not sufficiently soluble under the experimental conditions to carry out these experiments effectively. This highlighted the need for improved control reagents (either compounds with improved solubility or control resins).

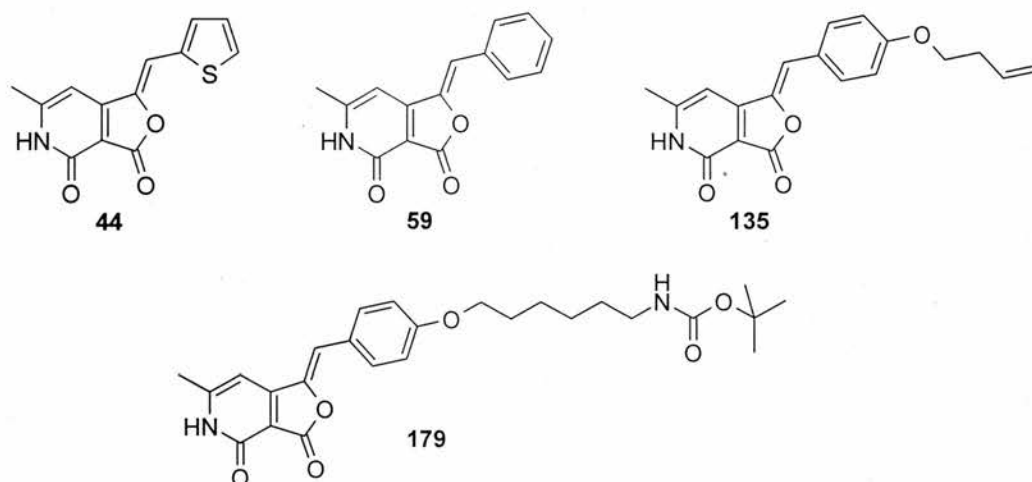


Figure 73. Compounds used as competitive enhancers in control experiments.

(v) While the protein target(s) of **44** from *T. gondii* extracts are of primary interest, the possibility that **44** may additionally target the host cell must not be forgotten. With this in mind, extracts from HFF cells were also incubated with **195R**.

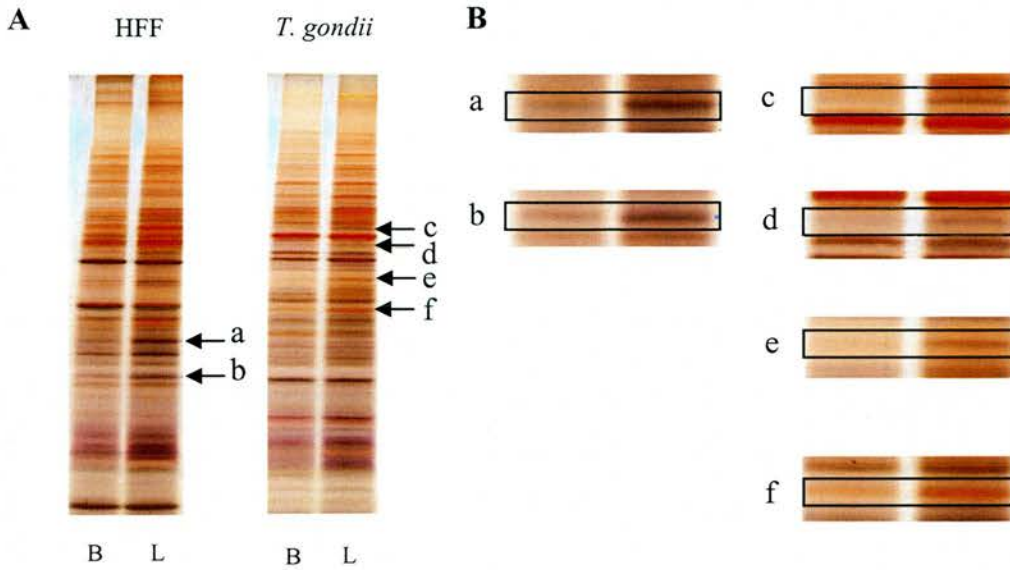


Figure 74. Proteins which appear to be binding to the active resin from both parasite and host cell extract. **A** Lanes 1 and 2 proteins from HFF cells, lanes 3 and 4 proteins from *T. gondii*. B =**193R**, L =**195R**. Arrows indicate the differences in observed bands. These gels are from the octylglucoside extract. **B** Enlarged images of the bands observed from incubation with **195R**, which do not appear to be present with the blank resin.

There appears to be a number of proteins which selectively bind to **195R** from the HFF lysate (Fig. 74). This suggests that **44** may have protein target(s) in the host cell which is interesting in light of the results briefly discussed in section 3.10 (fluorescence).

One clear conclusion that results from the experiments is that improvements are required in the following areas: reduction in the levels of non-specific binding, the method of washing used (step 4) and the controls used.

Other conclusions which can be made from these experiments are: (i) detergent selection is important; (ii) final elution conditions can dramatically affect the observed results; (iii) binding partners of **44** may be isolated from both *T. gondii* and host cells.

4.5 Developments for 2nd generation affinity chromatography experiments.

The experiments carried out at the University of Vermont highlighted areas in which improvements could be made to facilitate affinity chromatography experiments. Synthesis of either a control resin or a competitive control with greater solubility is required. Methods that reduce non-specific binding of proteins would also be expected to improve the clarity of the results.

4.5.1 Attempts to reduce non-specific binding

Different methods which could be used to reduce the level of non-specific binding include: (i) removing the major non-specifically binding proteins from the extract prior to the experiment (ii) altering the affinity reagent used (iii) increasing the stringency of the washing conditions.

4.5.1.1 Removing the non-specifically bound proteins from the protein lysate.

Initial incubation of the protein sample with a structurally related control resin should deplete the crude protein extract of those proteins which are non-specifically binding. The remaining proteins which do not non-specifically bind to the control resin could then be incubated with the active resin.

4.5.1.2 The use of compounds with differing distances from the resin

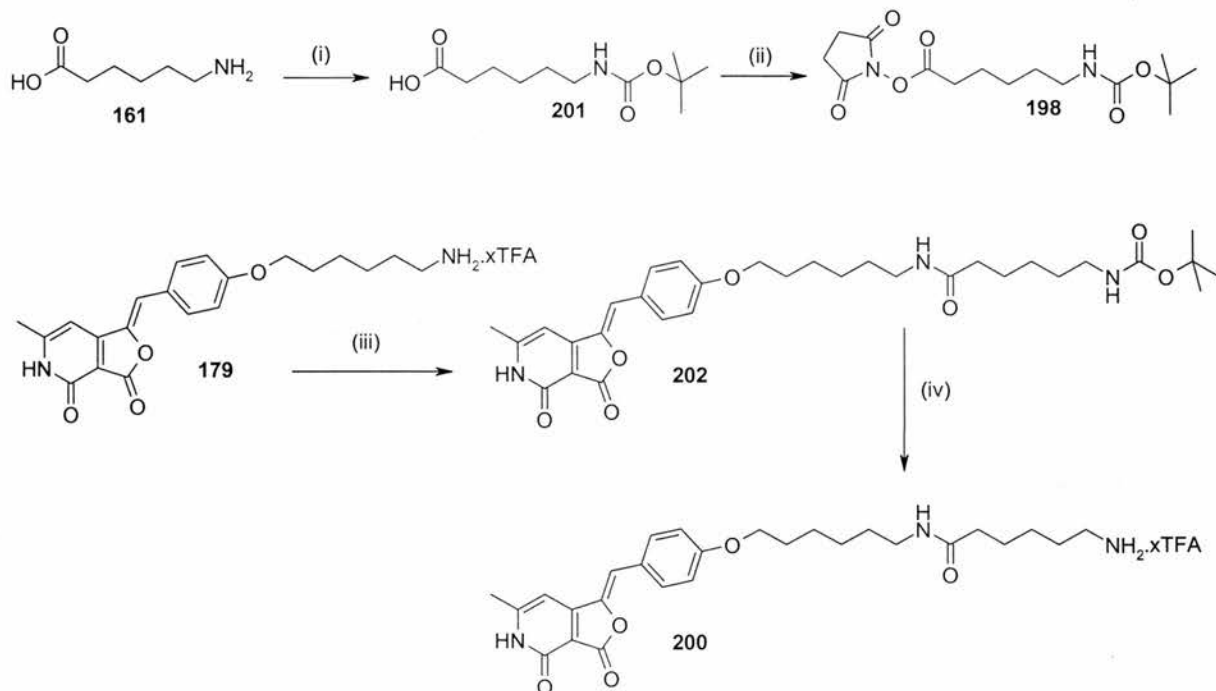
Non-specific binding can occur to both the compound and the unused or blocked loading sites on the resin. The further away the compound is from the resin the more likely the proteins are to interact with the compound versus other sites on the resin. Reduction of non-specific binding due to the resin could therefore be reduced through increasing the distance between the compound and the resin. Increasing this distance can be achieved through two approaches i) increasing the linker on the compound, ii) selecting a resin with an increased internal linker length.

The high levels of non-specific binding observed to the blank resin, **193R**, imply that the non-specific binding may be due to the resin not the compound. Selection of an alternative resin or solid support may therefore reduce this problem.

There are other supports available, which are not agarose based, such as: cellulose, glass, silica and polyethyleneimine.¹⁹⁹ However none of these supports have the same advantages of hydrophilicity, stability and commercial availability.¹⁹⁹ The only differences between the commercially available agarose resins is the agarose content or density, this is either at 4 or 6%. The smaller the percentage the greater the porosity and therefore the larger the ligands which can be accommodated. Given the previous success with the chemistry of the Affi-Gel resin (section 4.3.2.2.3), the advantages of agarose resins and the small differences in available agarose resins, all subsequent experiments used Affi-Gel resins.

4.5.1.2.1 Synthesis of a longer linker system.

Increasing the linker length has the potential to cause problems related to hydrophobic collapse or ion chelation (as discussed in section 4.2.1). Addition of units of 6-aminohexanoic acid is a proven method of increasing the linker length while avoiding such problems. Consequently **200** was selected as a synthetic target. **200** can be formed through the coupling of **177** with a derivative of 6-aminohexanoic acid (**161**). This amide bond can be formed through a number of different methods including synthesis of the acid chloride or fluoride of Boc-protected 6-aminohexanoic acid (**201**), use of carboimide coupling reagents (e.g. EDC, DCC or DIC), activation of **201** as its NHS ester or with 1,1'-carbonyl diimidazole (CDI) to form reactive imidazole carbamates. Selection of the method of coupling was driven by the insolubility of **177**. **177** is not sufficiently soluble in chlorinated solvents and the chemistry must be compatible with the use of DMSO or DMF as the solvent. It was therefore decided to prepare the NHS ester **198** (scheme 21).



Scheme 21. Synthesis of the active compound with an extended linker system. (i) di-*tert*-butyl dicarbonate, DCM, 0°C, 41% yield, (ii) DCC, N-hydroxysuccinimide, DCM, 79% yield, (iii) DCC, *tert*-butyl 6-[(2,5-dioxopyrrolidin-1-yl)oxy]-6-oxohexyl} carbamate, (**198**), TEA, DMF, 68%, (iv) TFA, >95%.

Synthesis of the NHS ester **198** was achieved in reasonable yield (scheme 21). Coupling of **198** with **177** under analogous conditions to those previously used to load resin occurred in good yield to give **202**. Subsequent deprotection of **202** yielded the target compound **200** which was then successfully loaded onto the resin in an analogous manner to **177**. As with resin **193R**, triethylamine was used to remove any remaining TFA. When coupled to Affi-Gel 10 resin (**193R**), this route provided a new resin **203R** which may lead to reduced non-specific binding of proteins. All of the above chemistry was carried out in the dark. Compound **202** was assayed in order to confirm that it retained activity. The results indicated that there was a further drop in activity LAC = 50 μ M when compared to **179** (LAC = 25 μ M).

4.5.1.2.2 Different internal resin linker lengths.

Alternative methods of increasing the linker length involve altering the internal linker present on the resin. So far the resin used has been Affi-Gel 10

(**193R**), this resin has a 10 atom linker unit. Another available resin is Affi-Gel 15 (**204R**) which has a 15 atom internal linker.

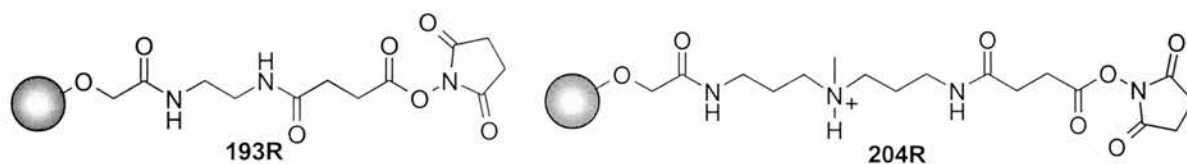


Figure 75. Different resins used with differing internal linker lengths. Affi-Gel 10 (**193R**) and Affi-Gel 15 (**204R**).

177 and **200** were therefore loaded on to Affi-Gel 15 (**204R**) to give resins **205R** and **206R** respectively. Four affinity chromatography resins have now been prepared with linker lengths of 17, 24, 22 and 29 atoms respectively.

The method used for loading Affi-Gel 15 (**204R**) was the same method as used for Affi-Gel 10 (**193R**). However, a decrease was observed in the efficiency of the loading. When attempting to load the Affi-Gel 10 (**193R**) resins at 20% it was possible to achieve 19% loading (section 4.3.2.2.3). In the case of Affi-Gel 15 (**204R**), the loading levels were in the range 14-17%.

4.5.2 Improvement of experimental controls.

There are two methods by which the controls for these experiments can be improved (i) the selection of competitive enhancers with improved solubility and (ii) synthesis of a control resin.

4.5.2.1 Competitive controls.

Access to enhancers with improved aqueous solubility can be achieved either through the use of the solubility information which has been obtained (section 3.8), or through the directed synthesis of compounds designed to have improved aqueous solubility.

Establishing the solubility of the derivatives synthesised (see section 3.8) enables selection of those derivatives which are more soluble in more aqueous environments for use as competitive enhancers. Out of the compounds synthesised

for the parallel synthesis two compounds **126** and **131** were observed to be completely soluble under the assay solution conditions. These compounds, an indole derivative (**126**) and the *p*-hydroxy derivative (**131**), will be used in any future competition experiments.

Another approach to obtaining competitive controls with improved solubility is through the synthesis of analogues where the functional groups of the aromatic ring have been selected to infer aqueous solubility. Taking this approach compounds **207**, **208**, and **209** were synthesised with the expectation that these would have improved aqueous solubility.

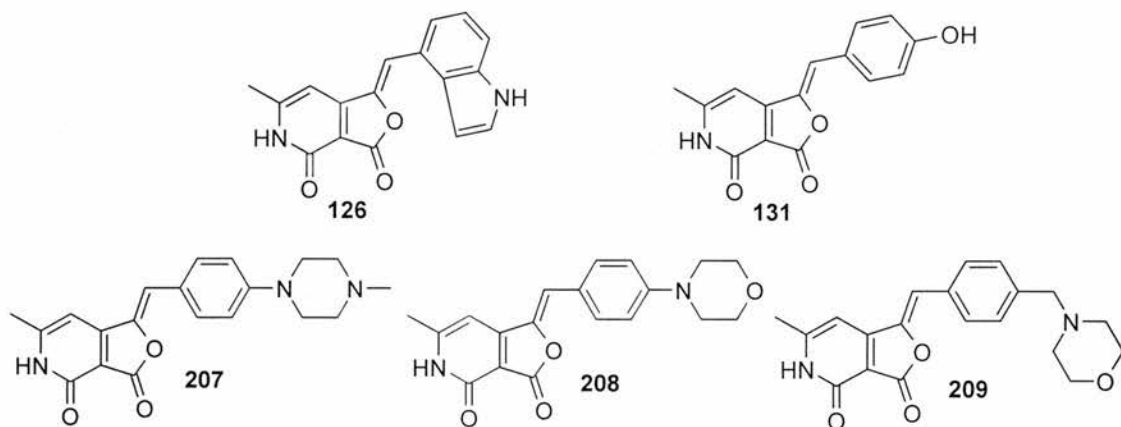


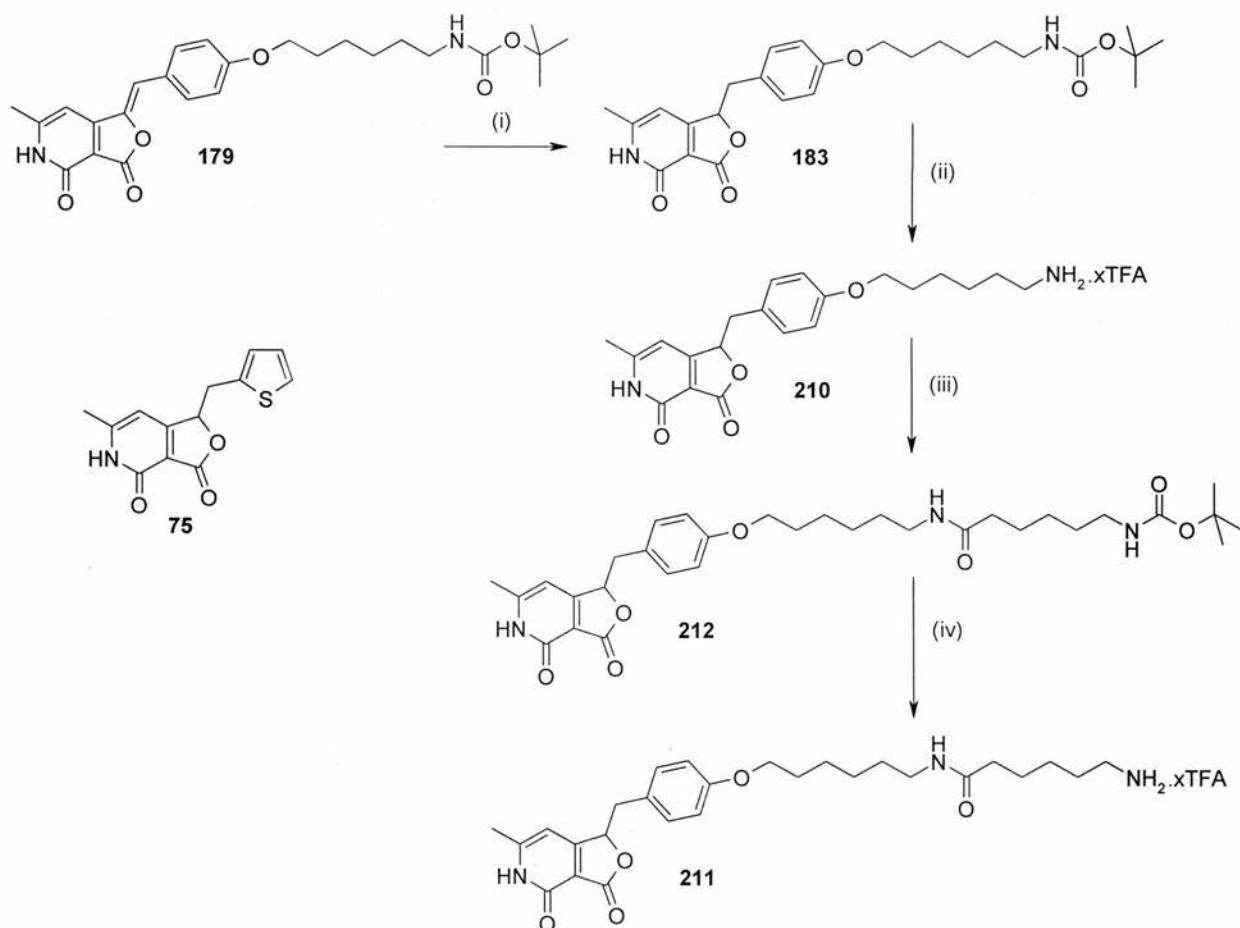
Figure 76. Compounds which can be used as competitive enhancers. Compounds **126** and **131** were identified from the parallel synthesis as being water soluble (see Chapter 3). Compounds **207**, **208** and **210** were synthesised in order to acquire derivative with good aqueous solubility.

Of these compounds, the morpholine derivatives **208** and **209** gave final concentrations of 84 μM , and 94 μM respectively (c.f. desired concentration of 100 μM), with **207** only reaching concentrations of 2 μM , most likely due to its very low solubility in DMSO (13mM c.f. desired concentration of 40 mM). Assaying these compounds revealed that **209** was active (LAC 12 μM), but surprisingly **208** was completely inactive. **207** was active (LAC 0.03 μM), with very good potency. It was concluded that **209** was the most appropriate compound to be used as a competitive enhancer.

4.5.2.2 Control resins

Control resins are loaded with inactive compounds that are structurally similar to the active compound (in this case **44**). Control resins are designed to enable the identification of proteins that non-specifically bind to the resin. Compound **183**, which has the exocyclic double bond reduced, has been shown to be inactive. It was therefore decided to prepare a control resin based on **183**.

The control resins were prepared using **220** and **211**, the reduced versions of **177** and **200** respectively. **220** was prepared by hydrogenation of **179**, followed by removal of the Boc group in **184** using TFA (scheme 22). A coupling procedure analogous to that used in the synthesis of **202** involving NHS-activated Boc protected 6-aminohexanoic acid **199** allowed for the successful synthesis of **212**, the protected precursor to **211**.



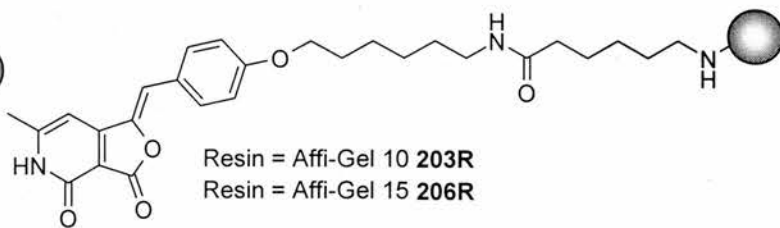
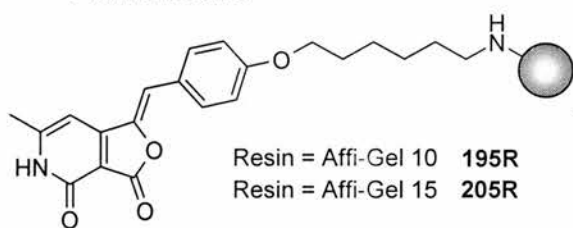
Scheme 22. Synthesis of compounds for a control resins. (i) H_2 , Pd/C, DMF, 72%, (ii) TFA, >95%, (iii) DCC, *tert*-butyl 6-[(2,5-dioxopyrrolidin-1-yl)oxy]-6-oxohexyl} carbamate (**199**), TEA, DMF, 54%, (iv) TFA, >95%.

Compounds **210** and **211** were loaded on to the resin using the same procedure as for **177** and **200**. Although **210** and **211** are white solids they do absorb UV light and the concentration of solutions containing these compounds could be determined using a UV spectrometer. The resin loading levels obtained for **210** and **211** were comparable to those obtained for **177** and **200**. Affi-Gel 10 (**195R**) gave loading levels of 19% for both compounds and Affi-Gel 15 (**204R**) gave 14% for **210** and 17% for **211** (intended levels 20%).

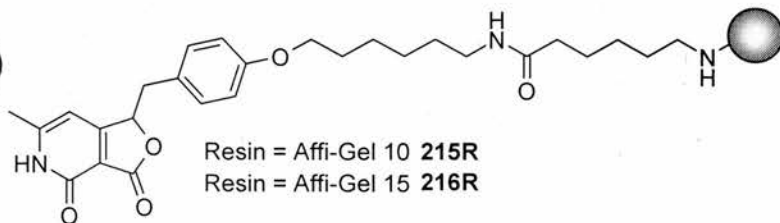
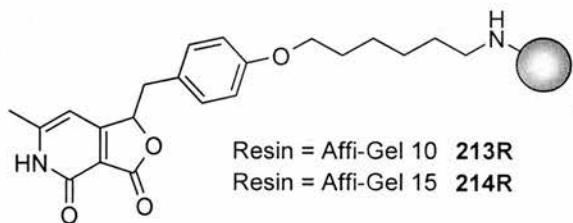
4.6 Chapter summary

Affinity chromatography resins were made using a derivative of **44**. These resins were used in initial affinity chromatography experiments. The experiments highlighted areas to be developed for further experiments. Consequently water soluble competitive compounds and control resins have been synthesised. Synthesis of resins with varying linker lengths has also been achieved in order to identify experimentally a linker length with the lowest level of non-specific binding. With four active and four control resins as well as competitive controls all the reagents required for a second generation of affinity chromatography experiments are in place. These experiments are currently ongoing in Prof. Ward's laboratory (see Chapter 5 for some preliminary results).

Active resins:



Control resins:



Competitive controls:

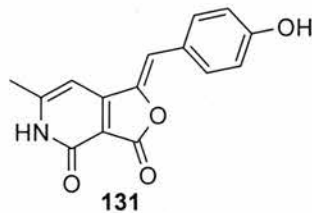
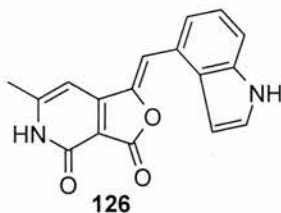
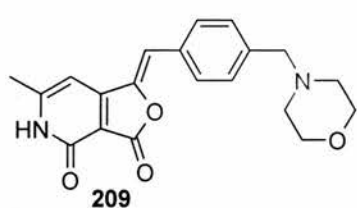


Figure 77. The reagents for affinity chromatography which have been synthesised. All the resins have intended loading levels of 20%.

Chapter 5: Affinity chromatography and covalent modification

This chapter describes investigations into the mechanism of action of **44**. Initial affinity chromatography results achieved with the resins described in Chapter 4 are also presented.

5.1 Covalent modification

The details of affinity chromatography experiments differ significantly depending on the mechanism of action of the compound. If the compound non-covalently binds to its cellular target(s) then the target(s) can be eluted off the resin in the final step of the procedure with ionic, acidic or basic solutions (chapter 4). However, if the compound covalently modifies its target(s) then the target(s) cannot be eluted off the resin. It is therefore very important to understand the possible mechanism of action of the compound of interest.

5.1.1 Known covalent modifiers

There are many examples of biologically active compounds which covalently modify their protein target. One of the best known examples is the β -lactam antibiotics, including penicillin G (**217**), which covalently bind to their target proteins, the penicillin binding proteins. The β -lactam ring has been shown to react with a serine residue of α -transpeptidase leading to covalent modification of the protein and inhibition of cell wall biosynthesis in bacteria.²⁰⁴

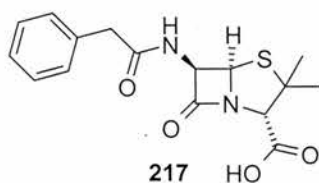


Figure 78. The β -lactam penicillin G (**217**).

Other drugs which covalently modify their cellular target include Acetaminophan **218**, Carbamazepine **219**, Diclofenac **220**, Halothane **221** and

Tamoxifen **222**.²⁰⁵ Possible cellular targets of these drugs (and their metabolites) have been identified through the use of with mass spectrometry, chromatographic separations and radiolabelling studies.²⁰⁵

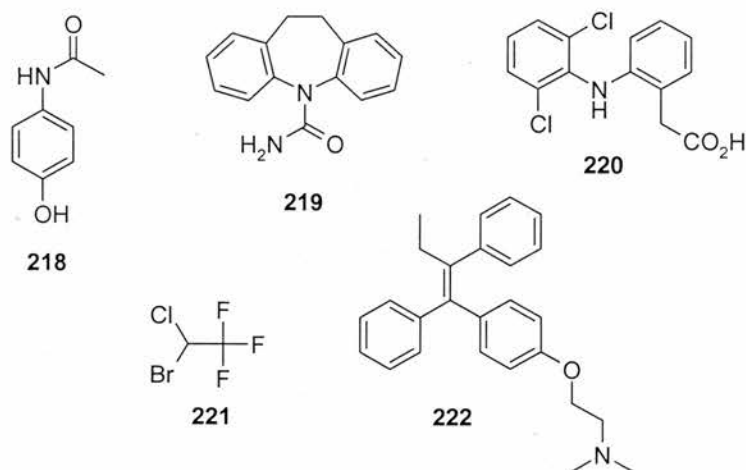


Figure 79. Drugs which have been identified as covalently modifying their cellular target.²⁰⁵ Acetaminophan **218**, Carbamazepine **219**, Diclofenac **220**, Halothane **221** and Tamoxifen **222**.

5.1.2 Target identification for covalent modifiers

Target identification techniques required for compounds which covalently modify their protein targets are different from those discussed in chapter 1 (sections 1.3.1 and 4.1).²⁰⁵ The differences can be as little as altering the stringency of the washing conditions or as significant as the use of an entirely new procedure. Covalent modifiers allow the use of experiments which are less feasible with non-covalent compounds.

5.1.2.1 Tagging

The importance of stringent and appropriate washing/elution conditions has already been highlighted (section 4.4). When a compound covalently modifies its protein substrate very harsh elution/washing conditions can be used without the fear of removing the target protein(s). This can be particularly important when taking a tagging approach (see section 1.3.1.3).

Fluorescently tagged compounds can be applied to a non-denaturing gel on which the crude protein extract is displayed. In order to identify successfully fluorescent spots/bands which contain the compound, there must be no residual fluorescence caused by insufficient washing of the excess protein extract. Therefore the harsher washing conditions allowed with covalent modifiers can prove particularly advantageous.

The use of the biotin/streptavidin binding pair has already been discussed (Chapter 1). Incubating a crude protein extract with a biotin tagged covalent modifier will allow labelled proteins to be easily pulled out using streptavidin resins. The labelled protein can then be identified using mass spectrometric techniques (see section 5.1.3).

5.1.2.3 Affinity chromatography

In general, affinity chromatography is not the target identification method of choice when dealing with covalent modifiers. Covalent binding of the protein to the resin means that removal of intact binding proteins is not possible. However resins can be incubated with trypsin which cleaves the proteins into peptide fragments which are released from the resin (Fig. 80). These peptide fragments can be identified through the use of mass spectrometry (see section 5.1.2.4) and therefore the identity of the protein established (see section 5.1.3).

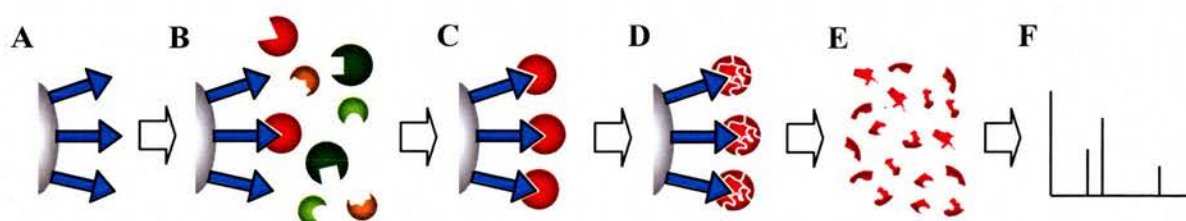


Figure 80. Schematic representing the approach taken to affinity chromatography when the compound covalently modifies its biological target. **A** Resin loaded with the active compound, **B** incubation of the resin with a crude protein extract, **C** target protein covalently binds to the resin, **D** the resin is incubated with trypsin, **E** the peptide fragments from the trypsin digest, **F** mass spectrometry is used to identify the protein.

5.1.3 Identifying the protein binding partners

Mass spectrometric techniques have been developed to allow the identification of peptide sequences. This is achieved through a technique known as tandem MS (also known as MS/MS or MS²).²⁰⁶ First the molecular weight of the peptide fragments (achieved from trypsin digest of the protein) is determined. An isolated peptide ion collides with an inert gas which causes the peptide to break apart. Mass spectra of the resulting fragments enables the peptide sequence to be determined. The data obtained from the mass spectra can then be used to search a sequence database using algorithms such as Mascot, Sequest and PeptideSearch.²⁰⁶ These compare the theoretical fragment sequences for proteins from the database to the fragments determined by mass spectrometry. The ability to identify the protein by mass spectrometry is not only dependent on the ability to gain the raw data but also the availability of the sequence data. The *T. gondii* protein sequence is not yet complete (approximately 95% assembled) and as such there will be a limit to the proteins which can be identified by mass spectrometry.²⁰⁷

5.2 Evidence for the covalent reaction of **44** with its protein target(s)

The affinity chromatography experiments which have been discussed so far (chapter 4) have assumed that compound **44** does not covalently modify its cellular target. However, experiments have shown that compound **44** can react with butylamine to give **93** (Chapter 2, section 2.3.1). Synthesis and analysis of a reduced compound **75** has also shown that the biological activity of this class of compounds is dependent on the presence of the exocyclic double bond (Chapter 2, section 2.2.3). It was proposed that the presence of the double bond in **44** would facilitate either the opening of the furanone ring with nucleophiles or allow conjugate addition at C8. In order to test the first hypothesis compound **75** was incubated with butylamine, under analogous conditions to the successful reaction of **44** to form **93**. No reaction was observed. The combination of these pieces of information suggests that there is a significant possibility that **44** covalently modifies its cellular target(s).

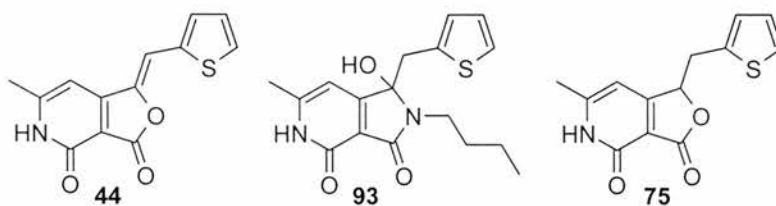
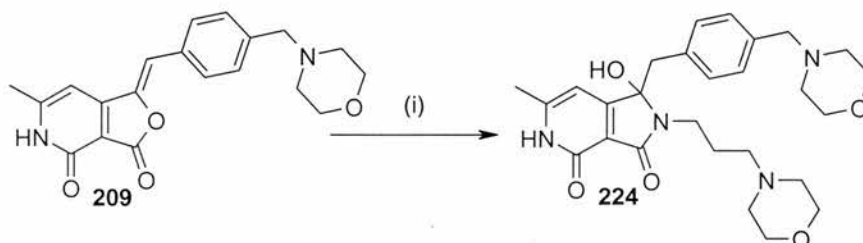


Figure 81. The structure of the original compound **44**, the product of **44** covalently modified with butylamine **93**, the reduced version of **44**, **75**.

These reactions of **44** and **75** with butylamine were set up in an organic solvent. This does not directly mimic the conditions of either the assay or the affinity chromatography experiments which are carried out under predominately aqueous conditions. To address this, analogues were synthesised with functional groups designed to increase the water solubility. Compound **209** was incubated with the water soluble amine 3-morpholino-1-propylamine (**223**) under analogous conditions to those used in the affinity chromatography, although a small quantity of DMSO needed to be added in order to ensure the compound was completely in solution (Scheme 23). The successful reaction of **209** with **223** to form **224** showed that these compounds can react with amines under aqueous conditions.



Scheme 23. Reaction of a water soluble derivative with a water soluble amine. (i) 3-Morpholino-1-propylamine (**223**), 150 mM KCl, 10 mM PIPES, 1.2% octylglucoside, 2% DMSO (as per the affinity chromatography experiments with DMSO added to aid solubility), 24%.

In addition to these arguments, a literature search revealed that compounds with a strained furanone unit similar to that in **44** have been shown to covalently modify their protein target.

Haloenol lactones such as **225** have been shown to be enzyme-activated irreversible inhibitors for serine proteases (Fig. 82).^{208,209} The initial step involves binding of **225** to α -chymotrypsin. This is followed by attack of the active site serine

at the lactone carbonyl carbon resulting in opening of the lactone ring. The final irreversible binding to the protein is caused by the displacement of the bromine atom.

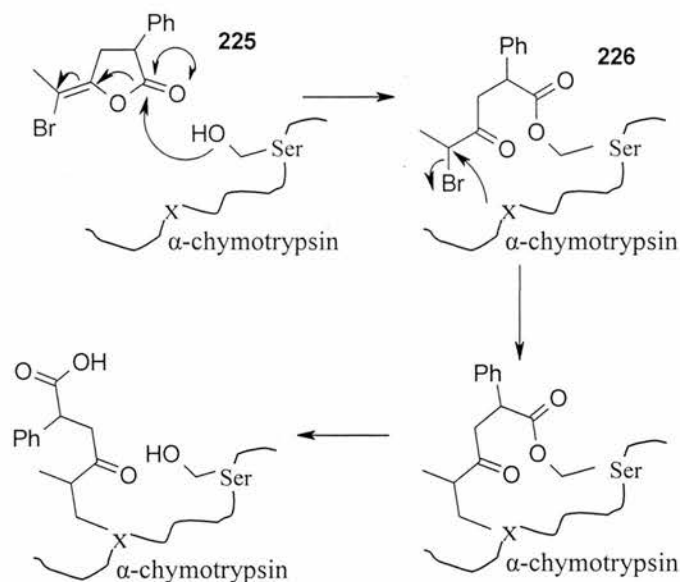


Figure 82. Ring opening of strained lactones with chymotrypsin. The mode of serine hydrolase inactivation by haloenol lactones. This process involves binding of **225** by ring opening of the lactone with a serine residue. Displacement of the bromine by a nearby nucleophilic residue results in the covalent modification of the protein. Figure is reproduced from Daniels *et al.*, *Biochemistry*, **1986**, *25*, 1436-1444.²⁰⁸

The possibility that the enzyme inhibition was due to formation of a stable acyl enzyme-inhibitor complex, **226**, was also investigated. It was discovered that although the first step was essential it was not the step which inactivated the enzyme. However, this example provides further evidence that strained lactones react with nucleophilic residues in proteins (eg. serine).

Compounds containing α,β -unsaturated carbonyl units have also been shown to be biologically active. They exhibit their mode of action through conjugate addition of a protein nucleophile.²¹⁰ Compound **227** has been shown to increase the induction of quinone reductase (QR) in Hepa 1c1c7 murine hepatoma cells.²¹⁰ The mechanism of which has been predicted to be through conjugate addition to the α,β -unsaturated system.

The reaction between enones and benzenethiol has been studied in great detail.²¹¹ The position of the equilibrium between the enone (**227**, **228**, **229**) and the addition product (**230**, **231**, **232**) was investigated for three compounds (Fig. 83) in

chloroform. The measured equilibrium constants ranged from $K=3.9 \times 10^3$ for **227** to $K=0.9$ for **229**. When there is a methyl group in position 3, compound **229**, then the equilibrium is such that there is an equal amount of **229** and **232**.

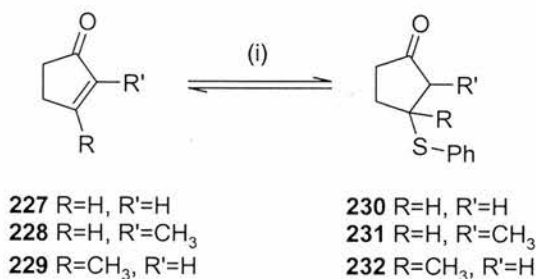


Figure 83. Michael-type addition of benzenethiol to 2-cyclopenten-1-ones. (i) benzenethiol, triethylamine, chloroform.²¹¹

Should compound **229** react with a cysteine residue on a protein in an analogous manner to how it reacts with benzenethiol then the position of the equilibrium will effect whether **229** binds irreversibly. If the protein-compound complex is washed then the equilibrium will be pulled towards the formation of the enone. This will result in the compound binding reversibly, despite the conjugate addition reaction forming a covalent bond between **229** and the protein.

The two literature examples show that attack of a nucleophilic residue of a protein can react with both lactones and α,β -unsaturated carbonyl systems both of which are present in **44**. This would involve attack of a nucleophile at two different positions i) at the carbon of the carbonyl group in **93** and ii) at C7a in one to give **233**, another possibility is an attack at C8 position to give **234**. Interestingly, reaction of **44** with butylamine resulted in the formation of **93** rather than either of the possible alternative products **233** or **234**.

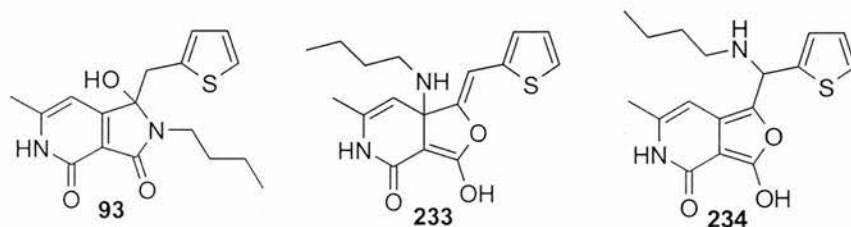


Figure 84. Possible products of compound **44** with butylamine. Compound **93** is the actual product of **44** with butylamine, compounds **233** and **234** are the possible products of a Michael-like addition of butylamine to **44**.

The majority of discussions relating to the covalent modification of a protein by **44** have focused on the attack of a nitrogen based nucleophile (page 139). However, attack of sulfur or oxygen based nucleophiles is also possible.

The nature of the nucleophile attacking **44** will have a significant impact on the position of attack. As discussed previously, a nitrogen based nucleophile is a hard nucleophile and would therefore be expected to attack at the furanone carbonyl carbon C3 as observed with butylamine. However, a sulfur based nucleophile is soft and will preferentially attack soft electrophiles. Sulfur based nucleophiles are therefore more likely to form structures such as, **237** or **238**.

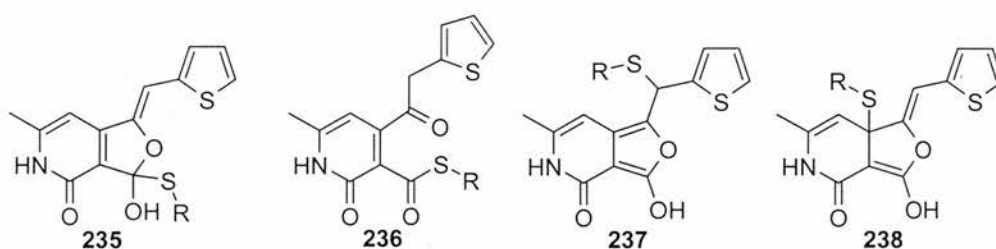


Figure 85. Possible products of **44** with sulfur based nucleophiles.

It is worth remembering that conjugate addition of a sulfur nucleophile is likely to be reversible which could potentially coincide with the observed biological activity of **44** although it remains unclear as to whether reversibility in a cell-based assay correlates directly with reversible binding to a single protein target.

Interestingly, the reaction of a compound containing a strained furanone, (*Z*)-ligustilide (**239**), with both nitrogen and sulfur based nucleophiles has been investigated previously.²¹² Beck *et al.* found that **238** reacted with benzylamine to form **240**.²¹² This is consistent with our observation of the formation of **93** from **44** on reaction with butylamine.

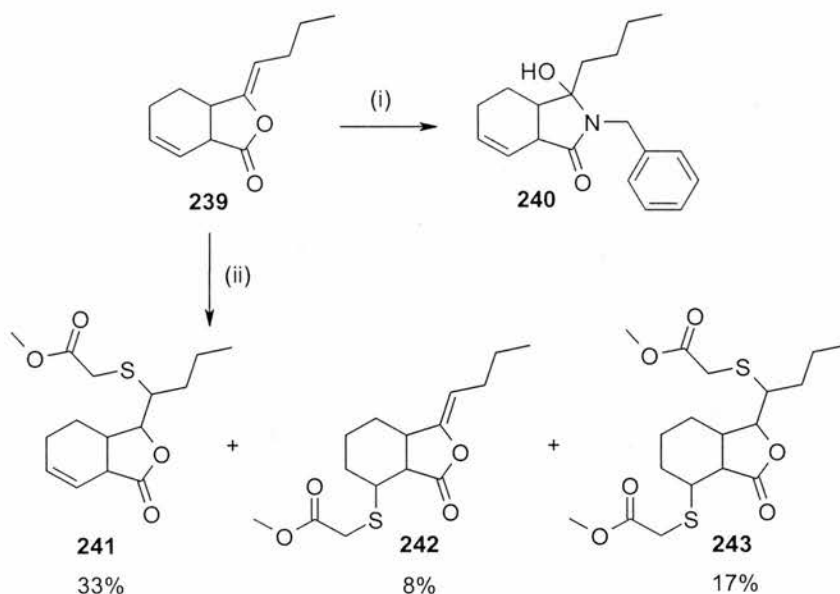


Figure 86. The reaction of (*Z*)-ligustilide (**239**) with nucleophiles.²¹² (i) benzylamine, (ii) methyl thioglycolate, triethylamine.

In addition when Beck *et al.* incubated (*Z*)-ligustilide (**239**) with a sulfur based nucleophile methyl thioglycolate, they observed the formation of three products **241**, **242** and **243** (Fig. 86) with the majority of the compounds resulting from 1,6-conjugate addition. The 1,6-conjugate addition at C8 to form **241** implies that the product of **44** with a sulfur based nucleophile would be **237**.

There has been an article published recently which discussed the use of ¹³C NMR to show the electrophilicity of the carbonyl and the likelihood of a conjugated carbonyl undergoing: conjugate addition of a nucleophile, reduction or photoisomerisation.²¹³ They used the carbon shift of the β-carbon to indicate which one of these reaction would be most likely to occur. Examining the carbon shift of the β-carbon in **44** showed it not to be in their range of investigation (118-139 ppm). This indicates that attack at position 7a to yield **233** is unlikely to occur. However, their investigations only focused on data from three scaffolds: pyrazolinones **243** and **244**, and benzothiazinones **245**. This relative lack of diversity of structure could explain why the carbon shifts of **44** do not fit.

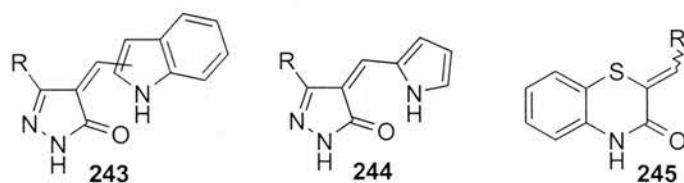


Figure 87. Scaffolds use to categorise the ^{13}C shift of the β -carbon of a α,β -unsaturated carbonyl system.²¹³

Based on this discussion, within a biological system covalent modification of a number of different amino acid residues, for example: the nitrogen based lysine, the sulfur containing cysteine and the oxygen centred serine could occur. These would be expected to react with **44** differently, however, the products from such reactions would induce the same protein mass increase of 259 Da regardless of structure. This means that preliminary target validation mass spectrometry experiments involving incubating **44** with an identified protein binding partner should lead to a protein mass increase of 259 Da, although the details of the chemistry cannot be inferred from this result.

Investigations were initiated to examine the likelihood of covalent modification of **44** by each of these three types of nucleophiles (nitrogen, sulfur and oxygen based) causing the observed biological activity. So far investigations have centred round the nitrogen centred nucleophile butylamine, which potentially mimics lysine.

5.3 NMR investigations into the mechanism of action of **44**

If **44** covalently modifies its cellular target then it is possible that the biological reactivity of the compound should be directly related to its chemical reactivity towards a nucleophile, assuming that little or no form of protein-ligand recognition is involved. An experiment was therefore derived to test this involving the use of ^1H NMR spectroscopy to assess the relative rates of reactions of analogues of **44**. The previous biological and physical characterisation of the collection of 22 compounds was used to help plan these experiments. Three selected analogues were

mixed together in DMSO- d_6 and their reactions with increasing amounts of butylamine were tracked using ^1H NMR techniques. The order of reactivity was then compared with the order of biological activity. Each set of three compounds was selected in such a way that the compounds within a set had analogous solubility but differing biological activity. Unfortunately due to the restrictions on solubility, it was not possible to select a number of sets of compounds with very distinct potencies. However six sets were chosen the majority of which had a near 4 fold difference between the most and least potent.

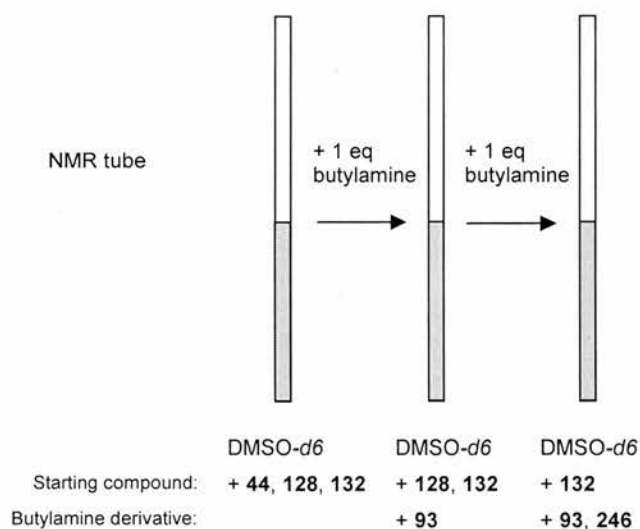


Figure 88. Schematic showing the procedure of the NMR experiment.

One equivalent of each of the compounds was dissolved in DMSO- d_6 in an NMR tube and one equivalent of butylamine was then added (Fig. 82). The ^1H NMR was recorded at various time points before a second equivalent of butylamine was added.

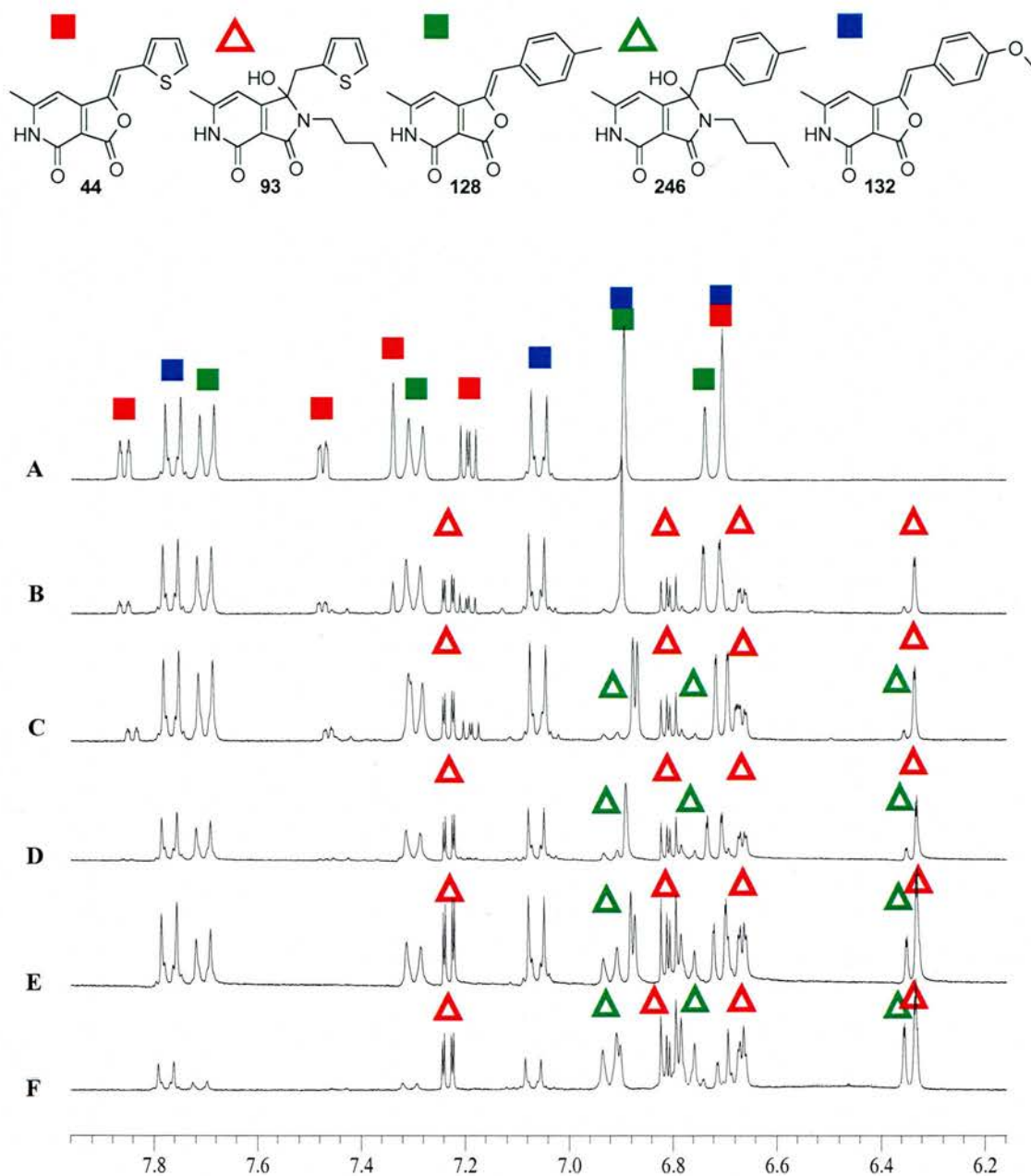


Figure 89. ^1H NMR competition experiment with three compounds **44**, **128** and **132**. **A** The first ^1H NMR spectrum of the three compounds before the addition of butylamine, **B** after the addition of 1 equivalent of butylamine, compound **44** is the first to react, **C** 12 hrs after the first addition almost all of **44** has reacted to form **93**, **D** after the addition of a 2nd equivalent of butylamine compound **128** begins to react, **E** 12 hrs after the second addition the amount of **246** has increased although there is some **128** remaining, **F** after the addition of the third equivalent of butylamine **128** has been converted to **246** and there is no initial reaction of **132**. The order in which the compounds reacted is therefore **44**, **128**, then **132**.

Figure 89 shows the results of one of these experiments. The three compounds **44**, **128** and **132** (Fig. 89) were selected as they had similar solubility and different biological activities, (2 μ M, 4 μ M and 6 μ M respectively). The results showed that **1** reacted first with butylamine, followed by **128**. **132** was the least reactive towards butylamine. Unfortunately, the relative order of reactivity in this case did not coincide with their biological activity. **44** was the most biologically active, however **128** was the least biologically active. The correlation between the reactivity of the compounds towards butylamine and their biological activity for all six set of compounds is outlined in table 7.

Set	Compound	Biological activity μ M	Order of reaction with butylamine
1	44	2	1
	132	4	3
	128	6	2
2	44	2	1
	132	4	3
	131	7	2
3	44	2	1
	132	4	2
	138	4	3
4	126	0.7	3
	135	7	2
	136	Weak	1
5	139	1	2
	122	2	3
	125	Weak	1
6	122	2	3
	129	3	2
	125	Weak	1

Table 7. The results of the competition NMR experiments.

Examining the six sets of compounds shows that out of the six sets in only one of them did the order of the reaction with butylamine mirror the order of biological activity. The most indicative results were those for sets 4, 5 and 6 where the compounds with the weakest activity **136** and **125** were the first to react with butylamine.

The aim of this experiment was to test the following hypothesis: the biological activity of analogues of **44** mirrors their reactivity towards nitrogen nucleophiles. This experiment does not support this conclusion. The results of the experiment indicate that the difference in biological activity observed between these compounds is not due entirely to their reactivity towards a nitrogen based nucleophile. This indicates one of three things i) the mode of action of these compounds is not covalent modification, ii) there are other factors such as a recognition elements, or cell permeability which contribute to the difference in biological activity or iii) the covalent modification of the compounds is occurring through a different mechanism (e.g. 1,6-conjugate addition of a sulfur based nucleophile).

Whichever of these rationales is correct the experiment has shown that these compounds are unlikely to unselectively covalently modify any random cellular target through ring opening of the furanone. This indicates that there is likely to be a specific cellular target(s) which these compounds bind to with some level of recognition. Therefore target identification studies should be able to identify a cellular target.

5.3.1 The use of IR to verify the outcome of the NMR investigations

Another method of gauging the reactivity of these compounds is to use the stretch in the IR spectrum which is assigned to the furanone carbonyl. The higher the stretch (in wavenumbers) the greater the reactivity of the carbonyl.²¹⁴ Therefore the relative reactivity of the furanone functionality of the compounds towards butylamine can be assessed, through examining the IR spectra.

Set	Compound	Biological activity μM	IR stretch of the furanone carbonyl	Order of IR stretch	Order of reaction with butylamine
1	44	2	1790	2	1
	132	4	1792	1	3
	128	6	1784	3	2
2	44	2	1790	2	1
	132	4	1792	1	3
	131	7	1784	3	2
3	44	2	1790	2	1
	132	4	1792	1	2
	138	4	1784	3	3
4	126	0.7	1748	3	3
	135	7	1787	2	2
	136	Weak	1789	1	1
5	139	1	1752	3	2
	122	2	1763	2	3
	125	Weak	1782	1	1
6	122	2	1763	3	3
	129	3	1783	1	2
	125	Weak	1782	2	1

Table 8. The comparative IR stretches for the furanone carbonyls.

The IR stretches which appear to be consistent with the ^1H NMR results are sets 4, 5 and 6, as the differences between the order of reaction of **139** and **122**, and between **129** and **125** are only slight (see appendix for NMR).

Although there is not a precise correlation between the order of reaction of the compounds with butylamine and the stretch in the IR signed to their furanone carbonyl, a similar trend is observed. More significantly the trends in the differences of the IR as stretches does not correlate with the trends in biological activities, thus reinforcing the conclusions from the NMR experiment.

Having investigated the possibility that **44** is reacting with a nitrogen based nucleophile through ring opening of the furanone ring, the next step was to investigate the reaction of **44** with either sulfur or oxygen based nucleophiles. Unfortunately, this works has not been completed due to time constraints and is on going in the Westwood laboratory.

Given that some of the experiments performed so far have suggested that **44** may be covalently modifying its biological target, it was decided that the overall approach to affinity chromatography experiments should include this possibility. This is described in the next section.

5.4 Covalent affinity chromatography experiments

Affinity chromatography experiments used resin **195R** (see Chapter 4) and **213R** as a control. The experiments were carried out at the University of Vermont by Prof. G. E. Ward and W. Dotzler. These resins were incubated with parasite extract for 1 hour, washed with buffer then washed repeatedly in boiling SDS. This was designed to remove all the proteins non-covalently bound to the resin. The resins were then washed with an aqueous solution (50 mM Tris, pH 7.4, 0.05% sodium azide) in order to remove any traces of detergent.

Detergents are not compatible with mass spectrometry as they disrupt the ability of the peptides to ionise and the signals produced by the detergent have the potential to swamp the signals for the peptide fragments.²⁰³ The resins were then sent to the University of St Andrews for analysis by mass spectrometry.

Suspending the resins in ammonium bicarbonate buffer insured the resins were at the correct pH to use trypsin. Trypsin was added to the resins and left overnight. Alcohol dehydrogenase had been previously added to the suspension before incubation with trypsin to act as an internal control for the trypsin digest. The supernatant was then removed and analysed by mass spectrometry.

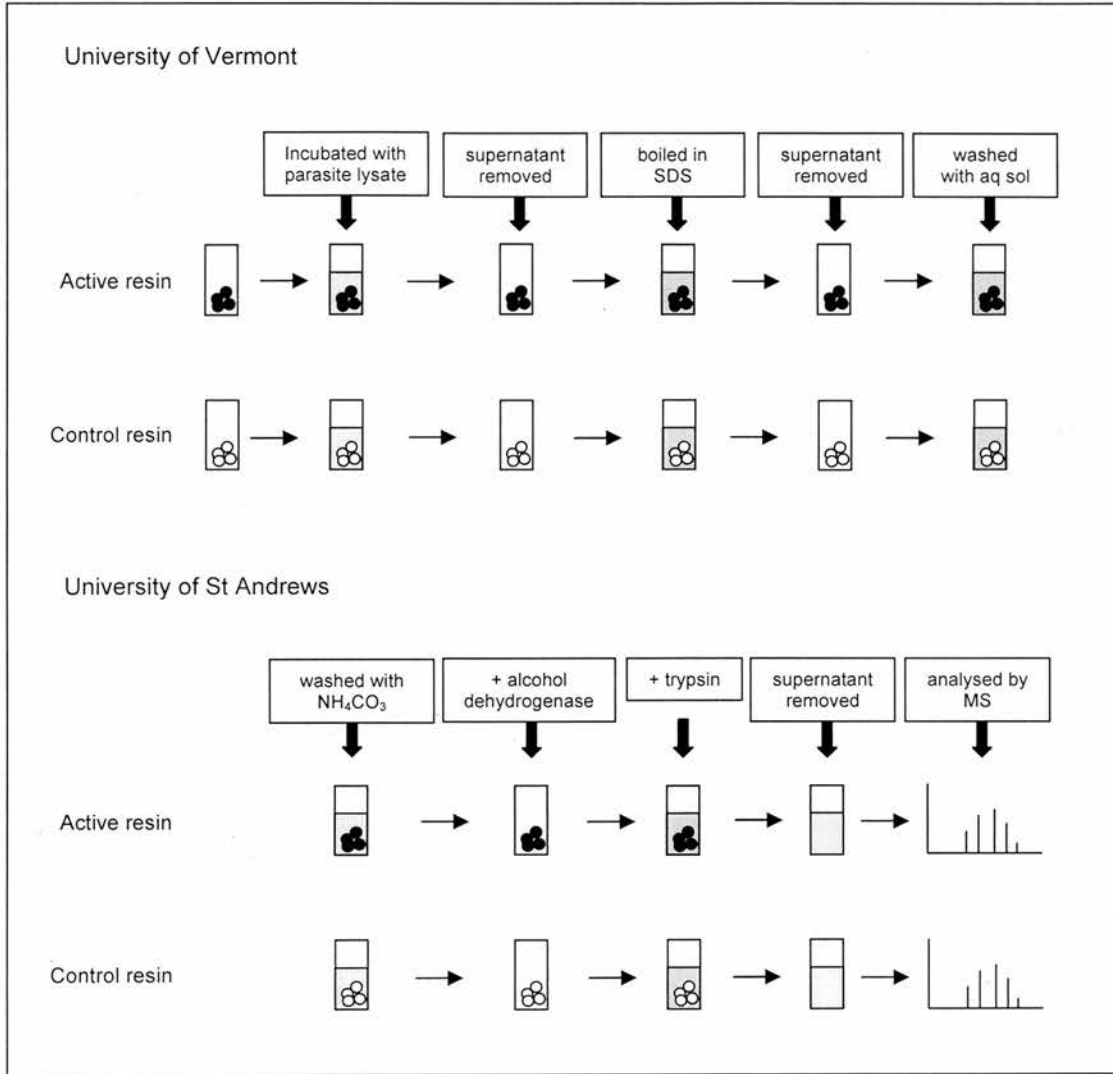


Figure 90. Covalent affinity chromatography experiments. At the University of Vermont the beads were incubated with parasite lysate. The supernatant was removed and the beads boiled in SDS. The beads were then washed with an aqueous solution of 50 mM Tris, pH 7.4, 0.05% sodium azide. At the University of St Andrews the beads were washed with ammonium carbonate. Alcohol dehydrogenase was added, followed by trypsin. The beads were incubated in trypsin over night. The resultant supernatant was removed. In experiment 1 this was analysed by mass spectrometry directly. In experiments 2 and 3 the supernatant was taken up in acid and then analysed by mass spectrometry.

This experiment was performed twice with slight variations between each experiment. The first experiment was performed as above. The supernatant from the second experiment was taken up in acid before analysis (in order to aid the chromatography).

The peptide fragments produced from the trypsin digest were analysed on a Q-STAR using LCMSMS analysis (see section 5.1.3). A large number of protein were identified as binding to **213R** (>100 proteins), however significantly less

proteins were found to bound to **195R** (~19) (see Appendix 3 for a full list of the proteins identified in all cases). A number of proteins were identified as binding to both **195R** and **213R**. Hypothetical protein pdi (Protein disulfide isomerase precursor), Non-transmembrane antigen, putative ribosomal protein S2 and an ADP/ATP carrier from *T. gondii* were amongst these. Identification was carried out by matching multiple peptide fragments (in the range of 6-25 peptides), which gave overall sequence coverages of 9-38%. Additionally actin from *Plasmodium yoelii* was identified. It is currently assumed that this is actually actin from *T. gondii* but that the sequence data are not in the database used (therefore *P. yoelii* is the closest match). A reanalysis of the data using a recently obtained version of the *T. gondii* genome is in progress.

There were only a few proteins which bound to the **195R** and not **213R**. The number of proteins range from 3 to 10 (see Appendix 3). These proteins were identified through peptide matches of 2 to 7 peptide fragments for each protein, with sequence coverages ranging from 5 to 20 % which is reflected in their scores (see Appendix 3). Only one of these proteins was identified as a *T. gondii* protein. *T. gondii* inflammatory profilin was identified as binding to the active resin **195R** and not the control **213R** in one of the two experiments (experiment 1). The match for the protein is a good match, with two peptide sequences identified in experiment 1 and three in experiment 3 (see below), giving sequences coverages of 15 and 26% respectively (see Appendix 3).

```

1 MSDWDPVVKE WLVDVTGYCCA GGIANAEDGV VFAAAADDDD GWSKLYKDDH
51 EEDTIGEDGN ACGKVSINEA STIKAAVDDG SAPNGVWIGG QKYKVVVRPEK
101 GFEYNDCTFD ITMCARSKGG AHLIKTPNGS IVIALYDEEK EQDKGNSRTS
151 ALAFAEYLHQ SGY

```

Figure 91. Peptide fragments matched to inflammatory profilin. The matched fragments are labelled in bold.

This means that of the peptide sequences identified from the extract some had very close identify to those theoretically expected from profilin. Inflammatory profilin is a small protein so only a few peptide sequences were found, which is evident by the comparatively high sequence coverage.

The experiment was repeated a third time, this time optimised for the identification of profilin (increasing the chances of getting an identification from the

control). The programming of the Q-STAR was varied in two ways. First, the peptide masses of six of the most abundant peptides were selected to be excluded. This was to decrease the amount of time spent identifying these peptides and so allow more time for the identification of other peptides. Second, the peptides from profilin were deliberately searched for. Profilin was again identified as the only *T. gondii* protein which bound to **195R** and not **213R**. Excitingly, this protein could therefore be the cellular target of **44**.

5.5 Inflammatory profilin as a possible target.

T. gondii inflammatory profilin (TGIP), also known as profilin *T. gondii* (PFTG), has been linked with its role in the immune response to *T. gondii*.¹⁰⁰ The protein has been successfully isolated from *T. gondii*.¹⁰⁰ Homologs have been identified in other apicomplexans including: *C. parvum*, *P. falciparum* and *E. tenella* (with homologies to *T. gondii* observed between 50 and 70%).¹⁰⁰ The exact function of the protein is not known but it has been predicted to have an actin binding site consistent with the role of profilin in other species.⁹⁹ Affinity chromatography experiments to determine actin binding proteins in *P. falciparum* identified, among other proteins, a 13 kDa protein which was postulated to be profilin.²¹⁵ This ability to bind actin suggests a possible link to motility and invasion although none has been reported to date for apicomplexa. In fact in the recent Science paper discussing the immune response to profilin the authors hypothesised that “profilins predicted actin-binding activity suggests that, like flagellin, they maybe involved in parasite motility and/or invasion”.¹⁰⁰ Should this prove to be true this is particularly exciting as it correlates very well with the phenotype observed for **44**.

5.5.1 Mammalian profilin

The role of mammalian profilin is relatively well understood. It is briefly review here to explore the possibility that the role of profilin in humans may provide some insight into its role in *T. gondii* and help evaluate profilin as a viable target for **44**.

Human profilin 1 and 2 are almost superimposable, however they are thought to have different properties.²¹⁶ Profilin 1 is found in the majority of cells in the human body, whereas profilin 2 is found only in the developing nervous system. Despite the possible variations in the protein ligands across species, all profilins are known to bind actin.

5.5.1.1 Profilin actin-binding activity

Profilin actin-binding activity has been shown to be important to the actin polymerisation process. Profilin binds to the actin monomer G-actin increasing the rate of nucleotide exchange. The profilin-actin-ATP (profilin bound to actin is also known as profilactin) complex can then bind to the growing end of the actin filament, where upon the actin is released by the profilin (Fig. 92).

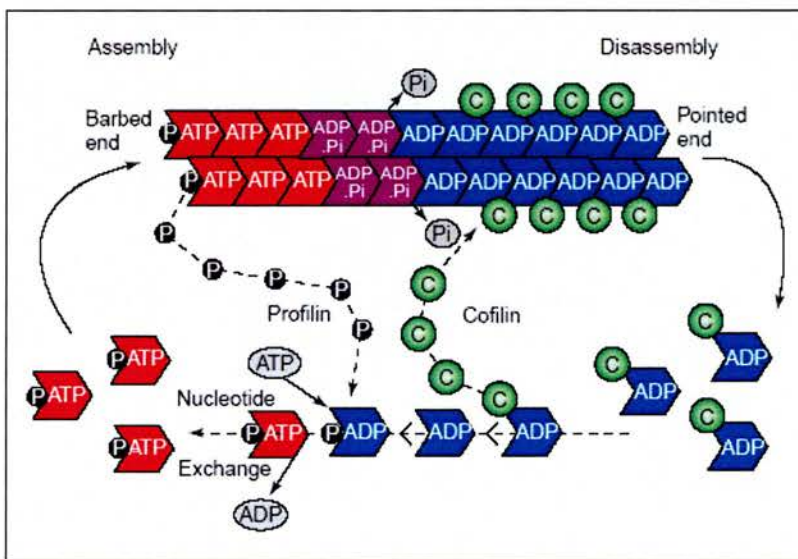


Figure 92. The role of profilin in actin polymerisation. Unpolymerised G-actin binds to profilin which increases the rate of nucleotide exchange and results in the generation of ATP-actin (from ADP-actin). The profilin-actin-ATP complex binds to the growing end of the actin. Overtime the ATP is hydrolysed to ADP. Cofilin can then bind to ADP-actin and increase the rate of dissociation from the actin filament and the process begins again. Figure reproduced from M. R. Holt *et al.*, *Trends Cell Biol.*, **2001**, *11*, 38-46.⁹⁹

Thus profilin is a key component to the actin polymerisation process and as discussed previously, actin has been identified as important factor in *T. gondii* motility (see section 1.5.2.1).

5.5.1.2 Other binding proteins of profilin

Proline-rich ligands have been implicated in facilitating the binding of profilactin to the cell membrane. Proteins such as WASP and ezrin have proline rich domains and have been shown to bind to both profilin and integral membrane proteins (such as Cdc42 and CD44 respectively).⁹⁹ This provides a method of localising actin polymerisation to a particular region.

The diverse number of ligands which bind profilin through the proline-rich binding site have led to profilin being identified as involved in various cellular process including: membrane trafficking, neuronal function and the Rac-Rho signalling pathway.²¹⁶ Although the exact nature of the binding partners of profilin and its role in these processes is not fully understood it is clear that profilin is important for a number of different reasons. This multifunctional capacity of profilin means that disrupting profilin would be expected to produce multiple phenotypic changes (perhaps consistent with **44** affecting multiple secondary assays).

5.5.1.3 The crystal structure of human profilin

The crystal structure of profilin and profilin complexes have visualised the binding site of actin (Fig. 93). The proline-rich binding site of profilin has also been identified with the two main binding sites being removed from each other.

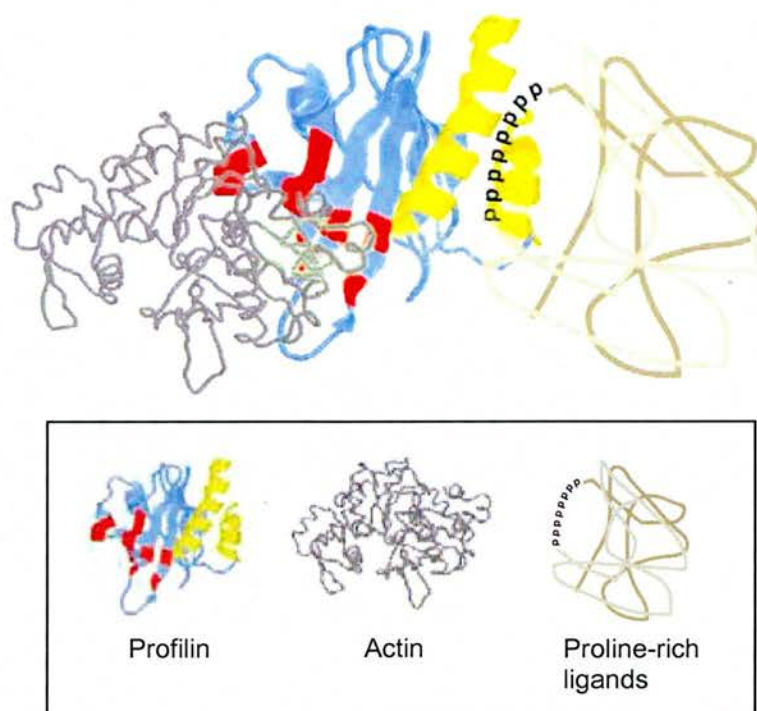


Figure 93. Annotated crystal structure of profilin. Profilin binds actin (red) and proline-rich ligands (yellow) the areas of binding are highlighted on the crystal structure. Figure taken from W. Witke, *Trends Cell. Biol.*, **2004**, *14*, 461-469.²¹⁶

The crystal structure of human profilin allows the amino acid residues at the different reactive sites to be determined. Given that the predicted mechanism of action of **44** involves attack of a nucleophilic residue onto the furanone it may be possible to examine the active sites to see if there are any residues suitable for such an attack.

5.5.1.3.1 The actin binding site

Examining the actin binding site shows that there are three main types of residues phenylalanine, arginine and aspartic acid. None of these residues would appear to be likely candidates for nucleophilic attack.

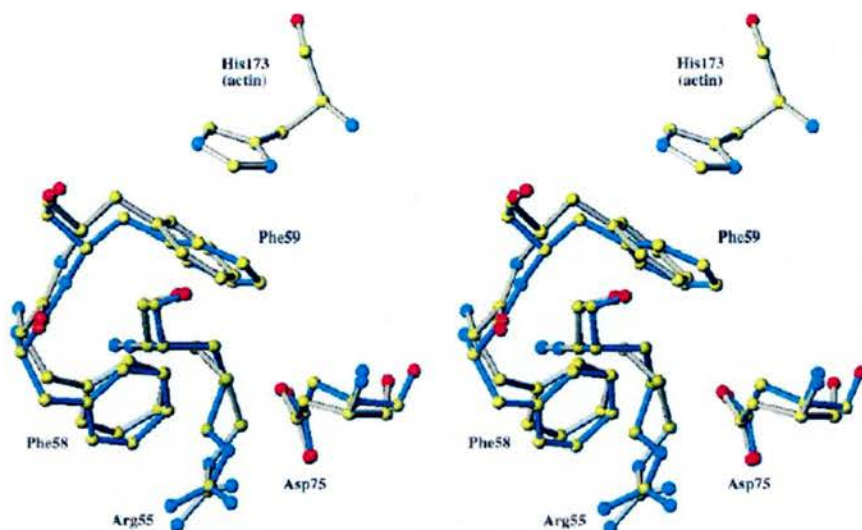


Figure 94. Structurally conserved residues at the actin binding site. Human profilin II (blue) and bovine profilin I bound complexed to β -actin (grey) superimposed. The Phe, Arg, Asp residues appear to be conserved. The His from actin stacks with the Phe of the profilin. Figure taken from Nodelman *et al.*, *J. Mol. Biol.*, **1999**, 294, 1271-1285.²¹⁷

5.5.1.3.2 Proline-rich binding site

The residues of the proline binding site of profilin I which interact with proline are Trp3, Tyr 6, Trp31, His 133 and Tyr 139 (Fig. 88). In profilin II His133 and Tyr139 are tyrosine and valine residues respectively. Again these residues do not appear to be obvious candidates for ring opening of the furanone of **44**, histidine is the possible exception.

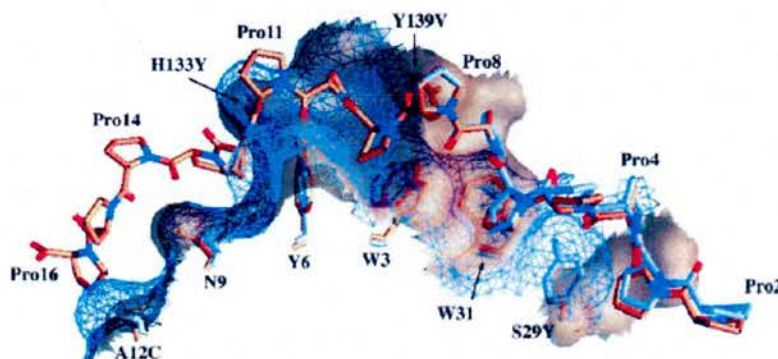


Figure 95. Binding site for proline-rich ligands. Human profilin I (brown) and human profilin II (blue) with poly-proline bound. The residues in profilin I which interact with proline are Trp3, Tyr 6, Trp31, His 133 and Tyr 139. In profilin II His133 and Tyr139 are tyrosine and valine residues respectively.

His133 and Tyr133 hydrogen bond to the carbonyl of Pro11 and stack against Pro13. Figure taken from Nodelman *et al.*, *J. Mol. Biol.*, **1999**, 294, 1271-1285.²¹⁷

5.5.1.3.3 Other conserved sites

Profilin also has a structurally conserved loop which follows the N-terminal helix. The crystal structure of profilin II shows the presence of three cysteines: Cys12, Cys15 and Cys16 (Fig. 96). Although in profilin I Cys15 is Thr15 the electrostatic interaction with Asp13 remains. This loop is not part of any of the obvious ligand binding sites of profilin. However should compound **44** bind to this site through one of the cysteines then it may exert a conformational change on profilin, thus effecting its binding to ligands at other sites (allosteric binding).

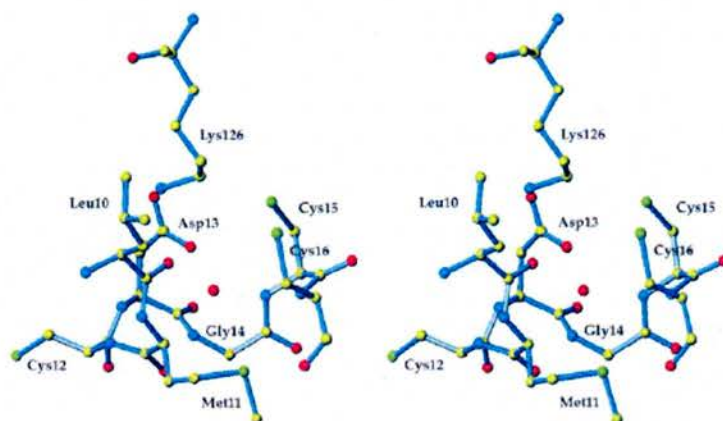


Figure 96. Structurally conserves cysteine rich loop in profilin. Profilin II has three cysteines in this loop: Cys12, Cys15 and Cys16. This loop has an extensive and conserved hydrogen bonding network including coordination of water to Asp13 and Leu10. Figure taken from Nodelman *et al.*, *J. Mol. Biol.*, **1999**, 294, 1271-1285.²¹⁷

Within the amino acid sequence for *T. gondii* inflammatory profilin there are no cysteine residues at positions 15 and 16, however two cysteines do exist in positions 17 and 18 which probably represents the analogous cysteine enriched area.

5.5.1.3.4 Summary from examining the crystal structure

Although the two main active sites the actin binding and proline binding sites do not appear to have a key nucleophilic residues which **44** could covalently modify

there are other conserved sites which have cysteines present which could provide the nucleophile required for covalent modification.

These discussions however have been based on the crystal structure of human profilin. The profilin of *T. gondii* may differ in the residues of these active sites. Without the crystal structure of *T. gondii* profilin or any further information regarding the nature of the binding of **44** with profilin it is difficult to draw any conclusions at this early stage.

5.5.2 Summary

Profilin appears to be an attractive target for **44**. The role of profilin in actin polymerisation could explain the effect of **44** on motility. Profilin has also been implicated in functions such as membrane trafficking and signalling pathways. The mechanism of secretion in *T. gondii* (see section 1.5.2.4) would suggest that interfering with pathways involved in the secretion process would alter the observed microneme secretion. Motility (in fact actin specifically) and secretion are required for invasion therefore anything seen to interact with profilin would be expected to effect invasion.

5.6 Chapter summary

The mechanism of action of **44** and its derivatives has been investigated. Chemical investigations indicated that **44** could covalently modify a substrate by attack of a nucleophile to ring open the furanone. The biological activity of analogues of **44** did not correlate with the observed reactivity of the compounds towards a nitrogen nucleophile. An alternative, as yet untested, possibility is that **44** covalently modifies its protein target(s) through conjugate addition of a sulfur based protein nucleophile.

Having concluded that **44** and its derivatives could potentially covalently modify their protein target(s) an alternative approach to affinity chromatography was selected. One *T. gondii* protein was shown to bind to the active resin and not the control: *inflammatory profilin*. Profilin covalently binds **44** and appears to be a viable

target based on its role in other organisms. Further experiments are required to confirm inflammatory profilin as the target of **44** but it appears a likely candidate.

Chapter 6: Experimental

6.1 General.

Chemicals and solvents were purchased from Acros Organics, Alfa Aaer-Avocado/Lancaster, Bamford Laboratories, Fisher Scientific or Sigma-Aldrich and were used as received unless otherwise stated. Tetrahydrofuran (THF) was dried by refluxing with sodium-benzophenone under an atmosphere of nitrogen and collected by distillation. Dichloromethane (DCM) was dried by heating under reflux over calcium hydride and distilled under an atmosphere of nitrogen.

Analytical thin-layer chromatography (TLC) was performed on pre-coated TLC plates SIL G-25 UV₂₅₄ (layer 0.25 mm silica gel with fluorescent indicator UV₂₅₄) (Aldrich). Developed plates were air dried and analysed under a UV lamp, Model UVGL-58 (Mineralight LAMP, Multiband UV_{254/365} nm). Flash column chromatography was performed using silica gel (40-63 μm , Fluorochem).

Microanalysis for carbon, hydrogen and nitrogen was performed by Mrs S. Williamson using EA 1110 CHNS CE Instruments, elemental analyser at the University of St. Andrews.

Melting points were recorded using an Electrothermal 9100 capillary melting point apparatus. Values are quoted to the nearest 0.5 °C and are uncorrected.

Fourier Transform Infra-red (FT IR) spectra were recorded on a Perkin Elmer Paragon 1000 FT spectrometer. Absorption maxima are reported in wavenumbers (cm^{-1}).

¹H Nuclear magnetic resonance (NMR) spectra were recorded either on a Brüker Avance 300 (300.1 MHz), Brüker Avance 500 (499.9 MHz), or a Varian Gemini 2000 (300.0 MHz) using the deuterated solvent as the lock and the residual solvent as the internal reference in all cases. ¹³C NMR spectra using the PENDANT sequence were recorded on a Bruker Avance 300 (75.5 MHz) or a Bruker Avance 500 (125.7 MHz) spectrometer. All other ¹³C spectra were recorded on a Varian Gemini 2000 (75.5 MHz) spectrometer using composite pulse ¹H decoupling. All NMR data was acquired using the deuterated solvent as the lock and the residual solvent as the

internal reference. The chemical shift data for each signal are given as (δ) in units of parts per million (ppm). Coupling constants (J) are quoted in Hz and are recorded to the nearest 0.5 Hz. The multiplicity used in the assignment of ^1H NMR spectra is indicated by the following abbreviations: s, singlet; d, doublet; dd, doublet of doublets; ddd, doublet of doublets of doublets; dt, doublet of triplets; t, triplet; q, quartet, m, multiplet and br, broad. Signals were assigned by means of two-dimensional (2D) NMR spectroscopy (^1H - ^1H COSY (Correlated Spectroscopy), ^1H - ^{13}C COSY (HSQC: Heteronuclear Single Quantum Coherence) and long-range ^1H - ^{13}C COSY spectra (HMBC: Heteronuclear Multiple Bond Connectivity). ^{19}F spectra were recorded on a Brüker Avance 300 (282.3 MHz).

Low and high-resolution mass spectral analysis (EI and CI) were recorded using a VG AUTOSPEC mass spectrometer or a Micromass GCT (Time-of-Flight), high performance, orthogonal acceleration spectrometer coupled to an Agilent Technologies 6890N GC system. Electrospray mass spectrometry (ES) was recorded on a high performance orthogonal acceleration reflecting TOF mass spectrometer operating in positive or negative ion mode, coupled to a Waters 2975 HPLC. Relative abundances and assignments are given in parantheses. The separation was carried out using a Waters XTerra, RP₁₈ 5 μM , 30 x 50 mm column with a flow rate of 0.5 ml/min and a 20 min gradient of 05-95% CH_3OH in H_2O , using a Waters 996 photodiode array detector. LCMS/MS of proteins samples was achieved using a Applied Biosystems Q-STAR Pulsar XL Quadrupole TOF instrument. Monoisotopic peptide masses were selected using Biolynx ProteinProbe (Micromass, Manchester, UK) and submitted for peptide mass matching against the NCNI databse using the Mascot search engine (Matrix Science).

X-Ray Crystallography data were recorded by Prof. A. Slawin on: a) Brüker SMART diffractometer with graphite-monochromated Mo-K α radiation ($\lambda = 0.71073$ Å), sealed tube and CCD detector, b) Mer-Rigaku, mercury detector 007 with Mo-K α radiation ($\lambda = 0.71073$ Å) generator (rotating anode), or c) Cop-Saturn 92 detector 007, Cu-K α radiation with rotation.

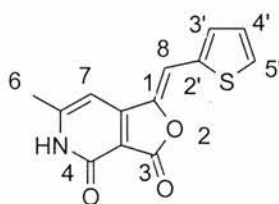
UV Absorption spectra were recorded on a Cary Varian model 300 absorption spectrophotometer, and time-integrated PL spectra were measured on a Jobin Yvon Fluoromax 2 fluorimeter. Measurements were recorded at a sample concentration of

3-10 mg L⁻¹, in HPLC grade DMSO. The optical densities of the samples were similar and small (~0.2).

6.2 Synthetic procedures and Analysis.

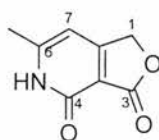
The compounds described in this section are arranged in order of the number assigned to the compound.

6.2.1 Synthesis of 6-methyl-1-[2-thienylmethylidene]furo[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**44**)¹⁵⁴



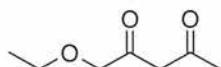
To lactone (**48**) (0.4 g, 2.4 mmol) was added methanol (40 cm³) and the resulting solution was heated to 60°C. The lactone does not dissolve completely in solution. Thiophene aldehyde (0.226 cm³, 2.4 mmol) and piperidine (0.22 cm³, 2.4 mmol) were added. The reaction mixture was heated at 90°C overnight. A yellow solid suspended in the reaction mixture was filtered and washed with methanol (3 x 10 cm³), to give the desired compound **44** as a yellow solid (0.454 g, 72%), mp >300 °C (decomposes), (lit.¹⁵⁴, 300 °C (decomposes)).

IR (KBr) 2388, 1790 (C(3)=O), 1673 (C(4)=O), 1615 (C=C), 1463 1177, 989, 697 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.17 (1H, brs, NH), 7.88 (1H, dd, *J*= 1.0, 5.0 Hz, H5'), 7.50 (1H, dd, *J*= 1.0, 4.0 Hz, H3'), 7.37 (1H, s, H8), 7.21 (1H, dd, *J*= 4.0, 5.0 Hz, H4), 6.72 (1H, s, H7), 2.34 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.6 (C4), 157.7 (C3), 157.7 (C6), 155.3 (C3a), 140.5 (C1), 135.6 (C2'), 132.0 (C3'), 132.19 (C5'), 128.4 (C4'), 106.0 (C8), 105.5 (C7a), 96.6 (C7), 20.0 (CH₃); *m/z* (EI+) 259 ([M⁺], 100%), 202 (6), 174 (6), 155 (10), 148 (18), 135 (6), 91 (18), 66(9); exact mass calcd for C₁₃H₉NO₃SNa 282.0201 (M + Na⁺) found 282.0197; Anal. Calcd for C₁₃H₉NO₃S: C, 60.22; H, 3.50; N, 5.40, found C, 59.69; H, 3.24; N, 5.19 (differing by 0.5%); λ_{max}(DMSO)/nm 378 (29474 ε/dm³ mol⁻¹ cm⁻¹); λ_{max} fluor (DMSO) 449 nm.

6.2.2 Synthesis of 6-methylfuro[3,4-c]pyridine-3,4(*1H,5H*)-dione (**48**)^{152,153,158}

Pyridone (**50**) (1 g, 5.2 mmol) was dissolved in the minimum amount of 10% sulfuric acid (30 cm³). This was heated at 125°C for 4 hours and the reaction mixture poured hot into water/ice mixture (50 cm³). The white needle-like crystals that formed were filtered and washed with water (2 x 20 cm³), ethanol (2 x 10 cm³) then ether (2 x 10 cm³) to give the desired compound **48** as a white solid (0.56 g, 85%), mp >300 °C (decomposes), (lit.¹⁵³ mp 300 °C (decomposes)).

IR (KBr) 2368, 2344, 1761 (C=O lactone), 1670 (C=O pyrone), 1617 (C=C), 1560, 1377, 1181, 1143, 1020 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.15 (1H, br s, NH), 6.34 (1H, s, H7), 5.16 (2H, s, CH₂), 2.31 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 168.2 (C4) 167.3 (C3), 158.2 (C3a), 154.9 (C6), 109.1 (C7a), 99.7 (C7), 68.1 (C1), 19.5 (CH₃); *m/z* (EI+) 165 ([M⁺], 100%), 155 (8), 136(39), 108 (14), 91 (11), 80 (5), 65 (10), 53 (8).

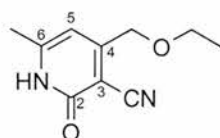
6.2.3 Synthesis of 1-ethoxypentane-2,4-dione (**51**)^{158,159}

Sodium hydride (60%w/w, 8 g, 208 mmol) was washed with dry toluene (30 cm³). The washed sodium hydride was then resuspended in dry toluene (90 cm³) and cooled to -10 °C, followed by the addition of ethyl ethoxyacetate (**49**) (25 cm³, 208 mmol) added. Acetone (14 cm³, 208 mmol) was added dropwise. The reaction mixture was stirred at 4-6 °C for two hours. The reaction mixture was stirred overnight and allowed to warm to room temperature. A white solid of the sodium salt of (**51**) precipitated. Diethyl ether was added (250 cm³) and washed with sulfuric acid (10%, 3 x 100 cm³), where upon the precipitate went into the aqueous layer. The organic layer was dried with magnesium sulfate, filtered and concentrated *in vacuo*. The

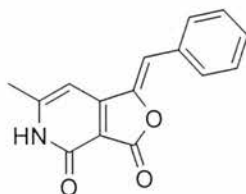
crude mixture was purified by flash column chromatography eluting with ethyl acetate and hexane to give the desired compound **51** as a yellow oil (10.2 g, 54%).

Rf (0.7) EtOAc/Hexane: 2/3; ^1H NMR (CDCl_3 , 300.1 MHz) 5.70 (1H, s, CH), 3.92 (2H, s, OCH_2CO), 3.45 (2H, q, $J=7.0$ Hz, CH_2CH_3), 1.98 (3H, s, COCH_3), 1.14 (3H, t, $J=7.0$ Hz, CH_2CH_3); ^{13}C NMR (CDCl_3 , 75.5 MHz) 187.5 (C=O), 185.8 (C=O), 92.1 (CH), 66.7 (OCH_2CO), 62.1 (CH_2), 19.6 (COCH_3), 9.9 (CH_2CH_3); m/z (ES+) 167 ($[\text{M}+\text{Na}^+]$, 100%), 145 ($[\text{M}+\text{H}]$, 40).

6.2.4 Synthesis of 3-cyano-4-ethoxymethyl-6-methyl-2-pyridone (**50**)^{152,153,158}



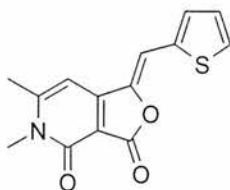
1-Ethoxypentane-2,4-dione (**51**) (2.3 g, 16 mmol) was added to cyanoacetamide (**54**) (1.34 g, 16 mmol) and dissolved in the minimum amount of ethanol (50 cm^3) and heated to 60 °C. Addition of piperidine (1.58 cm^3 , 16 mmol, 1eq) occurred and the reaction mixture was allowed to cool to room temperature overnight. A white solid precipitated from the solution. The white crystals were filtered and washed with ethanol (3 x 50 cm^3). The remaining reaction mixture was stirred over night. Further white solid precipitated from the solution, which was filtered and the process was repeated. Purification was achieved by recrystallisation from hot ethanol. To give the desired compound **50** as a white solid (2 g, 65%), mp 210-212 °C, (lit.¹⁵³ mp 210 °C). IR (NaCl, nujol) 2363, 2345, 2217 ($\text{C}\equiv\text{N}$), 1654 (C=O), 1623 (C=C), 1465, 1378, 1329, 1128, 1063 cm^{-1} . ^1H NMR (CD_3OD , 300.1 MHz) 6.51 (1H, s, H7), 4.60 (2H, s, CH_2O), 3.66 (2H, q, $J=7.0$ Hz, CH_2CH_3), 2.41 (3H, s, CH_3), 1.30 (2H, t, $J=7.0$ Hz, CH_2CH_3); ^{13}C NMR ($\text{DMSO}-d_6$, 75.5 MHz) 161.3 (C2), 161.0 (C3), 153.0 (C6), 115.4 (CN), 104.2 (C5), 96.7 (C4), 67.0 (OCH_2), 66.4 (CH_2CH_3), 19.5 (CH_3), 15.2 (CH_2CH_3); m/z (EI+) 192 ($[\text{M}^+]$, 27%), 148 ($[\text{M}-\text{OCH}_2\text{CH}_3]$, 100), 135 (5), 120 (20), 105 (6), 65(7).

6.2.5 Synthesis of 1-benzylidene-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (59)


To lactone (**48**) (1 g, 6 mmol), methanol (40 cm³) was added and heated to 60 °C. The lactone does not dissolve completely in solution. Benzylaldehyde (0.62 cm³, 6 mmol) and piperdine (0.6 cm³, 6 mmol) were added. The reaction mixture was heated to 90 °C overnight. A yellow solid suspended in the reaction mixture was filtered and washed with methanol (3 x 10 cm³), to give the desired compound **59** as a yellow solid (0.889 g, 58%), mp >293 °C (decomposes).

IR (KBr) 2929, 1774 (C(3)=O), 1670 (C(4)=O), 1609 (C=C), 1565, 1139, 942, 802, 688, 630 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.25 (1H, brs, NH), 7.83-7.80 (2H, m, AA' part of the AA'BB'C system, H2' & 6'), 7.52-7.47 (2H, m, BB' part of the AA'BB'C system, H3' & 5'), 7.43-7.38 (1H, m, C part of the AA'BB'C system, H4'), 6.96 (1H, s, H8), 6.79 (1H, s, H7), 2.36 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.9 (C4), 157.6 (C3), 156.2 (C6), 155.4 (C3a), 142.7 (C1), 132.7 (C1'), 129.2 (C2' & C6'), 128.9 (C4'), 127.4 (C3' & C5'), 111.2 (C8), 105.5 (C7a) 96.7 (C7), 19.5 (CH₃); *m/z* (EI+) 276 ([M+Na⁺], 100%), 254 ([M+H⁺], 6), 236 ([M-H₂O+H⁺], 10); exact mass calcd for C₁₅H₁₂NO₃ 254.0817 (M + H⁺) found 254.0825. Anal. Calcd for C₁₅H₁₂NO₃: C, 71.14; H, 4.38; N, 5.53, found C, 70.84; H, 4.13; N, 5.28; λ_{max}(DMSO)/nm 354 (31767 ε/dm³ mol⁻¹ cm⁻¹); λ_{max} fluor (DMSO) 443 nm.

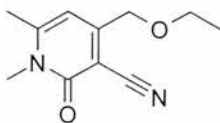
6.2.6 Synthesis of 5,6-dimethyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**60**)



6-Methyl-1-[2-thienylmethylidene]furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**44**) (0.2 g, 0.77 mmol) was dissolved in dimethylformamide (50 cm³). Potassium carbonate (0.128 g, 0.92 mmol) was added to the reaction solution, followed by iodomethane (0.057 cm³, 0.92 mmol). The reaction mixture was stirred overnight at room temperature, then concentrated *in vacuo*. The crude compound was purified by flash column chromatography, eluting with dichloromethane and methanol, to give the desired compound **60** as a yellow solid (0.125 g, 59%), mp >240 °C decomposes.

IR (KBr) 3079, 1778 (C(3)=O), 1651 (C(4)=O), 1578, 1423, 1375, 1278, 1209, 1038, 964, 927, 816, 794, 703, 591 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.87-7.84 (1H, m, H5'), 7.49-7.47 (1H, m, H3'), 7.31 (1H, s, H8), 7.20 (1H, dd, *J*= 4.0 5.0 Hz, H4'), 6.87 (1H, s, H7), 3.48 (3H, s, NCH₃), 2.51 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.8 (C4), 157.2 (C6), 157.2 (C3a), 152.5 (C3), 140.2 (C1), 135.6 (C1'), 132.0 (C5'), 131.9 (C3'), 128.1 (C4'), 105.7 (C8), 104.5 (C7a), 97.8 (C7), 30.7 (NCH₃), 22.1 (CH₃); *m/z* (ES+) 569 ([2M+Na⁺], 27%), 296 ([M+Na⁺], 100), 274 ([M+H⁺], 9); exact mass calcd for C₁₄H₁₁NO₃SNa 296.0357 (M + Na⁺) found 296.0356; Anal. Calcd for C₁₄H₁₁NO₃S: C, 61.53; H, 4.06; N, 5.12, found C, 61.72; H, 3.94; N, 4.95.

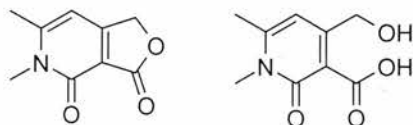
6.2.7 Synthesis of 4-(ethoxymethyl)-1,6-dimethyl-2-oxo-1,2-dihydropyridine-3-carbonitrile (**64**).¹⁶⁷



The 3-cyano-4-ethoxymethyl-6-methyl-2-pyridone (**50**) (0.2 g, 1 mmol) was dissolved in ethanol (10 cm³). Potassium hydroxide (0.086 g, 1.5 mmol) in water (1 cm³) and methyl iodide (1.42 cm³, 10 mmol) was added. The reaction mixture was heated under reflux for 2 hours and concentrated *in vacuo*. The product was purified using chromatography to eluting with ethyl acetate, hexane and methanol to give the desired compound **64** as a white solid (0.2g, 93%), mp 117-118 °C (lit.¹⁶⁷ mp 110-111 °C).

IR (nujol) 2359, 2345, 2287 (C≡N), 1653 (C=O), 1583, 1541 (C=C), 1458, 1424, 1377, 1127 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 6.59 (1H, s, H7), 4.57 (2H, s, CCH₂O), 3.67 (2H, q, *J*= 7.0 Hz, CH₂CH₃), 3.63 (3H, s, CH₃N), 2.58 (3H, s, CCH₃), 1.31 (3H, t, *J*= 7.0 Hz, CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 160.9 (C4), 158.2 (C3), 155.0 (C6), 115.6 (CN), 105.4 (C5), 96.2 (C4), 69.0 (CH₂O), 66.4 (CH₂CH₃), 31.7 (NCH₃), 21.5 (CH₃), 15.2 (CH₂CH₃); *m/z* (EI+) 229 (M+Na⁺, 100%) 207 (M+H⁺, 5).

6.2.8 Synthesis of 5,6-dimethylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**63**) and 4-(hydroxymethyl)-6-methyl-2-oxo-1,2-dihydropyridine-3-carboxylic acid (**65**).



4-(Ethoxymethyl)-1,6-dimethyl-2-oxo-1,2-dihydropyridine-3-carbonitrile (**64**) (1.15 g, 6 mmol) dissolved in the minimum amount of 50% sulfuric acid (150 cm³) was heated at 125 °C for 4 hours. The reaction mixture was then poured onto ice. The resulting aqueous solution was basified and extracted with dichloromethane (3 x 70 cm³). The organic layers were combined, dried with magnesium sulfate and concentrated *in vacuo*, to give the lactone compound **63** as a white solid (0.05 g, 5%) mp 177 °C decomposes. The remaining aqueous solution was acidified and extracted

with dichloromethane (3 x 70 cm³). The resulting organic layers were combined, dried with magnesium sulfate and concentrated *in vacuo*, to give the acid alcohol pyridone (**65**) as a white solid (0.38 g, 34%) mp 252-253 °C.

5,6-Dimethylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**63**)

IR (KBr) 3295, 2563, 2345, 1707 (C=O lactone), 1540 (C(4)=O), 1465, 1376, 1257, 1153, 1082, 1015 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.01 (1H, s, H7), 4.87 (2H, s, CH₂), 3.35 (3H, s, NCH₃), 2.55 (3H, s, CCH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 166.0 (C4), 165.2 (C3), 163.6 (C3a), 153.2 (C6), 108.5 (C7a), 108.0 (C7), 61.5 (C1), 32.5 (NCH₃) 21.4 (CH₃); *m/z* (EI+) 220 (M+MeCN, 100%), 179 (M+H, 60).

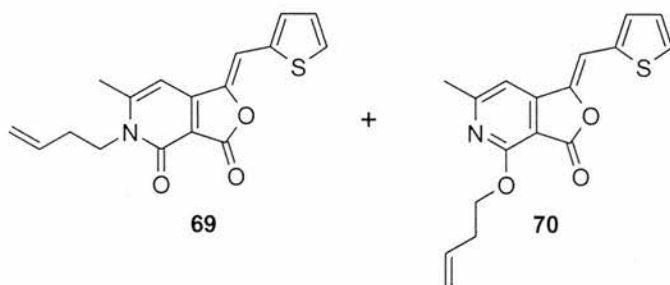
4-(Hydroxymethyl)-1,6-dimethyl-2-oxo-1,2-dihydropyridine-3-carboxylic acid (**65**)

IR (KBr) 3452, 1711 (C=O), 1541 (C(4)=O), 1481, 1404, 1257, 1154, 1084, 1015, 805, 682, 588, 479 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.79 (1H, s, H7), 6.46 (1H, s, OH), 5.16 (2H, s, CH₂OH), (3H, s, NCH₃), 2.46 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 165.6 (C2), 164.9 (C3), 163.3 (CO₂), 152.9 (C6), 108.0 (C4), 107.5 (C5), 61.1 (CH₂), 32.1 (NCH₃), 21.1 (CH₃); *m/z* (EI+) 220 ([M+Na⁺], 100%); exact mass calcd for C₉H₁₂NO₄ 198.0766 (M + H⁺) found 198.0769; Anal. Calcd for C₉H₁₁NO₄: C, 54.82; H, 5.62; N, 7.10, found C, 54.92; H, 5.49; N, 6.99.

6.2.9 Synthesis of 5,6-dimethylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**63**) through conversion of **65**.

4-(Hydroxymethyl)-1,6-dimethyl-2-oxo-1,2-dihydropyridine-3-carboxylic acid (**65**) (0.2 g, 1 mmol) and tosic acid monohydrate (0.193 g, 1 mmol) were heated under reflux for 10 hours in tetrahydrofuran (50 cm³) in the presence of powdered molecular sieves. The molecular sieves were removed and the reaction solution concentrated *in vacuo*. Dichloromethane (40 cm³) and water (40 cm³) were added. The aqueous layer was extracted with dichloromethane (3 x 30 cm³), dried with magnesium sulfate and concentrated *in vacuo* to yield the desired compound **65** as a white solid (0.12g, 66%), analysis as above.

6.2.10 Synthesis of 5-but-3-en-1-yl-6-methyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**69**) and 4-(but-3-en-1-yloxy)-6-methyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridin-3(1*H*)-one (**70**).



6-Methyl-1-[2-thienylmethylidene]furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**44**) (0.1 g, 0.39 mmol) was dissolved in dimethylformamide (50 cm³). Potassium carbonate (0.05 g, 0.39 mmol) was added followed by 4-bromo-1-butene (**72**) (0.04 cm³, 0.39 mmol). The reaction mixture was stirred overnight at room temperature. Potassium iodide was added (0.064 g, 0.39 mmol) with 4-bromo-1-butene (0.03 cm³, 0.3 mmol) and the reaction mixture was stirred for a further 12 hours, then concentrated *in vacuo*. The crude compound was purified by flash column chromatography eluting with dichloromethane and methanol to yield two compounds: 5-But-3-en-1-yl-6-methyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**69**) as a yellow solid (0.018 g, 15%), mp 204-205 °C and 4-(but-3-en-1-yloxy)-6-methyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridin-3(1*H*)-one (**70**) as a yellow solid (0.006 g, 4%), mp 145-146 °C.

5-But-3-en-1-yl-6-methyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**69**)

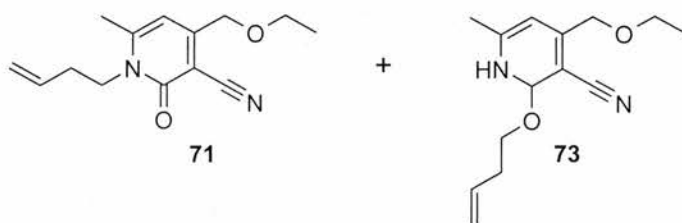
IR (KBr) 3068, 2925, 1773 (C(3)=O), 1672 (C(4)=O), 1572 (C=C), 1417, 1053, 941, 706 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.88-7.86 (1H, m, H3'), 7.49-7.48 (1H, m, H5'), 7.34 (1H, s, H8), 7.20 (1H, dd *J*= 4.0 5.0 Hz, H4'), 6.88 (1H, s, H7), 5.93-5.79 (1H, m, CH₂CH), 5.14-5.79 (2H, m, CH₂CH(CH₂)₂), 4.09-4.04 (2H, m, CH₂N), 2.36 (3H, s, CH₃), 2.42-2.34 (2H, m, CHCH₂CH₂N); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.7 (C4), 156.9 (C3), 156.3 (C6), 152.6 (C3a), 140.2 (C1), 135.5 (C2'), 134.6 (CH₂CH(CH₂)₂), 132.1 (C4' & C5'), 128.1 (C3'), 117.4 (CH₂CH(CH₂)₂N), 105.9 (C8), 104.9 (C7a), 98.3 (C7), 43.2 (NCH₂), 31.8 (NCH₂CH₂), 21.7 (CH₃); *m/z*

(ES⁺) 649 ([2M+Na⁺], 8%), 336 ([M+Na⁺], 100); exact mass calcd for C₁₇H₁₅NO₃NaS 336.0670 (M + Na⁺) found 336.0672.

4-(But-3-en-1-yloxy)-6-methyl-1-(2-thienylmethylene)furo[3,4-*c*]pyridin-3(1*H*)-one (70)

IR (KBr) 2962, 1794 (C(3)=O), 1605 (C=C), 1582, 1420, 1363, 1339, 1273, 1044, 978, 929, 741, 728 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.83-7.81 (1H, m, H3'), 7.47 (1H, s, H8), 7.46-7.45 (1H, m, H5'), 7.38 (1H, s, H7), 7.19 (1H, dd *J*= 4.0 5.0 Hz, H4'), 6.00-5.81 (1H, m, CH₂CH), 5.21-5.10 (2H, m, CH₂CH(CH₂)₂), 6.91 (2H, t *J*= 7.0 Hz, CH₂O), 2.52 (3H, s, CH₃), 2.57-2.50 (2H, m, CHCH₂CH₂O); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.4 (C4), 163.4 (C3), 162.8 (C6), 159.7 (C3a), 151.5 (C1'), 140.6 (C1), 135.4 (C2'), 134.4 (CH₂CH), 131.6 (C5'), 131.4 (C4'), 127.9 (C3'), 117.3 (CH₂CH), 107.2 (C8), 104.4 (C7), 101.9 (C7a), 65.4 (OCH₂), 32.6 (OCH₂CH₂), 24.8 (CH₃); *m/z* (ES⁺) 649 ([2M+Na⁺], 28%), 336 ([M+Na⁺], 100); exact mass calcd for C₁₇H₁₅NO₃NaS 336.0670 (M + Na⁺) found 336.0673.

6.2.11 Synthesis of 1-but-3-en-1-yl-4-(ethoxymethyl)-6-methyl-2-oxo-1,2-dihydropyridine-3-carbonitrile (71) and 2-(but-3-en-1-yloxy)-4-(ethoxymethyl)-6-methylnicotinonitrile (73).



Pyridone (50) (0.2 g, 1.04 mmol) was partially dissolved in acetonitrile (20 cm³). An aqueous solution of potassium hydroxide (0.68 cm³, 2 M, 1.35 mmol) and 4-bromo-1-butene (2 cm³, 19.7 mmol) were added to the suspension. The reaction mixture was heated at 90 °C for 24 hrs. Once the reaction mixture was cooled the unreacted pyridone (50) precipitated and was removed through filtration. The filtrate was then concentrated *in vacuo*. The crude mixture was purified by flash column chromatography eluting with hexane and ethyl acetate to two compounds the *N*-

butenylated pyridone (**71**) as a white solid (0.097 g, 38%) mp 100-101 °C and the *O*-butenylated pyridone (**73**) as a white solid (0.084 g, 33%), mp 55-56 °C.

1-But-3-en-1-yl-4-(ethoxymethyl)-6-methyl-2-oxo-1,2-dihydropyridine-3-carbonitrile (**71**)

IR (KBr) 2976, 2218, 1644 (C=O), 1582, 1550, 1420, 1200, 1101, 925, 862, 783, 645, 620, 455 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 6.42 (1H, s, H7), 5.90-5.76 (1H, m, CH₂CH), 5.11-5.02 (2H, m, CH₂CH(CH₂)₂), 4.47 (2H, s, CCH₂O), 4.04 (2H, t *J* 7.5, NCH₂), 3.57-3.50 (2H, m, CH₂CH₃), 2.51 (3H, s, CH₃), 2.40-2.33 (2H, m, CHCH₂CH₂N), 1.18 (3H, s, CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 160.2 (C2), 158.2 (C3), 153.8 (C6), 134.5 (CH₂CH), 117.4 (CH₂CH), 115.1 (C≡N), 105.6 (C5), 96.4 (C4) 68.9 (OCH₂), 67.1 (CH₂CH₃), 44.8 (NCH₂CH₂), 32.2 (CH₂CH₂O), 21.4 (CH₃), 15.0 (CH₂CH₃); *m/z* (ES+) 269 (M+Na⁺, 100%); exact mass calcd for C₁₄H₁₈N₂O₂Na 269.1266 (M+ Na⁺) found 269.1263.

2-(But-3-en-1-yloxy)-4-(ethoxymethyl)-6-methylnicotinonitrile (**73**)

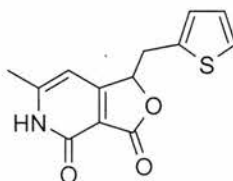
IR (KBr) 2974, 2225, 1596, 1566, 144, 1416, 1354, 1336, 1153, 1076, 738 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.05 (1H, s, H7), 5.94-5.80 (1H, m, CH₂CH(CH₂)₂), 5.19-5.05 (2H, m, CH₂CH(CH₂)₂), 4.55 (2H, s, CCH₂O), 4.41 (2H, t *J*= 7.0 Hz, OCH₂CH₂), 3.59-3.52 (2H, m, CH₂CH₃), 2.53-2.46 (2H, m, CHCH₂CH₂O), 2.46 (3H, s, CH₃), 1.18 (3H, s, CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.1 (C2), 161.5 (C6), 155.1 (C3), 134.4 (CH₂CH), 117.2 (CH₂CH), 114.7 (C5), 113.8 (C≡N), 69.1 (CH₂OCH₂CH₃), 67.0 (OCH₂CH₃), 66.3 (CH₂CH₂N), 33.2 (CH₂CH₂N), 24.9 (CH₃), 15.1 (CH₂CH₃); *m/z* (ES+) 515 ([2M+Na⁺], 12%), 269 ([M+Na], 100); exact mass calcd for C₁₄H₁₈N₂O₂Na 269.1266 (M + Na⁺) found 269.1262.

6.2.12 Selective synthesis of 2-(but-3-en-1-yloxy)-4-(ethoxymethyl)-6-methylnicotinonitrile (**73**)

Pyridone (**50**) (1 g, 5.2 mmol) was partially dissolved in tetrahydrofuran (120 cm³). 3-butene-1-ol (0.295 cm³, 3.5 mmol), tri-*tert*-butylphosphine (1.29 cm³, 5.2 mmol)

and 1,1'-azodicarbonyl dipiperidine (ADDP) (1.31 g, 5.2 mmol) were added. The reaction mixture was heated at 60 °C over night, allowing the pyridone (**50**) to be solubilised in the reaction mixture. Hexane (80 cm³) was added to the reaction mixture causing the dihydro by-product of ADDP (**74**), 1,1'-hydrazine carbonyl dipiperidine, to precipitate. This by-product was removed through filtration and the remaining reaction mixture concentrated *in vacuo*. The crude compound was purified by flash column chromatography eluting with hexane and ethyl acetate to give the desired compound **73** as a white solid (0.327 g, 38%). Analysis as above.

6.2.13 Synthesis of 6-methyl-1-(2-thienylmethyl)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**75**).

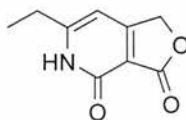


6-Methyl-1-(2-thienylmethylidene)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**44**) (0.15 g, 0.6 mmol) was taken with palladium on carbon (10% w/w, 0.08 g) under nitrogen. Dimethylformamide (40 cm³) and methanol (10 cm³) was added, the flask flushed with hydrogen and the reaction mixture was stirred overnight. The palladium was removed through filtering over cellite and the reaction mixture concentrated *in vacuo*. The crude compound was purified by flash column chromatography eluting with dichloromethane and methanol to give the desired compound **75** as a white solid (0.8 g, 53%), mp 191-192 °C (decomposes).

IR (KBr) 2927, 1752 (C(3)=O), 1668 (C(4)=O), 1613 (C=C), 1569, 1481, 1146, 1050, 694 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.13 (1H, brs, NH), 7.33 (1H, dd, *J*= 1.0, 5.0 Hz, H5'), 6.92 (1H, dd, *J*= 4.0, 5.0 Hz, H4'), 6.84 (1H, dd, *J*= 1.0, 4.0 Hz, H3'), 6.41 (1H, s, H7), 5.60 (1H, dd, *J*= 6.0 Hz, H1), 3.44 (2H, ddd, *J*= 4.0, 15.0, 50.0 Hz, H8), 2.29 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 167.7 (C4), 167.2 (C3), 157.7 (C3a), 154.7 (C6), 136.3 (C2'), 127.2 (C3'), 126.7 (C4'), 125.2 (C5'), 109.6 (C7a), 99.3 (C7), 77.6 (C1), 32.2 (CH₂), 19.3 (CH₃); *m/z* (ES⁺) 545 ([2M+Na⁺], 15%), 284 ([M+Na⁺], 100); exact mass calcd for C₁₃H₁₂NO₃ 262.0538 (M + H⁺)

found 262.0539; Anal. Calcd for $C_{13}H_{11}NO_3$: C, 59.76; H, 4.24; N, 5.36, found C, 59.13; H, 4.15; N, 5.18.

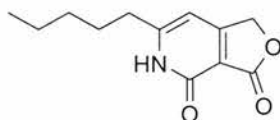
6.2.14 Synthesis of 6-ethylfuro[3,4-c]pyridine-3,4-(1H,5H)-dione (**78**).¹⁷⁵



Pyrone (**86**) (0.4 g, 2 mmol) dissolved in the minimum amount of 50% sulfuric acid (70 cm^3) was heated at $125\text{ }^\circ\text{C}$ for 4 hours. This reaction mixture was poured on to ice/water and the aqueous was extracted with dichloromethane ($3 \times 70\text{ cm}^3$). The organic layer was dried with magnesium sulfate and concentrated *in vacuo* to give the desired product (**78**) as a white solid (0.07 g, 20%), mp $283\text{--}284\text{ }^\circ\text{C}$ (decomposes). (lit.¹⁷⁵ mp $285\text{ }^\circ\text{C}$).

IR (KBr) $2933, 2232, 1772$ (C(3)=O), 1655 (C(4)=O), 1570 (C=C), $1438, 1151, 1017, 1002, 803, 656\text{ cm}^{-1}$; $^1\text{H NMR}$ (CD_3OD , 300.1 MHz) 6.27 (1H, s, Ar-H), 5.11 (2H, s, CH_2O), 2.78 (2H, q, $J = 7.5\text{ Hz}$, CH_2CH_3), 1.35 (3H, t, $J = 7.5\text{ Hz}$, CH_2CH_3); $^{13}\text{C NMR}$ ($\text{DMSO-}d_6$, 75.5 MHz) 168.2 (C4), 167.1 (C3), 159.7 (C6), 157.9 (C3a), 109.1 (C7a), 97.8 (C7), 67.9 (C1), 26.0 (CH_2CH_3), 12.6 (CH_3); m/z (ES+) 203 ($[\text{M}+\text{Na}^+]$, 100%).

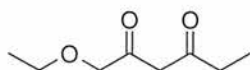
6.2.15 Synthesis of 6-pentylfuro[3,4-c]pyridine-3,4-(1H,5H)-dione (**79**).



4-Ethoxymethyl-2-oxo-6-pentyl-1,2-dihydro-pyridine-3-carbonitrile (**87**) (0.4 g, 1.6 mmol) dissolved in the minimum amount of 50% sulfuric acid (70 cm^3) was heated at $125\text{ }^\circ\text{C}$ for 4 hours. This was added to ice/water and the aqueous was extracted with dichloromethane ($3 \times 70\text{ cm}^3$). The organic layer was dried with magnesium sulfate and concentrated *in vacuo* to give the desired compound **79** as a white solid (0.08 g, 22%), mp $182\text{--}183\text{ }^\circ\text{C}$ (decomposes).

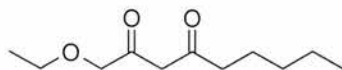
IR (KBr) 2928 (brs), 1751 (C(3)=O), 1655 (C(4)=O), 1569, 1487, 1440, 1352, 1238, 1183, 1143, 1104, 1020, 933, 798, 635 cm^{-1} ; ^1H NMR (CD_3OD , 300.1 MHz) 12.35 (1H, br s, NH), 6.24 (1H, s, H7), 5.10 (2H, s, H1), 2.75 (2H, t, $J=7.5$ Hz, COCH_2CH_2), 1.74–1.67 (2H, m, COCH_2CH_2), 1.4–1.34 (4H, m, $\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_3$), 0.92–0.88 (2H, m, CH_3); ^{13}C NMR ($\text{DMSO}-d_6$, 75.5 MHz) 168.2 (C4), 167.0 (C3), 158.5 (C3a), 158.0 (C6), 109.4 (C7a), 98.6 (C7), 67.9 (C1), 32.6 ($\text{CH}_2(\text{CH}_2)_4$), 30.5 ($\text{CH}_2\text{CH}_2(\text{CH}_2)_3$), 27.8 ($(\text{CH}_2)_2\text{CH}_2\text{CH}_2$), 21.7 (CH_2CH_3), 13.8 (CH_3); m/z (ES+) 465 ($[\text{2M}+\text{Na}]$, 82%), 244 ($[\text{M}+\text{Na}]$, 100); exact mass calcd for $\text{C}_{12}\text{H}_{15}\text{NO}_3\text{Na}$ 244.0933 ($\text{M} + \text{Na}^+$) found 244.0941.

6.2.16 Synthesis of 1-ethoxy-hexane-2,4-dione (**80**)¹⁷⁵



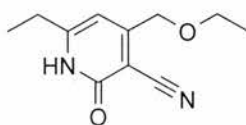
Sodium hydride (60%w/w, 1.8 g, 74 mmol) was washed with dry toluene (30 cm^3). The washed sodium hydride was then dissolved in dry toluene (80 cm^3). This was cooled to -10 °C and ethyl ethoxyacetate (**51**) (10 cm^3 , 74 mmol) added. Butan-2-one (**82**) (6.6 cm^3 , 74 mmol) was added dropwise. The reaction mixture was stirred overnight. Precipitation of a white solid of the sodium salt occurred. Diethyl ether added (250 cm^3) and washed with sulfuric acid (10%, 3x 100 cm^3), where upon the precipitate went into the aqueous layer. The organic layer was dried with magnesium sulfate and concentrated *in vacuo*. The crude compound was purified by flash column chromatography eluting with ethyl acetate and hexane to give the desired compound **80** as a yellow oil (3.97 g, 34%).

IR (NaCl) 2978, 1732 (C=O), 1711 (C=O), 1611, 1444, 1119, 810 cm^{-1} ; ^1H NMR (CDCl_3 , 300.1 MHz) 5.75 (1H, s, COCHCO), 4.00 (2H, s, OCH_2CO), 3.50 (2H, q, $J=7.0$ Hz, $\text{CH}_3\text{CH}_2\text{O}$), 2.30 (2H, q, $J=8.0$ Hz, COCH_2CH_3), 1.20 (3H, t, $J=7.0$ Hz, $\text{CH}_3\text{CH}_2\text{O}$), 1.09 (3H, t, $J=8.0$, COCH_2CH_3); ^{13}C NMR (CDCl_3 , 75.5 MHz) 195.3 (C2), 192.0 (C4), 95.9 (C3), 71.7 (C1), 67.0 (OCH_2CH_3), 31.4 (C5), 15.1 (OCH_2CH_3), 9.6 (C6); m/z (ES-) 157 ($[\text{M}+\text{H}]$, 100%).

6.2.17 Synthesis of 1-ethoxy-nonane-2,4-dione (**81**)

Sodium hydride (60%w/w, 2.8 g, 74 mmol) was washed with dry toluene (30 cm³). The washed sodium hydride was then dissolved in dry toluene (90 cm³). This was then cooled to -10 °C and ethyl ethoxyacetate (**51**) (10 cm³, 74 mmol) added. 2-Heptanone (**83**) (10 cm³, 74 mmol) was added dropwise. The reaction mixture was stirred overnight. Precipitation of a white solid of the sodium salt occurred. Diethyl ether added (250 cm³) and washed with sulfuric acid (10%, 3 x 100 cm³), where upon the precipitate went into the aqueous layer. The organic layer was dried with magnesium sulfate and concentrated *in vacuo*. The crude compound was purified by flash column chromatography eluting with ethyl acetate and hexane to yield the desired compound **81** as a yellow oil (8.57 g, 58%).

IR (NaCl) 2979, 2880, 1733 (C=O), 1711(C=O), 1615, 1445, 1327, 1119, 879, 810 cm⁻¹; ¹H NMR (CDCl₃, 300.1 MHz) 5.69 (1H, s, COCHCO), 3.89 (2H, s, OCH₂CO), 3.43 (2H, q, *J*= 7.0 Hz, CH₃CH₂O), 2.19 (2H, t, *J*= 7.0 Hz, COCH₂CH₂), 1.55–1.42 (2H, m, COCH₂CH₂), 1.24–1.15 (6H, m, CH₂CH₂CH₂CH₂CH₂CH₂CH₃), 1.12 (3H, t, *J*= 7.0 Hz, CH₃CH₃CO), 0.80–0.73 (3H, m, CH₂CH₃); ¹³C NMR (CDCl₃, 75.5 MHz) 194.3 (C2), 192.8 (C4), 96.63 (C3), 72.0 (C1), 67.3 (OCH₂CH₃), 38.3 (C5), 31.5 (C6) 25.6 (C7), 22.6 (C8), 15.3 (OCH₂CH₃), 14.1 (C9); *m/z* (ES+) 223 ([M+Na⁺], 100%); exact mass calcd for C₁₁H₂₀O₃Na 223.1310 (M + Na⁺) found 223.1311.

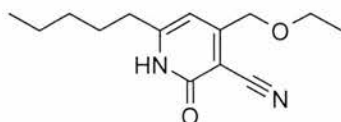
6.2.18 Synthesis of 4-ethoxymethyl-6-ethyl-2-oxo-1,2-dihydro-pyridine-3-carbonitrile (**86**).¹⁷⁵

1-Ethoxy-hexane-2,4-dione (**49**) (2.0 g, 1.3 mmol) was added to cyanoacetamide (**54**) (1.1 g, 7 mmol) and dissolved in the minimum amount of ethanol (150 cm³) and heated to 60°C. Addition of piperidine (1.3 cm³, 1.3 mmol, 1 eq) occurred and the

reaction mixture was allowed to cool to room temperature overnight. A white solid precipitated from the solution. The white crystals were filtered and washed with ethanol (3 x 50 cm³). To give the desired compound **86** as a white solid (0.75 g, 34%), mp 189-190 °C, (lit.¹⁷⁵ mp 190-191 °C).

IR (KBr) 2866, 2219 (C≡N), 1846, 1652 (C=O), 1542, 1484, 1450, 1380, 1343, 1311, 1198, 1140, 1120, 1085, 1057, 951, 851 cm⁻¹; ¹H NMR (CD₃OD, 300.1 MHz) 6.56 (1H, s, H5), 4.06 (2H, s, CCH₂O), 3.67 (2H, q, *J* = 8.0 Hz, OCH₂CH₃), 2.70 (2H, q, *J* = 8.0 Hz, CCH₂CH₃), 1.31 (6H, t, *J* = 8.0 Hz, CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 161.0 (C2), 160.9 (C3), 157.6 (C6), 115.0 (CN), 102.3 (C5), 96.8 (C4), 68.7 (CH₂O), 66.0 (OCH₂CH₃), 26.0 (CCH₂CH₃), 14.8 (OCH₂CH₃), 12.7 (CCH₂CH₃); *m/z* (ES⁺) 456 ([2M+2Na], 19%), 435 ([2M+Na], 12), 229 ([M+Na], 100).

6.2.19 Synthesis of 4-ethoxymethyl-2-oxo-6-pentyl-1,2-dihydro-pyridine-3-carbonitrile (**87**).

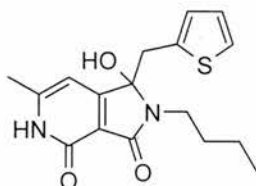


1-Ethoxy-nonane-2,4-dione (**81**) (2 g, 10 mmol) was added to cyanoacetamide (**54**) (0.8 g, 10 mmol) and dissolved in the minimum amount of ethanol (150 cm³) and heated to 60°C. Addition of piperidine (1.0 cm³, 10 mmol, 1 eq) occurred and the reaction mixture was allowed to cool to room temperature overnight. A white solid precipitated out of solution. The white crystals were filtered and washed with ethanol (3 x.50 cm³). To yield the desired compound **87** as a white solid (0.55 g, 22%), mp 153-154 °C.

IR (KBr) 2870, 2220 (C≡N), 1649 (C=O), 1614, 1540, 1486, 1455, 1371, 1351, 1337, 1229, 1196, 1137, 1118, 1060, 954, 836, 665, 569 cm⁻¹; ¹H NMR (CD₃OD, 300.1 MHz) 6.25 (1H, s, H5), 4.70 (2H, s, OCH₂), 3.66 (2H, q, *J* = 7.0 Hz, OCH₂CH₃), 2.66 (2H, t, *J* = 7.0 Hz, COCH₂CH₂), 1.73–1.69 (2H, m, COCH₂CH₂), 1.42–1.39 (4H, m, CH₂CH₂CH₂CH₂), 1.31 (3H, t, *J* = 8.0 Hz, CH₂CH₃), 1.00 (3H, t, *J* = 8.0 Hz, CH₂CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 167.0 (C2), 161.4 (C6), 161.1 (C3), 156.8 (C4), 115.4 (CN), 103.4 (C5), 69.0 (CCH₂OCH₂), 66.4 (OCH₂CH₃), 33.0 (CCH₂CH₂), 30.9 (CCH₂CH₂), 28.2 (CH₂CH₂CH₃), 22.1 ((CH₂)₃CH₂CH₃), 15.2

(OCH₂CH₃), 14.1 (CH₂CH₂CH₃); *m/z* (ES+) 519 ([2M+Na⁺], 14%), 271 ([M+Na⁺], 100); exact mass calcd for C₁₄H₂₀N₂O₄Na 271.1422 (M + Na⁺) found 271.1425.

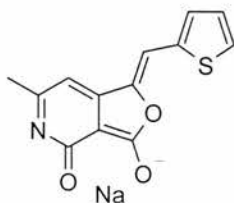
6.2.20 Synthesis of 2-butyl-1-hydroxy-6-methyl-1-(2-thienylmethyl)-1*H*-pyrrolo[3,4-*c*]pyridine-3,4(2*H*,5*H*)-dione (**93**).



6-Methyl-3,4-dioxo-1-(1-thienylidene)furo[3.4]pyridene (**44**) (0.15 g, 0.58 mmol) was dissolved in DMF (50 cm³). Butylamine was added (0.057 cm³, 0.58 mmol) and the reaction mixture stirred overnight at room temperature. The reaction mixture was concentrated *in vacuo* and the crude compound was purified by flash column chromatography eluting with dichloromethane and methanol to give the desired compound **93** as a white solid (0.034 g, 18%), mp 177-180 °C.

IR (KBr) 3258, 2958, 1707 (C(3)=O), 1637, 1609 (C(4)=O), 1573, 1412, 1259, 1128, 1055, 813, 698, 580 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 11.8 (1H, s, NH), 7.23 (1H, dd, *J*= 1.5 5.0 Hz, H5'), 6.81 (1H, dd, *J*= 3.5 5.0 Hz, H3'), 6.67-6.66 (1H, m, H4'), 6.57 (1H, s, H7), 6.33 (1H, s, OH), 3.53 (1H, d, *J*= 15.0 Hz, COHCH₂), 3.43 (1H, d, *J*= 15.0 Hz, COHCH₂), 3.43-3.10 (2H, m, NCH₂), 2.26 (3H, s, CH₃), 1.64-1.54 (2H, m, NCH₂CH₂), 1.37-1.24 (2H, m, CH₂CH₃), 0.90 (3H, t, *J* 7.0, CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.9 (C4), 163.7 (C3), 157.6 (C3a), 151.9 (C6), 136.1 (C2'), 127.2 (C4'), 126.2 (C5'), 125.1 (C3'), 114.7 (C7a), 99.3 (C7), 88.7 (COH), 38.1 (NCH₂), 35.6 (COHCH₂), 31.2 (CH₂CH₂CH₃), 20.0 (CH₂CH₃), 19.1 (CH₃), 13.8 (CH₂CH₃); *m/z* (ES+) 687 ([2M+Na], 98%), 355 ([M+Na], 100); exact mass calcd for C₁₇H₂₀N₂O₃NaS 355.1092 (M + Na+) found 355.1094.

6.2.21 Synthesis of sodium 6-methyl-4-oxo-1-(2-thienylmethylene)-1,4-dihydrofuro[3,4-*c*]pyridin-3-olate (**97**).



To a suspension of 6-methyl-3,4-dioxo-1-(1-thenylidene)furo[3.4]pyridene (**44**) (0.05 g, 0.19 mmol) in methanol (10 cm³) sodium methoxide was added (0.018 g, 0.33 mmol). The reaction mixture was refluxed for 2 hours, during which time a precipitate was continuously visible. The reaction was filtered and the yellow precipitate collected and identified as the sodium salt **97** (0.017 g, 33%), mp 272-273 °C (decomposes).

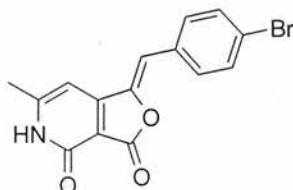
IR (KBr) 2926, 2833, 1777 (C(3)=O), 1644 (C(4)=O), 1614 (C=C), 1420, 1275, 1215, 1132, 937, 799, 633 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.67 (1H, d, *J*= 5.0 Hz, H5'), 7.33 (1H, d, *J*=4.0 Hz, H3'), 7.13 (1H, dd, *J*= 5.0, 4.0 Hz, H4'), 6.89 (1H, s, H8), 6.36 (1H, s, H7), 2.22 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 168.1 (C4), 167.6 (C3), 164.9 (C6) 151.3 (C3a), 142.8 (C1), 136.6 (C2'), 129.0 (C3' & C5'), 127.8 (C4'), 99.7 (C8), 98.6 (C7a), 95.6 (C7), 25.0 (CH₃); *m/z* (ES⁺) 541 ([2xM+Na⁺], 43%), 282 ([M+Na⁺], 100); exact mass calcd for C₁₃H₉NO₃NaS 282.0218 (M + Na⁺) found 282.0212.

6.2.22 General procedure for parallel synthesis of derivatives of 6-methyl-1-[1-thenylidene]furo[3,4-*c*]pyridine-3,4(*1H,5H*)-dione (**44**)

Lactone (**48**) (0.15 g, 0.91 mmol) was placed in each reaction tube. The relevant aldehyde (1.1 mmol) was added to each tube and the tubes placed in the Büchi syncore. The dispensing device was used to divide a solution of piperidine in methanol (360 cm³, 0.06 M) equally between the 24 tubes. The reactions were heated at reflux (the temperature of the syncore was set to 85 °C) for a total of 15 hours. Isolation of the solid precipitates was achieved through the use of a filtration device to remove the reaction solution. The precipitates were washed with methanol (3 x 240 cm³) and any remaining solvent removed by concentrating *in vacuo*, to yield the

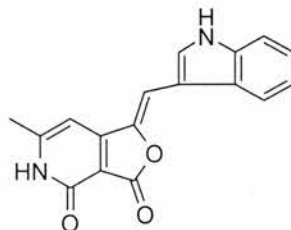
products. The initial filtrate was also concentrated *in vacuo*, which yielded compounds **120**, **122-136** and **138-141**.

6.2.23 Synthesis of 1-[4-bromo-benzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**120**).



Yellow solid (0.163 g, 31%), mp 318-319 °C (decomposes); IR (KBr) 3453, 2929, 2839, 1772 (C(3)=O), 1674 (C(4)=O), 1651, 1617 (C=C), 1490, 1129, 947, 812, 638, 567, 476 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.28 (1H, brs, NH), 7.77-7.71 (2H, m, AA' part of the AA'BB' system, H2' & H6'), 7.72-7.68 (2H, m, BB' part of the AA'BB' system, H3' & H5'), 6.95 (1H, s, H8), 6.76 (1H, s, H7); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 163.9 (C4), 157.6 (C3), 156.1 (C6), 155.6 (C3a), 143.2 (C1), 132.1 (C2', C3', C5' & C6'), 131.9 (C1'), 122.7 (C4'), 110.0 (C8), 105.5 (C7a), 96.9 (C7), 19.6 (CH₃); m/z (ES+) 353 ([M+Na], 100); exact mass calcd for C₁₅H₁₀NO₃79BrNa 353.9742 (M + Na⁺) found 353.9738, C₁₅H₁₀NO₃81BrNa 355.9721 (M + Na⁺) found 355.9723. Anal. Calcd for C₁₅H₁₀NO₃Br: C, 54.24; H, 3.03; N, 4.22, found C, 54.05; H, 2.74; N, 4.19; λ_{max} (DMSO)/nm 359 (27323 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{max} fluor (DMSO) 447 nm.

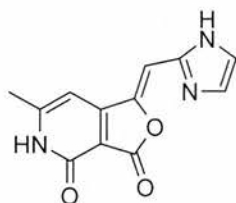
6.2.24 Synthesis of 1-[1*H*-indol-3-ylmethylene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**122**).



Orange solid (0.176 g, 37%), mp 323-324 °C (decomposes); IR (KBr) 3314, 2949, 1763 (C(3)=O), 1621 (C(4)=O), 1526, 1276, 1128, 1111, 985, 938, 835, 801, 748, 734, 627, 610, 562 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 11.95 (2H, br s, NH),

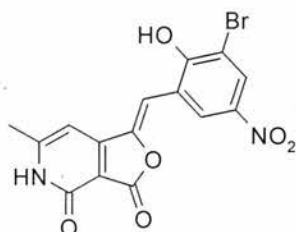
8.05–8.03 (1H, m, H7'), 8.03 (1H, s, H2'), 7.52–7.49 (1H, m, H4'), 7.37 (1H, s, H8), 7.27–7.18 (2H, m, H5' & 6'), 6.85 (1H, s, H7), 2.35 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.4 (C4), 158.1 (C3), 155.3 (C6), 153.9 (C3a), 138.9 (C1), 135.9 (C7a'), 130.2 (C2'), 126.4 (C3'), 122.5 (C6'), 120.5 (C7'), 118.5 (C4'), 112.2 (C5'), 109.1 (C3a'), 106.3 (C8), 104.3 (C7a), 96.4 (C7), 19.5 (CH₃); *m/z* (ES+) 607 ([2M+Na⁺], 38%), 315 ([M+Na⁺], 100); exact mass calcd for C₁₇H₁₂N₂O₃Na 315.0746 (M + Na⁺) found 315.0747; Anal. Calcd for C₁₇H₁₂N₂O₃: C, 69.86; H, 4.14; N, 9.58, found C, 69.45; H, 4.09; N, 9.37; λ_{max}(DMSO)/nm 432 (19710 ε/dm³ mol⁻¹ cm⁻¹); λ_{max} fluor (DMSO) 508 nm.

6.2.25 Synthesis of 1-[1*H*-imidazol-2-ylmethylidene]-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**123**).



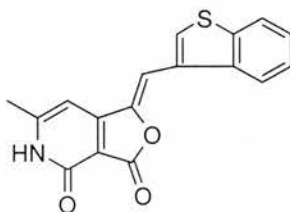
Brown solid (0.53 g, 12%), mp 322–323 °C (decomposes); IR (KBr) 3360, 3268, 1802 (C(3)=O), 1697 (C(4)=O), 1645, 1611, 1552, 1141 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.18 (1H, br s, NH), 7.35 (1H, s, NH), 7.18 (1H, s, ArH), 7.02 (1H, s, ArH), 6.96 (1H, s, H8), 6.89 (1H, s, H7), 2.34 (3H s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 157.6 (C3), 155.5 (C6), 141.8 (C2'), 140.2 (C1), 130.7 (CH), 121.0 (CH), 101.5 (C8), 97.0 (C7), 19.59 (CH₃) (missing 3 carbons due to compound insolubility); *m/z* (ES+) 244 ([M+H⁺], 100%); exact mass calcd for C₁₂H₉N₃O₃Na 266.0542 (M + Na⁺) found 266.0540; λ_{max}(DMSO)/nm 375 (23863 ε/dm³ mol⁻¹ cm⁻¹) and 307 (21213); λ_{max} fluor (DMSO) 485 nm.

6.2.26 Synthesis of 1-[3-bromo-2-hydroxy-5-nitrobenzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (124).



Red solid (0.326 g, 55%), mp 267-268 °C (decomposes); IR (KBr) 2725, 1765 (C(3)=O), 1641 (C(4)=O), 1614 (C=C), 1575, 1510, 1431, 1260, 1206, 1123, 968, 803, 699, 631, 589 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.03 (1H, brs, NH), 8.77 (1H, d, $J=3.0$ Hz, H4'), 8.18 (1H, d, $J=3.0$ Hz, H6'), 7.15 (1H, s, H8), 6.68 (1H, s, H7), 2.34 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 171.7 (C2'), 164.4 (C4), 157.9 (C3), 156.0 (C3a), 154.6 (C6), 140.4 (C1), 128.9 (C4'), 128.9 (C5'), 127.6 (C6'), 119.8 (C3'), 116.6 (C1'), 109.0 (C8), 104.3 (C7a), 96.3 (C7), 19.3 (CH₃); m/z (ES+) 480 ([M+2xNa⁺+K⁺], 95%), 478 ([M+2xNa⁺+K⁺], 100), 417 ([M+Na⁺], 43), 415 ([M+Na⁺], 43); exact mass calcd for C₁₅H₈N₂O₆⁷⁹Br 390.9566 (M - H⁺) found 390.9568, C₁₅H₈N₂O₆81Br 392.9545 (M - H⁺) found 392.9555; λ_{max} (DMSO)/nm 503 (19638 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$) and 429 (22662); λ_{max} fluor (DMSO) 590 nm.

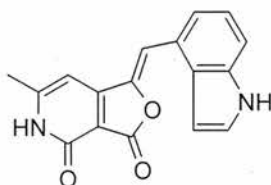
6.2.27 Synthesis of 1-[1-benzothien-3-ylmethylene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (125).



Yellow solid (0.215 g, 43%), mp 322-323 °C (decomposes); IR (KBr) 2936, 1782 (C(3)=O), 1670 (C(4)=O), 1655, 1617 (C=C), 1129, 956, 933, 799, 759, 632, 425 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.19 (1H, br s, NH), 8.40 (1H, s, H2'), 8.37 (1H, dd $J=7.0$ Hz, H7'), 8.09 (1H, dd, $J=7.0$ Hz, H4'), 7.57-7.53 (1H, m, H5'), 7.50-7.45 (1H, m, H6'), 7.38 (1H, s, H8), 7.05 (1H, s, H7), 2.37 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 164.1 (C4), 157.7 (C3), 156.0 (C3a), 155.3 (C6), 143.6 (C1), 138.8 (C7a'), 137.8 (C3a'), 131.1 (C2'), 127.7 (C3'), 125.1 (C6'), 124.8 (C5'), 123.0

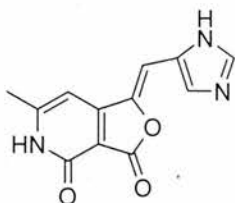
(C4'), 121.8 (C7'), 105.5 (C7a), 102.6 (C8), 97.2 (C7), 19.7 (CH₃); *m/z* (ES⁺) 641 ([2M+Na⁺], 37%), 332 ([M+Na⁺], 100); exact mass calcd for C₁₇H₁₁NO₃NaS 332.0357 (M + Na⁺) found 332.0346; Anal. Calcd for C₁₇H₁₁NO₃S: C, 66.01; H, 3.58; N, 4.53, found C, 66.04; H, 3.25; N, 4.40; λ_{\max} (DMSO)/nm 382 (9992 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 481 nm.

6.2.28 Synthesis of 1-[1*H*-indol-4-ylmethylene]-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**126**).



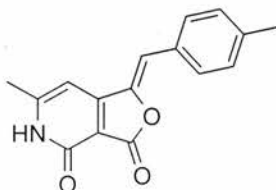
Orange solid (0.247 g, 52%), mp 302-303 °C (decomposes); IR (KBr) 3342, 1748 (C(3)=O), 1684 (C(4)=O), 1647 (C=O), 1598 (C=C), 1348, 1318, 1218, 1133, 1121, 982, 892, 802, 734, 619, 573 cm^{-1} ; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.08 (1H, br s, NH), 11.40 (1H, s, NH), 8.10 (1H, m, H4'), 7.66 (1H, dd, *J*= 9.0, 2.0 Hz, H7'), 7.51 (1H, d, *J*= 8.0 Hz, H8'), 7.44 (1H, t, *J*= 3.0, 3.0 Hz, H3'), 7.04 (1H, s, H8), 6.78 (1H, s, H7), 6.59–6.57 (1H, m, H2'), 2.35 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.5 (C4), 157.9 (C3a), 156.3 (C3), 154.6 (C6), 140.2 (C1), 136.4 (C7a'), 128.1 (C3a'), 126.8 (6'), 124.0 (C2'), 123.8 (C7'), 123.8 (C4'), 114.3 (C8), 112.1 (C3'), 104.7 (C7a), 102.2 (5'), 96.5 (C7), 19.5 (CH₃); *m/z* (ES⁺) 607 ([2M+Na⁺], 36%), 315 ([M+Na⁺], 100); exact mass calcd for C₁₇H₁₂N₂O₃Na 315.0746 (M + Na⁺) found 315.0734; Anal. Calcd for C₁₇H₁₂N₂O₃: C, 69.86; H, 4.14; N, 9.58, found C, 69.58; H, 3.87; N, 9.41; λ_{\max} (DMSO)/nm 399 (55957 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 496 nm.

6.2.29 Synthesis of 1-[*1H*-imidazol-5-ylmethylene]-6-methylfuro[3,4-*c*]pyridine-3,4(*1H,5H*)-dione (**127**).



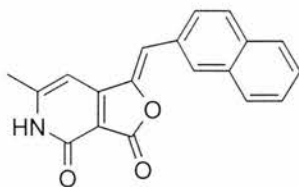
Dark yellow solid (0.190 g, 45%), mp 282-283 °C (decomposes); IR (KBr) 3155, 1754 (C(3)=O), 1654 (C(4)=O), 1612 (C=C), 1262, 1138, 1115, 949, 804, 636, 562 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 7.79 (1H, s, H8), 7.64 (1H, s, H2'), 6.97 (1H, s, H4'), 6.91 (1H, s, H7), 2.31 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 155.3 (C6), 97.7 (C7), 19.5 (CH₃) (missing 9 carbons due to insolubility); m/z (ES+) 244 ([M+H⁺], 100); exact mass calcd for C₁₂H₉N₃O₃Na 266.0542 (M + Na⁺) found 266.0533; λ_{max} (DMSO)/nm 378 (16948 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{max} fluor (DMSO) 464 nm.

6.2.30 Synthesis of 6-methyl-1-[4-methylbenzylidene]furo[3,4-*c*]pyridine-3,4(*1H,5H*)-dione (**128**).



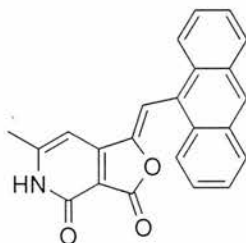
Yellow solid (0.285 g, 53%), mp 262-263 °C (decomposes); IR (KBr) 2922, 2835, 1784 (C(3)=O), 1672 (C(4)=O), 1651 (C=C), 1606, 1134, 967, 944, 811, 583 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.20 (1H, brs, NH), 7.72-7.69 (2H, m, AA' part of the AA'XX' system, H3' & 5'), 7.31-7.28 (2H, m, XX' part of the AA'XX' system, H2' & 6'), 6.94 (1H, s, H8), 6.77 (1H, s, H7), 2.37 (3H, s, CH₃ 4'), 2.36 (3H, s, CH₃ 6'); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 164.6 (C4), 159.1 (C3), 156.5 (C6), 155.7 (C3a), 142.4 (C1), 139.1 (C4'), 130.3 (C3' & 5'), 130.1 (C1'), 129.6 (C2' & 6'), 110.6 (C8), 104.3 (C7a), 96.5 (C7), 21.0 (Ar-CH₃) 20.3 (CH₃); m/z (ES+) 557 ([2 x M+ Na⁺], 55%), 289 ([M+Na⁺], 100); exact mass calcd for C₁₆H₁₃NO₃Na 290.0793 (M + Na⁺) found 290.0794; λ_{max} (DMSO)/nm 361 (27000 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{max} fluor (DMSO) 453 nm.

6.2.31 Synthesis of 6-methyl-1-[2-naphthylmethylene]furo[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**129**).



Yellow solid (0.208 g, 42%), mp 319-320 °C (decomposes); IR (KBr) 2924, 1783 (C(3)=O), 1649 (C(4)=O), 1619 (C=C), 1566, 1139, 1125, 962, 942, 899, 810, 744, 643, 577 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.27 (1H, br s, NH), 82.8 (1H, s, H1'), 8.05–7.94 (4H, m, H3', H4', H5' & H8'), 7.63–7.57 (2H, m, H6' & H7'), 7.13 (1H, s, H8), 6.83 (1H, s, H7), 2.38 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.1(C4), 157.2(C3), 156.2 (C3a), 155.5 (C6), 143.0 (C2'), 132.8 (C4a'), 130.5 (C1'), 130.4 (C8a'), 128.5 (CH), 128.5 (CH), 127.6 (CH), 127.4 (CH), 127.0 (CH), 126.8 (CH), 111.4 (C8), 105.4 (C7a), 96.9 (C7), 19.6 (CH₃); *m/z* (ES+) ([2M+Na⁺], 36%), 326 ([M+Na⁺], 100); exact mass calcd for C₁₉H₁₃NO₃Na 326.0793 (M + Na⁺) found 326.0785; Anal. Calcd for C₁₉H₁₃NO₃: C, 75.24; H, 4.32; N, 4.62, found C, 75.06; H, 4.23; N, 4.51; λ_{max}(DMSO)/nm 374 (22569 ε/dm³ mol⁻¹ cm⁻¹); λ_{max} fluor (DMSO) 455 nm.

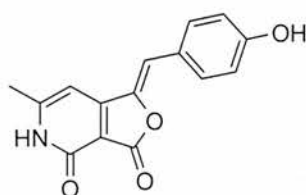
6.2.32 Synthesis of 1-[9-anthrylmethylene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**130**).



Orange solid (0.286 g, 53%), mp 327-328 °C (decomposes); IR (KBr) 2924, 2833, 1787 (C(3)=O), 1664 (C(4)=O), 1614 (C=C), 1570, 1126, 958, 940, 817, 739, 628 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.33 (1H, br s, NH), 8.74 (1H, s, H10), 8.21–8.09 (4H, m, H1', H4', H5' & H8'), 7.93 (1H, s, H8), 7.62–7.57 (4H, m, H2', H3', H6' & H7'), 7.18 (1H, s, H7), 2.45 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 163.9 (C4), 157.6 (C3), 155.8 (C6), 155.2 (C3a), 145.3 (C9), 130.8 (C8a' & C9a'),

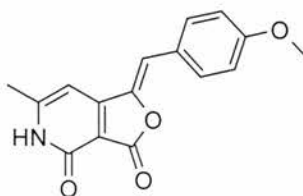
129.1 (C4a' & C10a'), 128.7 (CH), 128.1 (CH), 126.4 (CH), 125.9 (CH), 125.6 (CH), 108.5 (C8), 106.7 (C7a), 97.7 (C7), 19.7 (CH₃); *m/z* (ES⁺) 729 ([2M+Na⁺], 41%), 376 ([M+Na⁺], 100); exact mass calcd for C₂₃H₁₆NO₃ 354.1130 (M + H⁺) found 354.1140; Anal. Calcd for C₂₃H₁₅NO₃: C, 78.18; H, 4.28; N, 3.96, found C, 78.06; H, 4.22; N, 3.87; λ_{\max} (DMSO)/nm 389 (13759 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 557 nm.

6.2.33 1-[4-hydroxybenzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (131).



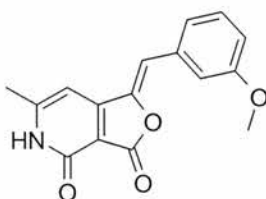
Yellow solid (0.211 g, 46%), mp 330-331 °C (decomposes); IR (KBr) 3137, 1784 (C(3)=O), 1645 (C(4)=O), 1604 (C=C), 1514, 1281, 1174, 1149, 979, 582 cm^{-1} ; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.70-7.65 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.90-6.85 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.87 (1H, s, H8), 6.72 (1H, s, H7), 3.18 (1H, s, OH), 2.34 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.3 (C4), 159.0 (C4'), 157.8 (C3), 150.2 (C3a), 154.7 (C6), 140.3 (C1), 132.5 (C2' & C6'), 123.8 (C1'), 116.1 (C3' & C5'), 112.3 (C8), 104.8 (C7a), 96.5 (C7), 19.5 (CH₃); *m/z* (ES⁺) 561 ([2xM+Na⁺], 13), 308 ([M+K⁺], 17), 292 ([M+Na⁺], 100); exact mass calcd for C₁₅H₁₂NO₄ 270.0788 (M + Na⁺) found 270.0771; λ_{\max} (DMSO)/nm 377 (22904 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 479 nm.

6.2.34 Synthesis of 1-[4-methoxybenzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**132**).



Yellow solid (0.186 g, 40%), mp 279-280 °C (decomposes); IR (KBr) (KBr) 2368, 2340, 1792 (C(3)=O), 1669 (C(4)=O), 1650 (C=O), 1614 (C=C), 1570, 1561, 1252, 1183, 1139, 1028, 972, 942, 821, 583 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.13 (1H, brs, NH), 7.79-7.74 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.09-7.04 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.93 (1H, s, H8), 6.74 (1H, s, H7), 3.84 (3H, s, OCH₃), 2.35 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 164.2 (C4), 160.2 (C4'), 157.8 (C3), 156.2 (C3a), 154.9 (C6), 140.9 (C1), 132.2 (C2' & 6'), 125.3 (C1'), 114.6 (C3' & C5'), 111.7 (C8), 104.9 (C7a), 96.6 (C7), 55.3 (OCH₃), 19.5 (CH₃); m/z (ES+) 589 ([2xM+Na⁺], 32%), 306 ([M+Na⁺], 100); exact mass calcd for C₁₆H₁₃NO₄Na 306.0742 (M + Na⁺) found 306.0746. Anal. Calcd for C₁₆H₁₃NO₄: C, 67.84; H, 4.63; N, 4.94, found C, 67.75; H, 4.46; N, 4.90; λ_{max} (DMSO)/nm 372 (19073 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{max} fluor (DMSO) 468 nm.

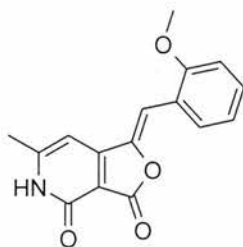
6.2.35 Synthesis of 1-[3-methoxybenzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**133**).



Yellow solid (0.128 g, 27%), mp 293-294 °C (decomposes); IR (KBr) (KBr) 2833, 1785 (C(3)=O), 1671 (C(4)=O), 1654, 1613 (C=C), 1251, 1134, 973, 958, 868, 801, 627, 595 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.26 (1H, br s, NH), 7.46-7.37 (3H, m, H4', H5' & H6'), 7.05-6.70 (1H, m, H2'), 6.93 (1H, s, H-7), 6.77 (1H, s, H-5), 3.82 (3H, s, OCH₃), 2.36 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 164.0 (C4), 159.4 (C3'), 157.7 (C3), 156.3 (C3a), 155.5 (C6), 142.8 (C1), 133.9 (C1'), 130.1 (C5'), 122.9 (C6'), 115.0 (C4'), 111.2 (C2'), 105.5 (C7a), 96.8 (C7), 55.1

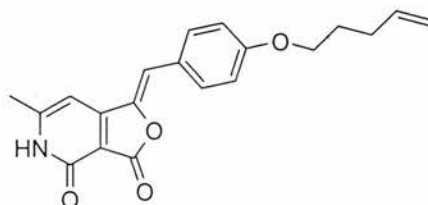
(OCH₃), 19.6 (CH₃); m/z (ES+) 589 ([2M+Na⁺], 56%), 306 ([M+Na⁺], 100); exact mass calcd for C₁₆H₁₄NO₄ 284.0923 (M + H⁺) found 284.0922; Anal. Calcd for C₁₆H₁₃NO₄: C, 67.84; H, 4.63; N, 4.94, found C, 67.64; H, 4.49; N, 4.93; λ_{\max} (DMSO)/nm 360 (42479 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 443 nm.

6.2.36 Synthesis of 1-[2-methoxybenzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**134**).



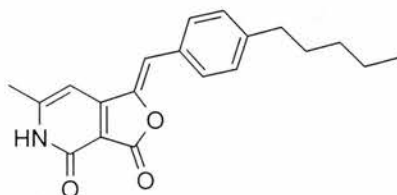
Yellow solid (0.225 g, 48%), mp 329-330 °C (decomposes); IR (KBr) 2830, 1785 (C(3)=O), 1661 (C(4)=O), 1642, 1614 (C=C), 1247, 1132, 1016, 962, 940, 804, 762, 626, 597 cm^{-1} ; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 8.05 (1H, dd $J= 7.5, 1.5$ Hz, H6'), 7.43 (1H, td $J= 7.5, 1.5$ Hz, H4'), 7.16–7.13 (1H, m, H3'), 7.13-7.07 (1H, m, H5'), 7.02 (1H, s, H8), 6.84 (1H, s, H7), 3.01 (3H, s, OCH₃), 2.36 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 157.2 (C2'), 155.6 (C6), 142.4 (C1), 131.9 (C1'), 130.7 (C4' & C6'), 120.0 (C5'), 111.5 (C3'), 103.6 (C8), 96.5 (C7), 55.8 (OCH₃), 19.5 (CH₃) (missing 4 carbons due to insolubility); m/z (ES+) 589 ([2M+Na⁺], 24%), 306 ([M+Na⁺], 100); exact mass calcd for C₁₆H₁₄NO₄Na 284.0923 (M + Na⁺) found 284.0919; Anal. Calcd for C₁₆H₁₃NO₄: C, 67.84; H, 4.63; N, 4.94, found C, 67.66; H, 4.01; N, 4.61 (differing by 0.6%); λ_{\max} (DMSO)/nm 375 (24163 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$), λ_{\max} fluor (DMSO) 476 nm.

6.2.37 Synthesis of 1-[4-(but-3-en-yloxy)benzylidene]-6-methylfuro[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**135**).



Yellow solid (0.004 g, 1%) mp >300 °C (decomposes); IR (KBr) 1787 (C(3)=O), 1668 (C(4)=O), 1644, 1601, 1513, 1251, 1179, 1135, 1023, 944, 872, 807, 592 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.78-7.75 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.08-7.04 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.88 (1H, s, H8) 6.71 (1H, s, H7), 5.96-5.82 (1H, m, CH₂CHCH₂) 5.20-5.04 (H, m, CHCH₂), 4.80 (2H, t, *J*= 7.0 Hz, OCH₂), 3.42 (2H, CHCH₂CH under DMSO), 2.32 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.1 (C4), 159.5 (C4'), 157.7 (C3), 156.2 (C6), 154.9 (C3a), 141.0 (C1), 134.6 (CH₂CHCH₂), 132.2 (C2' & C6'), 125.3 (C1'), 117.0 (CHCH₂), 115.1 (C3' & C5'), 111.5 (C8), 105.0 (C7a), 96.4 (C7), 66.9 (OCH₂), 32.9 (CH₂), 19.5 (CH₃); *m/z* (ES⁺) 346 ([M+ Na⁺], 100), 188 (10) 85 (10). exact mass calcd for C₁₉H₁₈NO₄ 324.1236 (M + H⁺) found 324.1240; λ_{max}(DMSO)/nm 372 (25606 ε/dm³ mol⁻¹ cm⁻¹); λ_{max} fluor (DMSO) 470 nm.

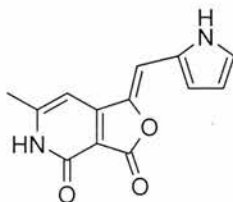
6.2.38 Synthesis of 6-methyl-1-[4-pentylbenzylidene]furo[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**136**).



Yellow solid (0.197 g, 38%), mp >288 °C (decomposes); IR (KBr) 2928, 1789 (C(3)=O), 1671 (C(4)=O), 1651, 1606 (C=C), 1133, 943, 802, 591 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.19 (1H, br s, NH), 7.73-7.70 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.30-7.32 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.91 (1H, s, H8), 6.75 (1H, s, H7), 2.60 (2H, t, *J*= 3.0 Hz, Ar-CH₂CH₂), 2.34 (3H, s, CH₃), 1.63-1.53 (2H, m, CH₂CH₂CH₂ CH₃), 1.35-1.24 (4H, m,

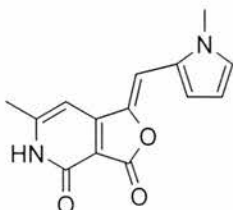
CH₂CH₂CH₃), 0.860 (3H, t, $J=7.0$, Hz, CH₂CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.1 (C4), 157.7 (C3), 156.3 (C3,a), 155.3 (C6), 144.2 (C4'), 142.1 (C1), 130.4 (C2' & C6'), 130.2 (C1'), 129.0 (C3' & C5'), 111.6 (C8), 105.3 (C7a), 96.7 (C7), 35.0 (Ar-CH₂), 30.8 (CH₂CH₂CH₂), 30.3 (CH₂CH₂CH₂CH₃), 21.9 (CH₂CH₃), 19.6 (CH₃), 13.9 (CH₂CH₃); m/z (ES+) 669 ([2M+Na⁺], 70%), 346 ([M+Na⁺], 100); exact mass calcd for C₂₀H₂₁NO₃Na 346.1419 (M + Na⁺) found 346.1419; Anal. Calcd for C₂₀H₂₁NO₃: C, 74.28; H, 6.55; N, 4.33, found C, 74.05; H, 6.18; N, 4.33; λ_{\max} (DMSO)/nm 362 (29241 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 437 nm.

6.2.39 Synthesis of 6-methyl-1-[1*H*-pyrrol-2-ylmethylene]furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**138**).



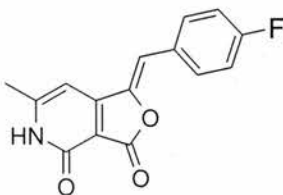
Brown solid (0.122 g, 29%), mp >280 °C (decomposes); IR (KBr) 3309 (NH), 3218 (NH), 2366, 2329, 1782 (C=O, lactone), 1641 (C=O, pyridione), 1606 (C=C), 1419, 1407, 1282, 1228, 1128, 1112, 1032, 981, 740, 693 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 11.92 (1H, brs, NH), 11.23 (1H, brs, NH), 7.09-7.07 (1H, m, H5'), 6.83 (1H, s, H8), 6.70-6.68 (1H, m, H3'), 6.62 (1H, s, H8), 6.28-6.25 (1H, m, H4'), 2.32 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.1 (C4), 157.9 (C3), 155.07 (C3a), 154.2 (C6), 137.6 (C1), 125.5 (C2'), 124.3 (C5'), 115.7 (C3'), 110.8 (C4'), 104.3 (C7a), 104.0 (C8), 96.0 (C7), 19.4 (CH₃); m/z (ES+) 507 ([2 x M+Na⁺], 40%), 265 ([M+Na⁺], 100); exact mass calcd for C₁₃H₁₀N₂O₃Na 265.0589 (M + Na⁺) found 265.0583; λ_{\max} (DMSO)/nm 429 (16037 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 495 nm.

6.2.40 Synthesis of 6-methyl-1-[(1-methyl-1*H*-pyrrol-2-yl)methylene]furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**139**).



Orange solid (0.168 g, 35%), mp 296-297 °C (decomposes); IR (KBr) 2933, 1752 (C(3)=O), 1652 (C(4)=O), 1621 (C=C), 1482, 1275, 1145, 1064, 976, 943, 801, 708, 631, 567 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 11.97 (1H, brs, NH), 7.08 (1H, t, *J*= 2.0 Hz, H5'), 6.94 (1H, s, H8), 6.89 (1H, dd, *J*= 4.0, 2.0 Hz, H3'), 6.32 (1H, s, H7), 6.26 (1H, dd, *J*=4.0, 2.0 Hz, H4'), 3.36 (3H, s, NCH₃), 2.32 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.3 (C4), 157.9 (C3), 155.5 (C3,a), 154.0 (C6), 138.4 (C1), 127.6 (C5'), 126.5 (C2'), 116.1 (C3'), 110.0 (C4'), 104.3 (C7a), 101.9 (C8), 96.5 (C7), 33.8 (N-CH₃), 19.5 (CH₃); *m/z* (ES⁺) 535 ([2 x M+Na⁺], 30%), 293 ([M+Na⁺], 100); exact mass calcd for C₁₄H₁₂N₂O₃Na 279.0746 (M + Na⁺) found 279.0740; Anal. Calcd for C₁₄H₁₂N₂O₃: C, 65.62; H, 4.72; N, 10.93, found C, 65.54; H, 4.55; N, 10.90; λ_{max}(DMSO)/nm 432 (20720 ε/dm³ mol⁻¹ cm⁻¹); λ_{max} fluor (DMSO) 502 nm.

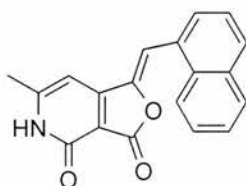
6.2.41 Synthesis of 1-[4-fluorobenzylidene]-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**140**).



Orange solid (0.081 g, 17%), mp 295-296 °C (decomposes); IR (KBr) 2831, 1783 (C(3)=O), 1648 (C(4)=O), 1615 (C=C), 1506, 1229, 1140, 977, 946, 827, 628, 580 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.25 (1H, br s, NH), 7.90-7.82 (2H, m, H2' & H6'), 7.38-7.30 (2H, m, H3' & H5'), 6.98 (1H, s, H8), 6.76 (1H, s, H7), 2.36 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.0 (C4), 163.9 (C4'), 160.6 (C4'), 157.7 (C3), 156.2 (C3a), 155.5 (C6), 142.4 (C1), 132.7 (C2'&6'), 132.6 (C2'&6'), 129.4 (C1'), 116.3 (C2' & C6'), 116.0 (C3' & C5'), 110.1 (C8), 105.4 (C7a), 96.8 (C7),

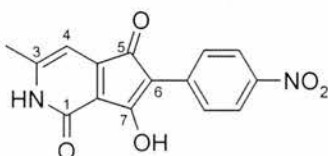
19.6 (CH₃); m/z (ES+) 282 ([M+Na⁺], 100%); exact mass calcd for C₁₅H₁₀NO₃FNa 294.0542 (M + Na⁺) found 294.0540; Anal. Calcd for C₁₅H₁₀NO₃F: C, 66.42; H, 3.72; N, 5.16, found C, 65.90; H, 3.44; N, 4.97 (differing by 0.5%); λ_{\max} (DMSO)/nm 354 (17620 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 447 nm.

6.2.42 Synthesis of 6-methyl-1-[1-naphthylmethylene]furo[3,4-c]pyridine-3,4(1*H*,5*H*)-dione (**141**).



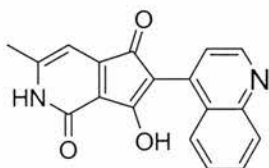
Yellow solid (0.168 g, 34%), mp 316-317 °C (decomposes); IR (KBr) 3046, 2929, 2835, 1787 (C(3)=O), 1669 (C(4)=O), 1615 (C=C), 1128, 960, 937, 801, 771, 629 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.30–12.20 (1H, br s, NH), 8.54 (1H, d, $J= 7.0$ Hz, H4'), 8.27 (1H, d, $J= 7.0$ Hz, H2'), 8.03 (2H, d, $J= 8.0$ Hz, H6' & H9'), 7.71–7.60 (2H, m, H7' & H8'), 7.71-7.60 (1H, m, H3'), 7.66 (1H, s, H8), 7.20 (1H, s, H7), 2.41 (1H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.3 (C4), 157.8 (C3), 156.3 (C3a), 155.5 (C6), 144.0 (C1), 133.3 (C4a'), 131.2 (C8a'), 129.8 (CH), 129.2 (CH), 128.8 (CH), 128.6 (CH), 126.9 (CH), 126.3 (CH), 125.8 (CH), 123.8 (CH), 107.2 (C8), 97.5 (C7), 19.7 (CH₃); m/z (ES+) 629 ([2M+Na⁺], 22%), 326 ([M+Na⁺], 100); exact mass calcd for C₁₉H₁₃NO₃Na 326.0793 (M + Na⁺) found 326.0782; Anal. Calcd for C₁₉H₁₃NO₃: C, 75.24; H, 4.32; N, 4.62, found C, 74.73; H, 4.17; N, 4.62 (differing by 0.5%); λ_{\max} (DMSO)/nm 379 (17759 $\epsilon/\text{dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$); λ_{\max} fluor (DMSO) 474 nm.

6.1.43 Synthesis of 7-hydroxy-3-methyl-6-(4-nitrophenyl)-1*H*-cyclopenta[*c*]pyridine-1,5(2*H*)-dione (**144**).



Red solid (0.04 g, 15%), mp 224-225 °C decomposes; IR (KBr) 2941, 1766 (C=O), 1597 (C=O), 1653 (C(1)=O), 1519, 1344, 1106, 947, 854, 774, 712, 582 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 8.60-8.57 (2H, m, AA' part of the AA'XX' system, H3' & H5'), 7.99-7.96 (2H, m, XX' part of the AA'XX' system, H2' & H6'), 6.11 (1H, s, H7), 2.20 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 191.9 (C7), 187.1 (C5), 157.5 (C1), 154.4 (C3), 152.9 (C7a), 145.8 (C4'), 138.8 (C1'), 123.6 (C3' & C5'), 120.7 (C2' & C6'), 117.4 (C4a), 99.5 (C6), 97.3 (C4), 19.3 (CH₃); m/z (ES-) 297 (M-H⁺, 100%); exact mass calcd for C₁₅H₉N₃O₅Na 297.0511 (M + Na⁺) found 297.0509.

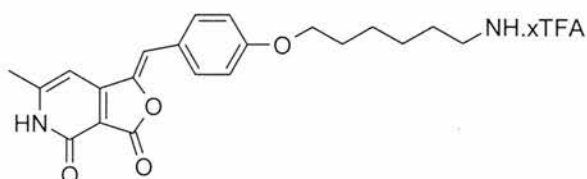
6.1.44 Synthesis of 7-hydroxy-3-methyl-6-quinolin-4-yl-1*H*-cyclopenta[*c*]pyridine-1,5(2*H*)-dione (**146**).



Yellow solid (0.08 g, 28%), mp 222-223 °C decomposes; IR (KBr) 3398, 1732 (C=O), 1652 (C=O), 1641 (C(1)=O), 1489, 1438, 1257, 1219, 1152, 804, 756, 584 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 8.52 (1H, d, $J=5.0$ Hz, H2'), 8.00 (1H, d, $J=7.0$ Hz, H8'), 7.61 (1H, d, $J=8.0$ Hz, H5'), 7.52 (1H, t, $J=7.0$ 8.0 Hz, H7'), 7.51 (1H, d, $J=5.0$, H3'), 7.27 (1H, t, $J=8.0$ Hz, H6'), 6.19 (1H, s, H7), 2.24 (3H, s, CH₃); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 190.7 (C7), 185.1 (C6), 157.9 (C1), 154.0 (C3), 152.6 (C7a), 147.4 (C2'), 147.0 (C8a'), 144.6 (C4'), 129.7 (C8'), 128.6 (C7'), 126.9 (C5'), 125.9 (C4a'), 123.3 (C6'), 119.5 (C3'), 116.7 (C4a), 100.7 (C6), 98.0 (C4), 19.4

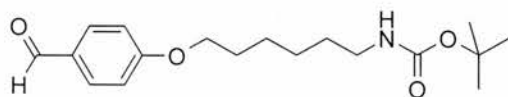
(CH₃). *m/z* (ES-) 303 (M-H⁺, 100%); exact mass calcd for C₁₈H₁₁N₂O₃ 303.0770 (M + H⁺) found 303.0778.

6.1.45 Synthesis of 1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**177**).



In the dark *tert*-butyl-(1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**179**) (0.1 g, 0.21 mmol) was dissolved in trifluoroacetic acid (10 cm³) and stirred for 1 hour at room temperature. The trifluoroacetic acid was removed by concentrating *in vacuo*, followed by repetition of the addition of toluene (5 x 20 cm³) then concentration *in vacuo*. This yielded the desired product (**177**) (partially as a TFA salt) as a yellow solid (0.08 g, >95%), mp 234-235 °C.

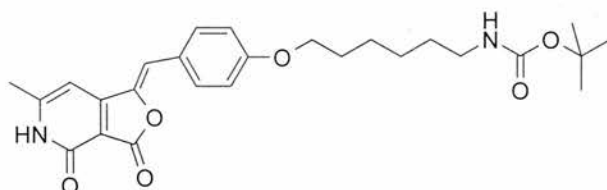
IR (KBr) 2940, 1765 (C(3)=O), 1643 (C(4)=O), 1598, 1512, 1471, 1309, 1260, 1177, 992, 828, 802, 722, 705, 632, 596 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.78-7.75 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.07-7.04 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.90 (1H, s, C8), 6.72 (1H, s, C7), 4.04 (2H, t, *J*=6.0 Hz, OCH₂CH₂), 2.84-2.74 (2H, m, NH₂CH₂CH₂), 2.33 (3H, s, CH₃), 1.76-1.68 (2H, m, OCH₂CH₂), 1.60-1.51 (2H, m, NH₂CH₂CH₂), 1.46-1.37 (4H, m, CH₂CH₂CH₂CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.2 (C4), 159.6 (C4'), 157.8 (C3), 156.2 (C3a), 154.9 (C6), 140.9 (C1), 132.3 (C2' & C6'), 125.2 (C1'), 115.0 (C3' & C5'), 111.7 (C8), 104.9 (C7a), 96.6 (C7), 67.5 (OCH₂), 28.3 (CH₂), 26.9 (CH₂), 25.4 (CH₂), 24.9 (CH₂), 19.5 (CH₃); ¹⁹F NMR (DMSO-*d*₆, 75.5 MHz) -77.6 (CF₃CO₂H); *m/z* (ES+) 737 ([2M+Na⁺], 34%), 369 ([M+Na⁺], 100); exact mass calcd for C₂₁H₂₄N₂O₄Na 369.1814 (M + Na⁺) found 369.1819; Anal. Calcd for C₂₁H₂₄N₂O₄: C, 68.46; H, 6.57; N, 7.60, Anal. Calcd for C₂₁H₂₄N₂O₄.C₂HO₂F₃: C, 56.26; H, 5.22; N, 5.81, (+ 1 mole trifluoroacetic acid), found C, 56.20; H, 5.03; N, 5.62; λ_{max}(DMSO)/nm 372 (20771 ε/dm³ mol⁻¹ cm⁻¹).

6.1.46 Synthesis of *tert*-butyl [6-(4-formylphenoxy)hexyl]carbamate (**178**).

tert-Butyl (6-hydroxyhexyl)carbamate (**181**) (0.17 g, 0.8 mmol), *p*-hydroxybenzaldehyde (**109**) and triphenylphosphene (0.1 g, 0.8 mmol) were dissolved in toluene (50 cm³). Diethyl azodicarboxylate (DEAD) (**182**) (0.130 cm³, 0.8 mmol) was added dropwise. The reaction was heated at 80 °C for 72 hrs. The reaction mixture was washed with sodium hydroxide (2 x 70 cm³) then water (1 x 70 cm³). The organic layer was dried with magnesium sulfate and concentrated *in vacuo* and purified with chromatography eluting with ethyl acetate and hexane to give the desired compound **178** as a clear oil, (0.25 g, 98%).

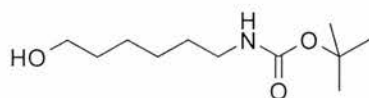
IR (KBr) 3375, 2936, 1682 (C=O), 1601 (C=O), 1508, 1262, 1159, 1015, 834, 651, 617, 515 cm⁻¹; ¹H NMR (CDCl₃, 300.1 MHz) 9.70 (1H, s, CCHO), 7.68-7.63 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.84-6.80 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 3.86 (2H, t, *J* = 6.0 Hz, CH₂O), 3.01-2.94 (2H, m, NHCH₂), 1.69-1.60 (2H, m, OCH₂CH₂), 1.42-1.20 (6H, m, CH₂CH₂CH₂), 1.29 (9H, s, CH₃); ¹³C NMR (CDCl₃, 75.5 MHz) 191.0 (CHO), 164.5 (C4'), 156.4 (NHCOO), 132.2 (C2' & C6'), 130.0 (C1'), 115.0 (C3' & C5'), 79.0 (C(CH₃)₃), 68.5 (OCH₂), 40.7 (HNCH₂), 30.3 (CH₂), 29.2 (CH₂), 28.7 (CH₃), 26.7 (CH₂), 25.9 (CH₂); *m/z* (ES+) 344 ([M+Na⁺], 100%); exact mass calcd for C₁₈H₂₇NO₄Na 344.1838 (M + Na⁺) found 344.1834.

6.1.47 Synthesis of *tert*-butyl-(1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**179**).



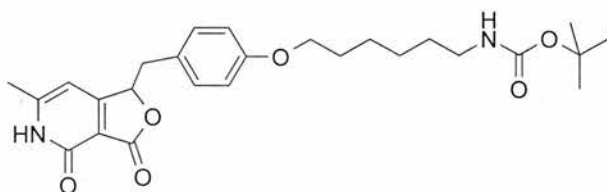
To lactone (**48**) (0.467 g, 2.8 mmol), methanol (14 cm³) was added and heated to 60°C. Piperidine (0.24 cm³, 2.4 mmol) was added followed by a solution of aldehyde (**178**) (1 g, 2.8 mmol) in methanol (26 cm³). The reaction mixture was heated at reflux overnight. This yielded a yellow solid suspended in the reaction mixture, the reaction was then allowed to cool fully to room temperature resulting in the further precipitation. The solid precipitate was filtered and washed with methanol (3 x 10 cm³), to yield the desired compound **179** as a yellow solid (0.894 g, 68%), mp >242 °C decomposes.

IR (KBr) 3361, 2937, 1795 (C(3)=O), 1677 (C=O boc), 1648 (C(4)=O), 1618, 1513, 1254, 1178, 1143, 970, 631, 591 cm⁻¹; ¹H NMR (CDCl₃, 300.1 MHz) 11.94 (1H, brs, 5-NH), 7.77-7.75 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.89-6.87 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.41 (1H, s, H7), 6.37 (1H, s, H8), 4.55 (1H, brs, NH), 3.99 (2H, t, *J*= 6.0 Hz, OCH₂), 3.18-3.08 (2H, m, NHCH₂), 2.53 (3H, s, CH₃), 1.85-1.75 (2H, m, OCH₂CH₂), 1.44 (9H, s, (CH₃)₃), 1.60-1.35 (6H, m, CH₂CH₂CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.7 (C4), 160.4 (C4'), 160.0 (C3), 156.4 (C3a), 154.1 (C6), 140.9 (C1), 132.8 (C2' & C5'), 124.9 (C1'), 114.8 (C3' & C6'), 113.0 (C8), 105.2 (C7a), 97.6 (C7), 77.4 (under CDCl₃, C(CH₃)), 67.9 (OCH₂), 40.5 (NCH₂), 29.7 (CH₂), 29.1 (CH₂), 28.4 ((CH₃)₃), 26.6 (CH₂), 25.8 (CH₂), 20.3 (CH₃) missing boc carbonyl; *m/z* (ES⁺) 959 ([2M+Na], 22%), 491 ([M+Na], 100); exact mass calcd for C₂₆H₃₂N₂O₆Na 491.2158 (M + Na⁺) found 491.2164; Anal. Calcd for C₂₆H₃₂N₂O₆: C, 66.65; H, 6.88; N, 5.98, found C, 66.13; H, 7.01; N, 5.92 (differs by 0.5%).

6.1.48 Synthesis of *tert*-butyl (6-hydroxyhexyl)carbamate (**181**)²¹⁶

6-Amino-1-hexanol (**180**) (2 g, 16.8 mmol) was dissolved in dichloromethane (160 cm³) at 0°C and di-*tert*-butyl dicarbonate (3.66 g, 16.8 mmol) added. The reaction mixture was allowed to warm to room temperature and stirred overnight. Water (120 cm³) was added and the aqueous extracted with dichloromethane (3 x 60 cm³). The combined organic extracts were dried with magnesium sulfate and concentrated *in vacuo*, to give the desired compound **181** as a clear oil, (2.5 g, 82%).

IR (NaCl) 3352, 2934, 1690 (C=O), 1515 (C=C), 1366, 1266, 1172, 739 cm⁻¹; ¹H NMR (CDCl₃, 300.1 MHz) 4.69 (1H, br s, NH), 3.57 (2H, t, *J* = 6.0 Hz, HOCH₂) 3.03 (2H, m, *J* = 7.0 Hz, CH₂NH), 2.53 (1H, s, OH), 1.51–1.26 (8H, m, CH₂CH₂CH₂CH₂), 1.37 (9H, s, COO(CH₃)₃); ¹³C NMR (CDCl₃, 75.5 MHz) 155.4 (NHCOO), 79.0 (C(CH₃)₃), 62.3 (OCH₂), 40.6 (NCH₂), 32.7 (CH₂), 30.2 (CH₂), 28.6 (CH₃), 26.7 (CH₂), 25.6 (CH₂); *m/z* (ES⁺) 457 ([M+M+Na⁺], 30%), 281 ([M+MeCN+Na⁺], 75), 240 ([M+Na⁺], 100).

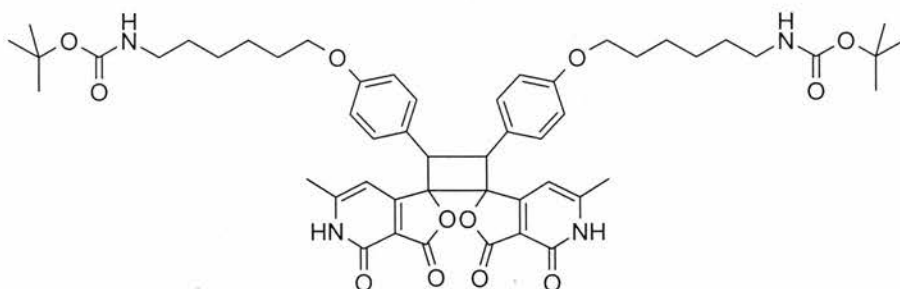
6.1.49 Synthesis of *tert*-butyl-(1-{4-[(5-aminohexyl)oxy]benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**183**).

tert-Butyl-(1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**179**) (0.6 g, 1.3 mmol) was taken with palladium on carbon (10% w/w, 0.08 g) under nitrogen. Dimethylformamide (70 cm³) and methanol (10 cm³) was added, the flask flushed with hydrogen and the reaction mixture was left stirring overnight. The palladium was removed through filtering over celtite and the reaction mixture concentrated *in vacuo*. The crude compound was

purified by flash column chromatography eluting with dichloromethane and methanol to give the desired compound **183** as a white solid (0.404 g, 67%), mp 194-195 °C.

IR (KBr) 3353, 2933, 2860, 1751 (C(3)=O), 1672 (C(4)=O), 1614, 1569, 1513, 1479, 1364, 1250, 1176, 1146, 1052, 1014, 997, 809, 723, 646, 591 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.13 (1H, brs, N(5)H), 7.09-7.06 (2H, m, AA' of a AA'XX' system, H2' & H6'), 6.81-6.78 (2H, m XX' of a AA'XX' system, H3' & H5'), 6.77-6.78 (1H, m, NH), 6.35 (1H, s, H7), 5.57-5.55 (1H, m, H1), 3.89 (2H, t $J=6.0$ Hz, OCH₂), 3.26-3.19 (1H, m, H8), 2.95-2.86 (2H, m, NCH₂), 2.90-2.88 (1H, m, H8), 2.29 (3H, s, CH₃), 1.71-1.62 (2H, m OCH₂CH₂), 1.36 (9H, s, C(CH₃)₃), 1.44-1.23 (6H, m, CH₂(CH₂)₃CH₂N); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 168.4 (C3), 167.2 (C4), 157.8 (C3a), 157.5 (C4'), 155.6 (C(O₂C(CH₃)₃)), 154.6 (C6), 130.6 (C2' & C6'), 127.0 (C1'), 114.0 (C3' & C5'), 109.4 (C7a), 99.4 (H7), 78.6 (C1), 77.3 (C(CH₃)₃), 67.2 (OCH₂), 39.2 (NHCH₂), 37.3 (C8), 29.4 (CH₂), 28.6 (CH₂), 28.2 (C(CH₃)₃), 26.0 (CH₂), 25.2 (CH₂), 19.3 (CH₃); m/z (ES+) 963 ([2M+Na], 22%), 493 ([M+Na], 100); exact mass calcd for C₂₆H₃₄N₂O₆Na 493.2315 (M + Na+) found 493.2314.

6.1.50 15,16-Bis(*tert*-butyl-(5,6-{4-[(5-aminohexyl)oxy]benzylidene})-5,13-methyl-1,9-dioxadipiro[4.0.4.2]hexadeca-2,9-(1*H*,5*H*)-dione[3,4-*c*]pyridine carbamate (**184**)

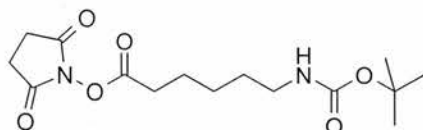


tert-Butyl-(1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**179**) (0.1 g, 0.2 mmol) was dissolved in chloroform (20 cm^3) and left in the sunlight for one month. The solution was concentrated *in vacuo* to yield the transformed compound **184**, mp 190-191 °C decomposes.

IR (KBr) 3413, 2926, 2367, 1785 (C(3)=O), 1676 (C(4)=O), 1605, 1569, 1514, 1466, 1251, 1173, 1049, 828, 580 cm^{-1} ; ^1H NMR (CDCl₃, 300.1 MHz) 7.21-7.16 (2H, m,

AA' part of the AA'XX' system, H2' & H6'), 6.85-6.82 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 5.68 (1H, s, H7), 4.75 (1H, s, H8), 4.53 (1H, brs, NH), 3.87 (2H, t, $J=6.0$ Hz, OCH₂), 3.16-3.09 (2H, m, CH₂NH), 2.49 (3H, s, CH₃), 1.78-1.68 (2H, m, OCH₂CH₂), 1.42 (9H, s, C(CH₃)₃), 1.53-1.28 (6H, m, CH₂CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 165.2 (C4), 162.3 (C3), 160.3 (C3a), 159.1 (C4'), 156.4 (C6), 129.2 (C2' & C5'), 124.7 (C1'), 114.8 (C3' & C6'), 110.8 (C7a (or1)), 102.6(C7), 89.7 (C1 (or C3,4c), 77.7 (C(CH₃)), 67.8 (OCH₂), 51.5 (C8), 40.5 (NCH₂), 30.0 (CH₂), 29.06 (CH₂), 28.4 (C(CH₃)₃), 26.5 (OCH₂CH₂), 25.7 (CH₂), 20.2 (CH₃) missing boc carbonyl; *m/z* (ES+) 960 ([M+Na⁺], 100%); exact mass calcd for C₅₂H₆₄N₄O₁₂Na 959.4418 (M + Na+) found 959.4406.

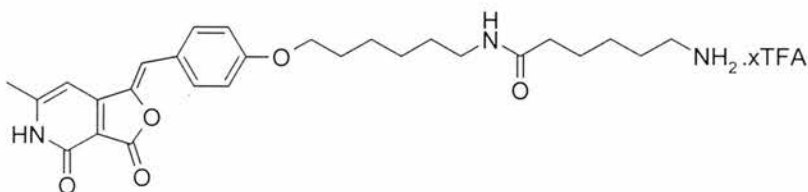
6.1.51 Synthesis of *tert*-butyl 6-[(2,5-dioxopyrrolidin-1-yl)oxy]-6-oxohexyl} carbamate (**198**).²¹⁸



6-[(*tert*-Butoxycarbonyl)amino]hexanoic acid (**201**) (0.5 g, 2 mmol), 1,3-dicyclohexylcarbodiimide (0.492 g, 2.2 mmol) and *N*-hydroxysuccinimide (0.249 g, 2.2 mmol) were dissolved in dichloromethane (50 cm³). The reaction mixture was stirred at room temperature overnight, then concentrated *in vacuo*. The crude compound was purified by flash column chromatography eluting with hexane and ethyl acetate to give the desired compound **198** as a white low melting point solid (0.561 g, 81%).

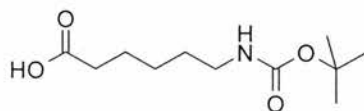
IR (KBr) 3391, 2922, 1832 (C=O), 1789 (C=O), 1731 (C=O), 1694 (C=O), 1626, 1519, 1365, 1253, 1215, 176, 1071, 993, 868, 654 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 3.69 (4H, s, (CH₂)₂CO), 3.12-3.04 (2H, m, NHCH₂), 2.32 (2H, t $J=7.0$ Hz, CH₂CO₂), 1.67-1.57 (2H, m, CH₂CH₂N), 1.52-1.31 (6H, m, CH₂(CH₂)₃CH₂), 1.41 (9H, s, C(CH₃)₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 170.2 (C2' & C5'), 168.9 (CO₂), 155.5 (CO₂(CH₃)₃), 77.3 (C(CH₃)₃), 33.3 (CO₂CH₂), 30.1 (CH₂N), 28.9 (CH₂), 28.2 (CH₃), 25.4 (C3' & C4'), 24.4 (CH₂), 23.9 (CH₂); *m/z* (ES+) 351 ([M+Na⁺], 100%).

6.1.52 Synthesis of 1-{4-(6-amino-*N*-(6-phenoxyhexyl)hexanamide)benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**200**).



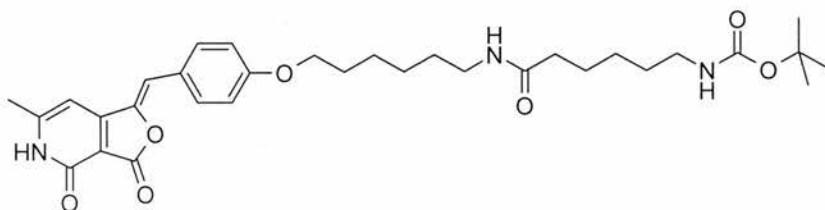
In the dark, *tert*-butyl-(1-{4-(6-amino-*N*-(6-phenoxyhexyl)hexanamide)benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione)carbamate (**202**) (0.14 g, 0.21 mmol) was dissolved in trifluoroacetic acid (10 cm³) and stirred for 1 hour at room temperature. The trifluoroacetic acid was removed by concentrating *in vacuo*, followed by repetition of the addition of toluene (5 x 20 cm³) then concentration *in vacuo*. This yielded the desired product (**200**) (partially as a TFA salt) as a yellow solid (0.117 g, >96%), mp 210-211 °C.

IR (KBr) 2937, 1781 (C(3)=O), 1647 (C(4)=O), 1598 (C=O), 1510, 1260, 1177, 1131, 942, 800, 721, 595 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.14 (1H, brs, NH), 7.77-7.74 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.72 (2H, brs, NH₂), 7.06-7.03 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.90 (1H, s, H8), 6.71 (1H, s, H7), 4.02 (2H, t, *J*= 6.0 Hz, OCH₂), 3.06-3.00 (2H, m, CH₂), 2.81-2.71 (2H, m, CH₂), 2.32 (3H, s, CH₃), 2.05 (2H, t, *J*= 7.0 Hz, CH₂), 1.77 (2H, m, CH₂), 1.56-1.20 (12H, m, CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 171.6 (NHCOCH₂), 164.2 (C4), 159.7 (C4), 157.8 (C4'), 156.2 (C3a), 154.9 (C6), 140.9 (C1), 132.3 (C2' & C5'), 125.1 (C1'), 115.0 (C3' & C5'), 111.7 (C8), 104.9 (C7a), 96.6 (C7), 67.6 (OCH₂), 35.1 (COCH₂), 29.1 (CH₂), 28.5 (CH₂), 26.8 (CH₂), 26.1 (CH₂), 25.5 (CH₂), 25.2 (CH₂), 24.7 (CH₂), 19.5 (CH₃); *m/z* (ES⁺) 482 ([M+H⁺], 100); exact mass calcd for C₂₇H₃₆N₃O₅ 482.2655 (M + H⁺) found 482.2656; Anal. Calcd for C₂₇H₃₅N₃O₅: C, 67.34; H, 7.33; N, 8.73, Anal. Calcd for C₂₇H₃₅N₃O₅·C₂HO₂F₃: C, 58.48; H, 6.09; N, 7.05, (+ 1 mole trifluoroacetic acid), found C, 51.07; H, 6.22; N, 5.56; λ_{max}(DMSO)/nm 372 (18856 ε/dm³ mol⁻¹ cm⁻¹).

6.1.53 Synthesis of 6-[(*tert*-butoxycarbonyl)amino]hexanoic acid (**201**)²¹⁷

6-Aminohexanoic acid (**161**) (5 g, 38 mmol) and sodium hydroxide (1.52 g, 38 mmol) were dissolved in dioxane:water (50:50, 130 cm³). The reaction mixture was cooled to 0 °C and di-*tert*-butyl dicarbonate (7.51 g, 34 mmol) added. The reaction mixture was stirred overnight warming to room temperature. Slight acidification of the reaction mixture was achieved with 10% hydrochloric acid. The solution was then extracted with dichloromethane (3 x 100 cm³), dried with magnesium sulfate and concentrated *in vacuo* to yield the desired compound **201** as a clear oil (7.25 g, 66%). IR (KBr) 2935, 1709 (C=O), 1522, 1410, 1366, 1167, 943, 872, 737, 613 cm⁻¹; ¹H NMR (CD₃OD, 300.1 MHz) 3.07 (2H, t, *J*= 7.0 Hz, OCH₂), 2.33 (2H, t, *J*=7.0 Hz, NCH₂), 1.70-1.61 (2H, m, CH₂), 1.57-1.43 (2H, m, CH₂), 1.47 (9H, s, C(CH₃)₃), 1.43-1.33 (2H, m, CH₂); ¹³C NMR (CD₃OD, 75.5 MHz) 177.9 (CH₂CO₂), 158.9 (NHCO₂), 80.2 (C(CH₃)₃), 68.6 (OCH₂), 41.6 (NHCH₂), 35.3 (CH₂), 31.1 (CH₂), 29.3 (CH₃), 27.8(CH₂), 26.2 (CH₂); *m/z* (ES⁺) 254 ([M+Na⁺], 100%).

6.1.54 Synthesis of *tert*-butyl-(1-{4-(6-amino-*N*-(6-phenoxyhexyl)hexanamide)benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione)carbamate (**202**).

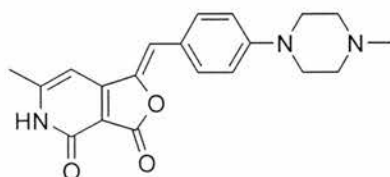


In the dark, 1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**177**) (0.3 g, 0.8 mmol) and *tert*-butyl {6-[(2,5-dioxopyrrolidin-1-yl)oxy]-6-oxohexyl}carbamate (**198**) (0.267 g, 0.8 mmol) were dissolved in dimethylformamide (60 cm³). Triethylamine was added (0.114 cm³, 0.8 mmol) and

the reaction mixture stirred at room temperature over night. The crude reaction mixture was concentrated *in vacuo* and then purified by flash column chromatography eluting with dichloromethane and methanol to yield the desired compound **202** as a yellow solid (0.161 g, 30%), mp 223-224 °C.

IR (KBr) 3302, 2933, 1783 (C(3)=O), 1648 (C(4)=O), 1598, 1512, 1390, 1365, 1258, 1177, 1132, 968, 943, 802, 630, 596 cm^{-1} ; ^1H NMR (DMSO- d_6 , 300.1 MHz) 12.19 (1H, brs, N(5)H), 7.77-7.74 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.82-7.78 (1H, m, NH), 7.06-7.03 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.90 (1H, s, H8), 6.78-6.74 (1H, m, NH), 6.72 (1H, s, H7), 4.02 (2H, t, $J=6.0$ Hz, OCH₂), 3.05-2.99 (2H, m, NHCH₂), 2.90-2.83 (2H, m, NHCH₂), 2.38 (3H, s, CH₃), 2.02 (2H, t, $J=7.0$ Hz, COCH₂), 1.76-1.67 (2H, m, CH₂), 1.27-1.50 (10H, m, CH₂), 1.36 (9H, s, C(CH₃)), 1.23-1.13 (2H, m, CH₂); ^{13}C NMR (DMSO- d_6 , 75.5 MHz) 171.8 (NHCO), 164.2 (C4), 159.7 (C4'), 157.8 (C3), 156.2 (C3a), 155.3 (COC(CH₃)₃), 154.9 (C6), 140.1 (C1), 132.3 (C2' & C6'), 125.2 (C1'), 115.0 (C3' & C5'), 111.7 (C8), 104.9 (C7a), 96.6 (C7), 77.2 (C(CH₃)₃), 67.5 (OCH₂), 38.3 (CH₂), 35.4 (NHCH₂), 29.3 (CH₂), 29.1 (CH₂), 28.5 (CH₂), 28.2 (C(CH₃)₃), 26.1 (CH₂), 26.0 (CH₂), 25.2 (CH₂), 25.1 (CH₂), 19.6 (CH₃); m/z (ES⁺) 604 ([M+Na], 100%); exact mass calcd for C₃₂H₄₃N₃O₇Na 604.2999 (M + Na⁺) found 604.2990; Anal. Calcd for C₁₉H₁₃NO₃: C, 66.07; H, 7.45; N, 7.22, found C, 65.48; H, 7.40; N, 7.09 (differing by 0.6%).

6.1.55 Synthesis of (1Z)-6-methyl-1-[4-(4-methylpiperazin-1-yl)benzylidene]furo[3,4-*c*]pyridine-3,4(1H,5H)-dione (**207**).

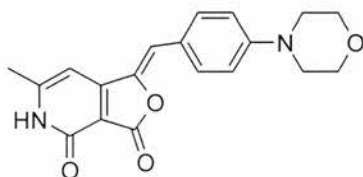


To lactone (**48**) (0.150 g, 0.9 mmol), methanol (15 cm^3) was added and the suspension heated to 60 °C. 4-Morpholinobenzaldehyde (0.18 g, 0.9 mmol) and piperidine (0.09 cm^3 , 0.9 mmol) were added. The reaction mixture was heated at 90 °C overnight. A orange solid suspended in the reaction mixture was filtered and washed with methanol

(3 x 10 cm³), to give the desired compound **207** as an orange solid (0.267 g, 87%), mp >269 °C decomposes.

IR (KBr) 3503, 2813, 1774 (C(3)=O), 1665 (C(4)=O), 1642, 1599, 1519 (C=C), 1454, 1380, 1249, 1226, 1191, 1144, 979, 814, 634, 574 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.01 (1H, brs, NH), 7.69-7.66 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.04-7.01 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.83 (1H, s, H8), 6.68 (1H, s, H7), 3.31-3.27 (2H, m, H2'' & H6''), 2.47-2.45 (2H, m, C3'' & C5''), 2.31 (3H, s, CH₃), 2.22 (3H, s, NCH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.4 (C4), 157.9 (C3), 156.0 (C3a), 154.4 (C6), 151.3 (C4'), 139.8 (C1), 132.1 (C2' & 6'), 122.2 (C1'), 114.4 (C3' & C5'), 112.7 (C8), 104.4 (C7a), 96.4 (C7), 54.3 (C3'' & C5''), 46.7 (C2'' & C6''), 45.7 (NCH₃), 19.5 (CH₃); *m/z* (ES+) 725 ([2M+Na], 40%), 1374 ([M+Na], 100), 352 ([M+H⁺], 19%); exact mass calcd for C₂₀H₂₁N₃O₃Na 374.1478 (M + Na⁺) found 374.1468; λ_{max}(DMSO)/nm 432 (36552 ε/dm³ mol⁻¹ cm⁻¹).

6.1.56 Synthesis of (1*Z*)-6-methyl-1-(4-morpholin-4-ylbenzylidene)furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**208**).

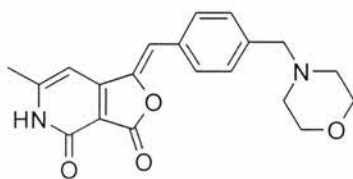


To lactone (**48**) (0.150 g, 0.9 mmol), methanol (15 cm³) was added and the suspension heated to 60 °C. 4-(4-Methylpiperidiazino)benzaldehyde (0.18 g, 0.9 mmol) and piperidine (0.09 cm³, 0.9 mmol) were added. The reaction mixture was heated at 90 °C overnight. A orange solid suspended in the reaction mixture was filtered and washed with methanol (3 x 10 cm³), to give the desired compound **208** as an orange solid (0.282 g, 88%), mp >301 °C decomposes.

IR (KBr) 3036, 2947, 1750 (C(3)=O), 1678 (C(4)=O), 1639, 1594, 1518, 1244, 1225, 1191, 1133, 1110, 976, 921, 830, 806, 620 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.04 (1H, brs, NH), 7.71-7.68 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.06-7.02 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.84 (1H, s, H8), 6.68 (1H, s, H7), 3.77-3.76 (2H, m, H3''), 3.27-3.24 (2H, m, H2''), 2.32 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 165.4 (C4), 157.9 (C3), 156.0 (C3a), 154.4 (C6),

151.4 (C4'), 140.0 (C1), 132.0 (C2' & C6'), 122.7 (C1'), 114.3 (C3' & C5'), 112.5 (C8), 104.5 (C7a), 96.4 (C7), 65.8 (OCH₂), 47.1 (CH₂N), 19.5 (CH₃); *m/z* (ES+) 699 ([2M+Na], 20%), 361 ([M+Na], 100); exact mass calcd for C₁₉H₁₈N₂O₄Na 361.1164 (M + Na⁺) found 361.1171; Anal. Calcd for C₁₉H₁₈N₂O₄: C, 67.45; H, 5.36; N, 8.28, found C, 67.07; H, 5.00; N, 8.16; λ_{\max} (DMSO)/nm 436 (33235 $\epsilon/\text{dm}^3 \text{mol}^{-1} \text{cm}^{-1}$).

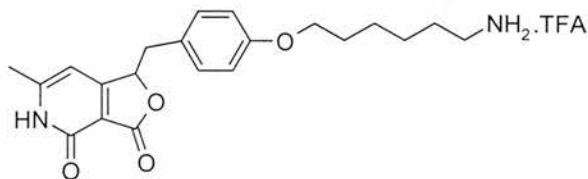
6.1.57 Synthesis of 6-methyl-1-[4-(morpholin-4-ylmethyl)benzylidene]furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**209**).



To lactone (**48**) (0.150 g, 0.9 mmol), methanol (15 cm³) was added and the suspension heated to 60 °C. 4-(Morpholinomethyl)benzaldehyde (0.186 g, 0.9 mmol) and piperidine (0.09 cm³, 0.9 mmol) were added. The reaction mixture was heated at 90 °C overnight. A yellow solid suspended in the reaction mixture was filtered and washed with methanol (3 x 10 cm³), to give the desired compound **209** as a yellow solid (0.145 g, 46%), mp >270 °C decomposes.

IR (KBr) 2911, 2832, 1799 (C(3)=O), 1669 (C(4)=O), 1646, 1613 (C=C), 1567, 1494, 1455, 1395, 1345, 1311, 1214, 1115, 1005, 960, 943, 870, 802, 633, 595, 565 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.20 (1H, brs, NH), 7.78-7.75 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 7.43-7.41 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.93 (1H, s, H8), 6.76 (1H, s, H7), 3.59-3.56 (2H, m, (H3'' & H5'')), 3.49 (2H, s, CH₂), 2.37-2.34 (2H, m, H2'' & H6''), 2.34 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.1 (C4), 157.7 (C3), 156.3 (C3a), 155.4 (C6), 142.4 (C1), 139.5 (C4'), 131.5 (C1'), 130.3 (C3' & C5'), 129.5 (C2' & C6'), 111.3 (C8), 105.4 (C7a), 96.8 (C7), 66.1 (C3'' & C5''), 62.1 (CH₂N), 53.2 (C2'' & C6''), 19.5 (CH₃); *m/z* (ES+) 727 ([2M+Na⁺], 38%), 375 ([M+Na⁺], 100), 353([M+H⁺], 19); exact mass calcd for C₂₀H₂₀N₂O₄Na 375.1321 (M + Na⁺) found 375.1313; Anal. Calcd for C₂₀H₂₀N₂O₄: C, 68.17; H, 5.72; N, 7.95, found C, 67.96; H, 5.37; N, 7.72; λ_{\max} (DMSO)/nm 359 (16741 $\epsilon/\text{dm}^3 \text{mol}^{-1} \text{cm}^{-1}$).

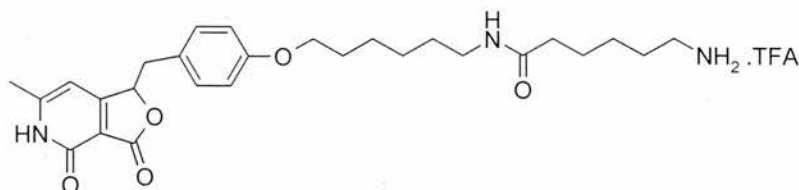
6.1.58 Synthesis of 1-{4-[(5-aminohexyl)oxy]benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**210**).



tert-Butyl-(1-{4-[(5-aminohexyl)oxy]benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**183**) (0.21 g, 0.4 mmol) was dissolved in trifluoroacetic acid (10 cm³) and stirred for 1 hour at room temperature. The trifluoroacetic acid was removed by concentrating *in vacuo*, followed by repetition of the addition of toluene (5 x 20 cm³) then concentration *in vacuo*. This yielded the desired compound **210** (partially as a TFA salt) as a white solid (0.159 g, >95%), mp 209-210 °C decomposes.

IR (KBr) 3063, 2940, 1759 (C(3)=O), 1672 (C(4)=O), 1610, 1513, 1249, 1202, 1137, 832, 798, 721, 640 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.74 (2H, brs, NH₂), 7.09-7.06 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.82-6.79 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.39 (1H, s, H7), 5.56 (1H, dd, *J* 4.0, 6.0, H1), 3.91 (2H, t, *J*=6.0 Hz, OCH₂), 3.08 (2H, ddd, *J*=4.5, 6.5, 14.0, Hz, CHCH₂), 2.78 (2H, t, *J*=7.0 Hz, CH₂NH₂), 2.30 (3H, s, CH₃), 1.72-1.63 (2H, m, CH₂), 1.57-1.49 (2H, m, CH₂), 1.44-1.31 (4H, m CH₂CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 168.4 (C3), 167.3 (C4), 157.8 (C3a), 157.4 (C4'), 154.6 (C6), 130.6 (C2' & C6'), 126.9 (C1'), 114.0 (C3' & C5'), 109.4 (C7a), 99.4 (C7), 78.6 (C1), 67.0 (OCH₂), 37.2 (C8), 28.4 (CH₂), 26.9 (CH₂), 25.4 (CH₂), 25.0 (CH₂), 19.3 (CH₃); *m/z* (ES⁺) 393 ([M+Na⁺], 18%), 371 ([M+H⁺], 100); exact mass calcd for C₂₁H₂₇N₂O₄ 371.1970 (M + H⁺) found 371.1970; Anal. Calcd for C₂₁H₂₆N₂O₄: C, 68.09; H, 7.02; N, 7.56, Anal. Calcd for C₂₁H₂₆N₂O₄.C₂HO₂F₃: C, 57.02; H, 5.62; N, 5.78, (+ 1 mole trifluoroacetic acid), found C, 55.97; H, 5.48; N, 5.45; λ_{max}(DMSO)/nm 327 (7158 ε/dm³ mol⁻¹ cm⁻¹).

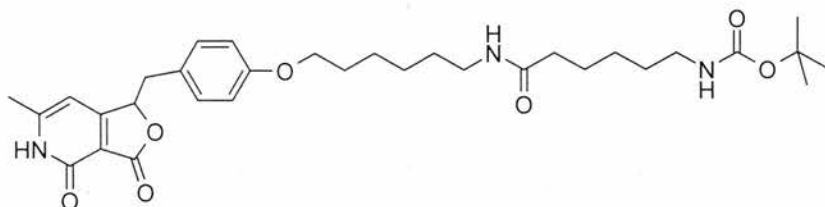
6.1.59 Synthesis of 1-{4-(6-amino-*N*-(6-phenoxyhexyl)hexanamide)benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**211**).



tert-Butyl-(1-{4-(6-amino-*N*-(6-phenoxyhexyl)hexanamide)benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione)carbamate (**212**) (0.1 g, 0.18 mmol) was dissolved in trifluoroacetic acid (10 cm³) and stirred for 1 hour at room temperature. The trifluoroacetic acid was removed by concentrating *in vacuo*, followed by repetition of the addition of toluene (5 x 20 cm³) then concentration *in vacuo*. This yielded the desired compound **211** (partially as a TFA salt) as a white solid (0.08 g, >95%), mp 196-197 °C.

IR (KBr) 2937, 1749 (C(3)=O), 1705 (NHC=O), 1668 (C=O boc), 1633 (C(4)=O), 1606, 1514, 1204, 1144, 1048, 811, 722, 642 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.13 (1H, brs, NH), 7.78-7.74 (1H, m, CH₂NHCO), 7.66 (2H, brs, NH₂), 7.09-7.07 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.82-6.79 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.38 (1H, s, H7), 5.56 (2H, dd, *J*=4.0 7.0 Hz, H1), 3.89 (1H, t, *J*=7.0 Hz, OCH₂), 3.08 (2H, ddd, *J*=4.0 7.0 15.0 Hz, COCH₂), 3.05-2.99 (2H, m, CH₂), 2.79-2.72 (2H, m, CH₂), 2.30 (3H, s, CH₃), 1.21-1.62 (2H, m, CH₂), 1.56-1.20 (14H, m CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 171.6 (NHCO), 168.4 (C4), 167.3 (C3), 157.8 (C3a), 157.4 (C2'), 154.6 (C6), 130.6 (C2' & C6'), 127.0 (C4'), 114.0 (C3' & C5'), 109.4 (C7a), 99.4 (C7), 78.6 (C8), 67.2 (OCH₂), 37.3 (CH₂), 35.1 (CH₂), 29.1 (CH₂), 28.6 (CH₂), 26.8 (CH₂), 26.2 (CH₂), 25.5 (CH₂), 25.2 (CH₂), 24.7 (CH₂), 19.3 (CH₃); *m/z* (ES+) 506 ([M+Na⁺], 42%), 484 ([M+H⁺], 100); exact mass calcd for C₂₇H₃₈N₄O₅ 484.2825 (M + H⁺) found 484.2819; λ_{max}(DMSO)/nm 327 (15418 ε/dm³ mol⁻¹ cm⁻¹).

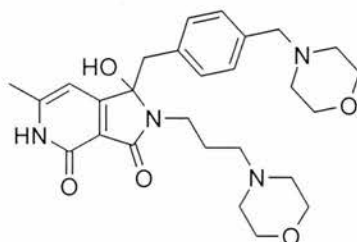
6.1.60 Synthesis of *tert*-butyl-(1-{4-(6-amino-*N*-(6-phenoxyhexyl)hexanamide)benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione)carbamate (**212**).



1-{4-[(5-Amino-hexyl)oxy]benzyl}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**210**) (0.141 g, 0.39 mmol) and the boc protected acid, **201** (0.125 g, 0.39 mmol) were dissolved in dimethylformamide (30 cm³). Triethylamine was added (0.058 cm³, 0.4 mmol) and the reaction mixture stirred at room temperature over night. The crude reaction mixture was concentrated *in vacuo* and then purified by flash column chromatography eluting with dichloromethane and methanol to yield the desired compound **212** as a white solid (0.120 g, 54%), mp 155-156 °C.

IR (KBr) 3318, 2936, 1748 (C(3)=O), 1708 (C=O), 1636 (C(4)=O), 1514, 1230, 1146, 1048, 999, 8098, 648 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 12.30 (1H, brs, NH), 7.73-7.70 (1H, m, NHCOCH₂), 7.10-7.07 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.82-6.79 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.77-6.73 (1H, m, CH₂NHCO), 6.36 (1H, dd, *J*=5.0 7.0 Hz, H1), 3.89 (2H, t, *J*= 6.0 Hz, OCH₂), 3.26-2.88 (2H, ddd, *J*= 4.0 7.0 14.0 Hz, CHCH₂), 3.04-3.00 (2H, m, CH₂), 2.90-2.83 (2H, m, CH₂), 2.29 (3H, s, CH₃), 2.01 (2H, t, *J*= 7.0 Hz, COCH₂), 1.72-1.62 (2H, m CH₂), 1.36 (9H, s, C(CH₃)₃) 1.50-1.15 (12H, m, CH₂); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 171.8 (NHCO), 168.4 (C3), 167.2 (C4), 157.8 (C3a), 157.5 (C4'), 154.5 (C6), 130.5 (C2' & C6'), 127.0 (C1'), 114.0 (C3' & C5'), 109.4 (C7a), 99.4 (C7), 78.6 (C1), 77.2 (C(CH₃)₃), 67.2 (OCH₂), 38.3 (CH₂), 37.3 (C8), 35.4 (CH₂), 29.3 (CH₂), 29.1 (CH₂), 28.6 (CH₂), 28.2 (C(CH₃)₃), 26.2 (CH₂), 26.0 (CH₂), 25.2 (CH₂), 25.0 (CH₂), 19.3 (CH₃); *m/z* (ES-) 582 ([M-H⁺], 100%); exact mass calcd for C₃₂H₄₅N₃O₇Na 606.3155 (M + Na⁺) found 606.3153.

6.1.61 Synthesis of 1-hydroxy-6-methyl-1-[4-(morpholin-4-ylmethyl)benzyl]-2-(3-morpholin-4-ylpropyl)-1*H*-pyrrolo[3,4-*c*]pyridine-3,4(2*H*,5*H*)-dione (**224**)



6-Methyl-1-[4-(morpholin-4-ylmethyl)benzylidene]furo[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione (**209**) (0.07 g, 0.2 mmol) was dissolved in DMSO (1 cm³), to which the buffer solution used for the affinity chromatography experiments (150 mM KCl, 10 mM PIPES, 1.2% octylglucoside) (50 cm³) was added. 3-Morpholino-1-propylamine (**223**) (0.06 cm³, 0.4 mmol) was added to the solution and the mixture stirred overnight at room temperature. The reaction mixture was then concentrated *in vacuo* and then purified by flash column chromatography eluting with dichloromethane and methanol. To yield the desired compound **224** as a pale yellow oil (0.024 g, 24%).

IR (KBr) 3414, 1699 (C(3)=O), 1651 (C(4)=O), 1612, 1573, 1409, 1113, 1025, 1002, 864, 730, 579 cm⁻¹; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.07-7.05 (2H, m AA' part of the AA'BB' system, H2' & H6'), 6.87-6.85 (2H, m BB' part of the AA'BB' system, H3' & H5'), 6.26 (1H, s, H7), 5.76 (1H, s, OH), 3.58-3.51 (10H, m, H3''&H5'', H8), 3.34 (2H, s, CCH₂NCH₂), 3.19-3.18 (2H, m, NCH₂CH₂), 2.36-2.27 (8H, m, H2''&H6''), 2.29-2.24 (4H, m, NCH₂CH₂CH₂NCH₂), 2.25 (3H, s, CH₃); ¹³C NMR (DMSO-*d*₆, 75.5 MHz) 164.6 (C4), 164.2 (C3), 157.6 (C3a), 151.6 (C6), 136.0 (C4'), 133.4 (C1'), 129.5 (C2' & C6'), 128.3 (C3' & C5'), 114.3 (C7a), 99.6 (C7), 89.1 (C1), 66.1 (C3'' & C5''), 62.0 (CCH₂N), 56.0 ((CH₂)₂CH₂NCH₂), 53.1 (C2'' & C6''), 41.1 (NCH₂(CH₂)₂), 36.9 (C8), 25.1 (CH₂CH₂CH₂), 19.1 (CH₃); *m/z* (ES⁺) 1016 ([2M+Na⁺], 45%), 519 ([M+Na⁺], 100), 497 ([M+H⁺], 6).

6.2.63 General procedure for determining solubility

The generation of the extinction coefficients was achieved by dissolving 1 mg of the compound into spectroscopic grade DMSO (100 cm^3), this solution was then diluted 10 fold. A series of 5, 2 fold dilutions of this new solution were then made, giving a total of 7 solutions with concentrations of: 0.01, 0.001, 5×10^{-4} , 2.5×10^{-4} , 1.25×10^{-4} , 6.25×10^{-5} , $3.125 \times 10^{-5} \text{ mg/cm}^3$. These solutions were analysed by UV spectrometry and the absorbance of the solutions at a given wavelength plotted on a graph to give the extinction coefficient as the gradient of the line of best fit.

The actual concentrations of “40 mM” DMSO solutions were determined by taking 1 mg of compound and adding the volume of DMSO required for a theoretical 40 mM solution. Any undissolved compound was then pelleted by centrifugation, the supernatant (0.05 cm^3) was removed and dissolved in DMSO (1 cm^3). The resultant solution was then analysed by UV spectrometry.

The actual concentration of a “100 μM ” aqueous solution was determined by taking 1 mg of compound and adding the volume of DMSO required for a theoretical 40 mM solution. An aliquot (0.042 cm^3) of this “40 mM” solution was added to water (0.1 cm^3), to give a solution at “168 μM ”. An aliquot (0.015 cm^3) of this “168 μM ” solution was added to water (0.01 cm^3) and any undissolved compound was then pelleted by centrifugation. The supernatant (0.02 cm^3) was removed and dissolved in DMSO (1 cm^3). The resultant solution was then analysed by UV spectrometry.

The extinction coefficients generated by this procedure were used to determine the loading levels of the resins. It should be noted however that the extinction coefficients were generated for the TFA salts of these compounds. During the resin loading procedure triethylamine is added to generate the free amine, which may have a slightly different UV absorption. However given the reactive nature of these compounds towards free amines measuring the absorption of the free amines was considered unsuitable.

6.2.64 General Resin loading procedure

In the dark the amine was dissolved in DMSO (1 cm³) and triethylamine (1.5 eq) added. A slurry of the Affi-Gel resin (1 cm³) was washed with DMSO (10 x 1 cm³) and then added to the amine solution and incubated at room temperature over night. The resin was removed from the solution, washed with DMSO (5 x 2 cm³), to yield the desired resin. The DMSO washings were collected and further DMSO added (40 cm³) to give a final volume of 50 cm³. The solution was analysed by UV spectrometry and the quantity of unloaded compound determined. The resin loading levels were determined as: 9.4 (**194R**), 19.9 (**195R**), 29.1 (**196R**), 97.8 (**197R**), 19.2 (**203R**), 14.1 (**205R**), 15.3 (**206R**), 18.6 (**213R**), 13.7 (**214R**), 19.1 (**215R**) and 16.7 (**216R**) %.

6.2.65 NMR experiments to determine a suitable blocking procedure

6.2.65.1 Experiment to mimic the blocking of 100% of the resin

The NHS-ester **196** (0.0012 g, 0.0037 mmol) and *tert*-butyl-(1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**179**) (0.0043 g, 0.00092 mmol) were dissolved in DMSO-*d*₆ (1 cm³) and the solution analysed by ¹H NMR spectrometry. A solution of ethanolamine in DMSO-*d*₆ (0.009 M, 0.5 cm³, 0.0046 mmol) was added, the solution was analysed by ¹H NMR after 5 hours. The ethanolamine had reacted with both **196** to give **240** and **179** to give **241**.

Compound **240**; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.76 (1H, brs, NH), 6.78-6.74 (1H, m, NHCO₂), 3.36 (2H, t, *J*= 6.0 Hz, NHCOCH₂), 3.11-3.05 (2H, m, OCH₂), 2.90-2.84 (2H, m, CH₂), 2.03 (2H, t, *J*= 7.0 Hz, CH₂NHCOCH₂), 1.50-1.40 (2H, m, CH₂), 1.40-1.29 (2H, m, CH₂), 1.36 (9H, s, C(CH₃)₃), 1.29-1.13 (2H, m, CH₂).

Compound **241**; ¹H NMR (DMSO-*d*₆, 300.1 MHz) 7.80-7.73 (2H, m, AA' part of the AA'XX' system, H2' & H6'), 6.67-6.64 (2H, m, XX' part of the AA'XX' system, H3' & H5'), 6.33 (1H, s, H7), 2.73-2.69 (2H, m, CH₂), 2.32-2.30 (2H, m, CH₂), 2.26

(3H, s, CH₃), 1.73-1.64 (2H, m, CH₂), 1.64-1.55 (2H, m, CH₂), 1.36 (9H, s, C(CH₃)₃) signals missing as underneath the signals for the other reagents.

6.2.65.2 Experiment to mimic the blocking of 30% of the resin

An analogous procedure as in 6.2.3.1 was used with NHS-ester **196** (0.0012 g, 0.0037 mmol), *tert*-butyl-(1-{4-[(5-aminohexyl)oxy]benzylidene}-6-methylfuro[3,4-*c*]pyridine-3,4(1*H*,5*H*)-dione) carbamate (**179**) (0.0043 g, 0.00092 mmol) and ethanolamine solution (0.009 M, 0.15 cm³, 0.0014 mmol). The solution was analysed by ¹H NMR. Only the formation of the ester **240** was observed.

6.2.66 General procedure for the NMR competition experiments

Three compounds were dissolved together in DMSO-*d*₆ (1 cm³) so that each compound was at a concentration of 0.019 M. This solution was then analysed by ¹H NMR spectroscopy. To the compound mixture a solution of butylamine in DMSO-*d*₆ (0.19 M, 0.1 cm³, 0.019 mmol) was added and the solution immediately analysed by ¹H NMR. The solution was left at room temperature for 12 hours and then analysed again by ¹H NMR. A further aliquot of the butylamine solution (0.1 cm³, 0.019 mmol) was added and again the solution was immediately analysed by ¹H NMR. After a further 12 hours the solution was analysed by ¹H NMR, this process was repeated once more (see the appendix for the ¹H NMR spectrum produced by this experiment).

6.3 Biological Procedures

6.3.1 Assay against *T. gondii* invasion.¹³⁷

6.3.1.1 Cell Lines and Parasites

African Green Monkey renal epithelial cells (BS-C-1; CCL 26, ATCC, Rockville, MD), cultured as recommended by ATCC, were used as the host cells in all invasion assays. *T. gondii* HXGPRT knockout RH strain stably expressing a tandem repeat of the yellow fluorescent protein (YFP) driven by the *T. gondii* alpha-tubulin promoter (1) were cloned by limiting dilution and cultured in human foreskin fibroblasts (HFFs; CRL 1634, ATCC) as previously described²¹⁹ under 20 mM chloramphenicol selection.

Parasites were harvested either by syringe release of infected cells through a 27 gauge needle, or by collecting parasites from the supernatant following spontaneous lysis of infected monolayers. In either case, the parasites were centrifuged at 1000xg for 4 min, resuspended in Hanks Buffered Salt Solution (HBSS; Invitrogen Life Technologies, Carlsbad, CA) containing 10 mM HEPES, pH 7.0 (HH) and 1% (v/v) dialyzed fetal bovine serum (dFBS, Invitrogen), and filtered through a 3 mm Nuclepore filter (Whatman, Clifton, NJ) prior to use.

6.3.1.2 High-throughput invasion assay

BS-C-1 cells were briefly trypsinized, released into DMEM (Dulbecco's Modified Eagle's Medium; Hyclone, Logan, UT) containing 10% (v/v) FBS and dispensed into 384-well tissue culture-treated Special Optics plates (kindly provided by Corning Costar, Cambridge, MA) at a density of 5×10^3 cells/ 50 cm^3 / well using a Multidrop 384-well dispenser (Thermo Labsystems, Franklin, MA). The plates were incubated overnight at 37 °C in a 5% CO₂ incubator during which time confluent monolayers developed. Immediately prior to the invasion assay, parasites were harvested and filtered as described above and resuspended at $1-1.5 \times 10^7$ tachyzoites/ cm^3 in HH containing 1% (v/v) dFBS. Culture medium was removed from the BS-C-1 cells in the 384-well plate and replaced with HBSS (15 cm^3) containing the

compounds to be tested. Parasites (10 cm^3) were immediately added to each well and incubated with the host cells and the compound for 15 min at 23-25 °C, followed by a 60-90 min incubation at 37 °C. The compounds were assayed at final concentrations of 100 μM to 0.39 μM . The compound containing media was then removed and replaced with HHB (15 cm^3). The infected cells were incubated with 0.5 mg/cm^3 MAb 11-132 (Argene, N. Massapequa, NY) for 15 min at 23-25 °C, then washed twice with HHB (50 cm^3 /well) and incubated for 15 min at 23-25 °C in a mixture of Alexa 488-conjugated goat anti-mouse IgG and Alexa 546-conjugated goat anti-mouse IgG (Molecular Probes, Eugene, OR; each at 2 mg/cm^3 in HHB; Molecular Probes). Following antibody incubation, the cells were washed twice with HH, fixed for 30 min at 23-25 °C in HH containing 3.1% (v/v) formaldehyde and 0.06% (v/v) glutaraldehyde, and washed again with HH. Phosphate-buffered saline (PBS) containing 75% (v/v) glycerol and 0.05% (v/v) sodium azide was added to each well, and plates were stored at 4 °C in the dark until imaged. The plates were imaged on a Nikon TE300 inverted microscope equipped with epifluorescence illumination, using a 20X PlanFluor objective (N.A. 0.5) and FITC and HYQ TR filter set (Chroma Technology, Brattleboro, VT) for Alexa 488 and Alexa 546 fluorescence, respectively. The plates were scored by two individuals independently and the level of invasion described as strong, medium or weak in comparison to the control wells (contain DMSO).

6.3.3 Affinity chromatography experiments using *T. gondii* parasites and HFF cells

6.3.3.1 *T. gondii* lysate

T. gondii parasites were collected by centrifugation at 2500 rpm for 4 min. Then resuspended in PBS (10 cm^3) pelleted again, resuspended in PBS (5 cm^3) and counted. The parasites were pelleted once more and resuspended in lysis buffer (150mM KCl, 10mM PIPES (pH 7.5), 1.2% octylglucoside and 10 μl sigma protease inhibitors) to a density of 2.4×10^8 parasites per cm^3 . The parasites were incubated in the lysis buffer for 1 hour at 4 °C. The lysate was then centrifuged for 10 min at 15,000xg, the supernatant removed and added to the affinity resins. The same method

was used for the NP40 generated extracts, with the 1.2% octylglucoside substituted with 1.2% NP40.

6.3.3.2 HFF cell lystate

HFF cells were collected by centrifugation at 2500 rpm for 4 min, then resuspended in PBS (10 cm³) pelleted again and finally resuspended in PBS (5 cm³). The cells were pelleted once more and resuspended in lysis buffer (150mM KCl, 10mM PIPES (pH 7.5), 1.2% octylglucoside and 10µl sigma protease inhibitors) to a density of 1mg per cm³. The cells were then incubated in lysis buffer for 1 hour at 4 °C. The lysate was then centrifuged for 10 min at 15,000xg, the supernatant removed and added to the affinity resins.

6.3.3.3 General affinity chromatography procedure

Just before use the affinity resin (0.02 cm³) was washed three times with 500 cm³ detergent buffer (150 mM KCl, 10 mM PIPES (pH 7.5), 1.2% octylglucoside and 10 cm³ sigma protease inhibitors) and resuspended in 0.500 cm³ of this buffer. The parasite extract (300 cm³) was then added and the tubes rotated at room temperature for 1 hour. The suspension was centrifuged at 2000 g and the supernatant removed. The resin was then washed five times with buffer. Resin bound proteins were recovered by boiling the resins with SDS buffer (0.5 M Tris, 40 % glycerol; 0.002 % bromophenol blue and 8% SDS) for 5 min. The eluted proteins were separated using standard SDS-PAGE protocols on a 12% acrylamide gel and visualised by conventional silver staining procedures. The above experiments were performed in the dark in order to avoid light induced modification of the compound on the resin.

6.3.3.4 Affinity chromatography procedure for covalent modification

At the University of Vermont the resins were incubated with parasite extract and washed as described in 6.3.3.3. The resin was boiled in SDS buffer 5 times, then washed with an aqueous solution of 50mM Tris, pH 7.4, 0.05% sodium azide. At the University of St Andrews the beads were washed with ammonium carbonate (0.10

cm³, 50 mM). Alcohol dehydrogenase (250 fM) was added, followed by trypsin. The beads were incubated in trypsin over night. The resultant supernatant was removed. In experiment 1 this was analysed by mass spectrometry directly. In experiments 2 and 3 the supernatant was taken up in acid and then analysed by mass spectrometry.

Chapter 7: Conclusions and Future work

7.1 Conclusions

The specific goals this research set out to address were:

- i) Resynthesis of **44** to confirm its biological activity.
- ii) Synthesis of derivatives of **44** to determine a position which could be altered while retaining the biological activity.
- iii) Synthesis of a collection of derivatives of **44** in order to optimise the physical and biological properties.
- iv) Synthesis of a derivative of **44** for loading onto an affinity resin; v) synthesis of the resin based on **44**.
- v) The use of these reagents in affinity chromatography experiments with the ultimate goal of identifying the biological target(s) of **44**.

Taking each of these in turn:

- i) **44**, an enhancer of *T. gondii* invasion, was synthesised and confirmed as the active component (Chapter 2). The geometry of the exocyclic double bond in **44** was shown to be *Z* in some cases using nOe and crystallographic techniques. Characterising the DMSO and aqueous solubility of **44** and its fluorescence properties provided information which directed the development of this chemical tool.
- ii) Derivatives of **44** have been prepared and screened providing structure activity relationships (Chapter 2). Positions which can be modified while retaining the desired biological activity were identified, thus highlighting the positions around which to develop both a small library of analogues and derivatives for affinity chromatography. An inactive analogue of **44** was identified providing an analogue for development as a control for affinity chromatography experiments.
- iii) The parallel synthesis of 22 analogues of **44** provided an insight into both the physical and biological properties of these types of compounds (Chapter 3). The objective of synthesising compounds with improved physical and biological properties was achieved. This work has provided a valuable set of tools with

characterised properties which can be used in various future experiments (including target validation).

iv) Identification of a suitable position for modification directed the synthesis of an analogue of **44** which could be loaded onto an affinity resin, compound **177** (Chapter 4).

v) Resins were loaded with **177** and attempts were made to characterise the resin produced. Preliminary affinity chromatography investigations using this resin directed the development of a second generation of affinity chromatography reagents. This included the synthesis of four resins loaded with active derivatives of **44**, four control resins loaded with an inactive derivative of **44** and three analogues which could be used as competitive controls (Chapter 4).

vi) Investigations into the mechanism of action of **44** have been initiated and as a result affinity chromatography experiments assuming covalent modification of a protein target were instigated (Chapter 5). *T. gondii* inflammatory profilin has been identified as a protein binding partner of **44**.

The aims of the investigation were successfully achieved. A possible cellular target of **44** has been identified, inflammatory profilin.

7.2 Future work

Validation of inflammatory profilin as the cellular target of **44** is the primary objective of any future work. Affinity chromatography experiments which do not assume covalent modification should also be revisited with the second generation affinity chromatography reagents.

7.2.1 Validation of the cellular target

First, investigations will be initiated to confirm that profilin is a true binding partner of **177** and therefore **44**.

i) The affinity chromatography experiments should be repeated using control resins, blank resin and competitive controls to verify the results observed.

ii) Following this, it should be possible to verify this binding interaction between profilin and **44** in vitro using recombinant inflammatory profilin. Incubation of profilin with **44** would be expected to lead to formation of a covalent complex which

can be analysed by mass spectrometry. Covalent modification of inflammatory profilin with **44** should cause an observed mass increase of the protein of 259 Da. Depletion experiments can also be used to confirm the binding partnership, by incubating a solution of inflammatory profilin with both active and control resins and comparing the eluents from each. The amount of protein eluted off the active resin should be depleted in comparison to that of the control. A request has been made to Drs F. Yarovinsky and A. Sher for a sample of recombinant inflammatory profilin with which to do these experiments.

If these experiments confirm profilin as a binding partner *in vitro* validation of the protein as the target of **44** in cells needs to be achieved. This means establishing that binding to and therefore disrupting the profilin function causes the observed phenotype. This can be achieved through a variety of methods including: generation of resistant mutants,²⁶ tetracyclin inducible knockouts,^{220,221} comparison of the SAR observed with the compounds against the isolate protein with their activity in cells.

The library of compounds can be used to support the argument that profilin is the target in cells by incubation of the 22 compounds with isolated inflammatory profilin in order to see if the activity profile observed is consistent with that observed in the whole cell assay. Similarly the derivatives identified as inactive can be confirmed as inactive against the isolated protein. This requires a functional *in vitro* assay to be established.

Combining observations from mutations and compound activity profiling should enable the validation of inflammatory profilin as the cellular target of **44**.

7.2.2 Completing chemical investigations

Initial investigations into the mechanism of action of compounds related to **44** examined the possibility of covalent modification of **44** by a nitrogen based nucleophile. Further investigations should therefore examine the possibility of covalent modification through a sulfur or oxygen based protein nucleophile.

Once the target of **44** has been established a more rational approach to the synthesis of compounds can be taken. This may involve the further development of some of the chemistry introduced in chapter 2.

During the course of this research, reactions which are of potential interest from a synthetic chemistry perspective, such as the [2+2] cycloaddition, have been

observed. While these observations do not feature in the overall goals of this project it would be of interest to examine these further.

References

- 1) Schreiber, S. L. *Bioorg. Med. Chem.*, **1998**, *6*, 1127-1152.
- 2) Ward, G. E.; Carey, K. L.; Westwood, N. J. *Cell. Microbiol.*, **2002**, *4*, 471-482.
- 3) Yeh, J-R. J.; Crews, C. M. *Dev. Cell*, **2003**, *5*, 11-19.
- 4) Shokat, K.; Velleca, M. *Drug Discov. Today*, **2002**, *7*, 872-879.
- 5) Koh, B.; Crews, C. M. *Neuron*, **2002**, *36*, 563-566.
- 6) Westwood, N. J. *Phil. Trans. R. Soc. Lond. A*, **2004**, *362*, 2761-2774. and references therein.
- 7) Crews, C. M.; Splittgerber, U. *Trends Biochem. Sci.*, **1999**, 317-320.
- 8) Landro, J. A.; Taylor, I. C. A.; Stirtan, W. G.; Osterman, D. G.; Kristie, J.; Hunnicutt, E. J.; Rae, P. M. M.; Sweetman, P. M. *J. Pharmacol. Toxicol. Methods*, **2000**, *44*, 273-289.
- 9) Mayer, T. U. *Trends Cell Biol.*, **2003**, *13*, 270-277.
- 10) Lokey, R. S. *Curr. Opin. Chem. Biol.*, **2003**, *7*, 91-96.
- 11) Bray, P. G.; Ward, S. A.; O'Neill, P. M. *Curr. Top. Microbiol. Immunol.*, **2005**, *295*, 3-38.
- 12) McFadden, D. C.; Camps, M.; Boothroyd, J. C. *Drug Resist. Updat.*, **2001**, *4*, 79-84.
- 13) Mackinnon, M. J. *Acta Trop.*, **2005**, *94*, 207-217.
- 14) Mäser, P.; Lüscher, A.; Kaminsky, R. *Drug Resist. Updat.*, **2003**, *6*, 281-290.
- 15) Ouellette, M.; Drummelsmith, J.; Papadopoulou, B. *Drug Resist. Updat.*, **2004**, *7*, 257-266.
- 16) Mayer, T. U.; Kapoor, T. M.; Haggarty, S. J.; King, R. W.; Schreiber, S. L.; Mitchison, T. J. *Science*, **1999**, *286*, 971-974.
- 17) Blangy, A.; Lane, H. A.; d'Herin, P.; Harper, M.; Kress, S M.; Nigg, E. A. *Cell*, **1995**, *83*, 1159-1169.
- 18) Straight, A. F.; Cheung, A.; Limouze, J.; Chen, I.; Westwood, N. J.; Sellers, J. R.; Mitchison, T. J. *Science*, **2003**, *299*, 1743-1747.

-
- 19) Lucas-Lopez, C.; Patterson, S.; Blum, T.; Straight, A. F.; Toth, J.; Slawin, A. M. Z.; Mitchison, T. J.; Sellers, J. R.; Westwood, N. J. *Eur. J. Org. Chem.*, **2005**, 1736-1740.
- 20) Patterson, S.; Lucas-Lopez, C.; Westwood, N. J. *Beilstein-Institut*, **2004**, 147-166.
- 21) Dzierszinski, F.; Mortuaire, M.; Cesbron-Delauw, M. F.; Tomavo, S. *Mol. Microbiol.*, **2000**, *37*, 574-582.
- 22) Blackwell, H. E.; Pérez, L.; Stavenger, R. A.; Tallarico, J. A.; Eatough, E. C.; Foley, M. A.; Schreiber, S. L. *Chem. Biol.*, **2001**, *8*, 1167-1182.
- 23) Walters, W. P.; Ajay; Murcko, M. A. *Curr. Opin. Chem. Biol.*, **1999**, *3*, 384-387.
- 24) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney P. J. *Adv. Drug Deliv. Rev.*, **1997**, *23*, 3-25.
- 25) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.*, **2001**, *46*, 3-26.
- 26) Dobrowolski, J. M.; Sibley, L. D. *Cell*, **1996**, *84*, 933-939.
- 27) Wang, S.; Sim, T. B.; Kim, Y-S.; Chang, Y.-T. *Curr. Opin. Chem. Biol.*, **2004**, *8*, 371-377.
- 28) Bertone, P.; Snyder, M. *FEBS J.*, **2005**, *272*, 5400-5411.
- 29) Merkel, J. S.; Michaud, G. A.; Salcius, M.; Schweitzer, B.; Predki, P. F. *Curr. Opin. Biotech.*, **2005**, *16*, 447-452.
- 30) MacBeath, G. *Nat. Genet. Suppl.*, **2002**, *32*, 526-532.
- 31) Stockwell, B. R. *Nat. Rev. Genet.*, **2000**, *1*, 116-125.
- 32) Kumble, K. D. *Anal. Bioanal. Chem.*, **2003**, *377*, 812-819.
- 33) Hart, C. P. *Drug Discov. Today*, **2005**, *10*, 513-519.
- 34) Chen, J. K.; Lane, W. S.; Schreiber, S. L. *Chem. Biol.*, **1999**, *6*, 221-235.
- 35) Painter, G. F.; Thuring, J. W.; Lim, Z.-Y.; Holmes, A. B.; Hawkins, P. T.; Stephens, L. R. *Chem. Commun.*, **2001**, 645-646.
- 36) Knockaert, M.; Meijer, L. *Biochem. Pharmacol.*, **2002**, *64*, 819-825.
- 37) Knockaert, M.; Gray, N.; Damiens, E.; Chang, Y.-T.; Grellier, P.; Grant, K.; Fergusson, D.; Mottram, J.; Soete, M.; Dubremetz, J.-F.; Le Roch, K.; Doerig, C.; Schultz, P. G.; Meijer, L. *Chem. Biol.*, **2000**, *7*, 411-422.
- 38) Knockaert, M.; Wieking, K.; Schmitt, S.; Leost, M.; Grant, K. M.; Mottram, J. C.; Kunick, C.; Meijer, L. *J. Biol. Chem.*, **2002**, *277*, 25493-25501.
-

-
- 39) Rosania, G. R.; Merlie Jr, J.; Gray, N.; Chang, Y-T.; Schultz, P. G.; Heald, R. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 4797-4802.
- 40) Byrd, C. A.; Bornmann, W.; Erdjument-Bromage, H.; Tempst, P.; Pavletich, N.; Rosen, N.; Nathan C. F.; Ding, A. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 5645-5650.
- 41) Davioud-Charvet, E.; Berecibar, A.; Girault, S.; Landry, V.; Drobecq H.; Sergheraert, C. *Bioorg. Med. Chem. Lett.*, **1999**, *9*, 1567-1572.
- 42) Gavigan, C. S.; Kiely, S. P.; Hirtzlin, J.; Bell, A. *Int. J. Parasitol.*, **2003**, *33*, 987-996.
- 43) Graves, P. R.; Kwiek, J. J.; Fadden, P.; Ray, R.; Hardeman, K.; Coley, A. M.; Foley, M.; Haystead, T. A. *J. Mol. Pharmacol.*, **2002**, *62*, 1364-1372.
- 44) Ambroise, Y.; Mioskowski, C.; Leblanc, G.; Rousseau, B. *Bioorg. Med. Chem. Lett.*, **2000**, *10*, 1125-1127.
- 45) Gurnett, A. M.; Liberator, P. A.; Dulski, P. M.; Salowel, S. P.; Donald, R. G. K.; Anderson, J. W.; Wiltsiel, J.; Diaz, C. A.; Harris, G.; Chang, B.; Darkin-Rattray, S. J.; Nare, B.; Crumley, T.; Blum, P. S.; Misura, A. S.; Tamas, T.; Sardana, M. K.; Yuan, J.; Biftu, T.; Schmatz, D. M. *J. Biol. Chem.*, **2002**, *277*, 15913-15922.
- 46) Fernández-Gacio, A.; Fernández-Marcos, C.; Swamy N.; Ray, R. *Bioorg. Med. Chem. Lett.*, **2003**, *13*, 213-216.
- 47) Vergnon, A. L. ; Chu, Y.-H. *Methods*, **1999**, *19*, 270-277.
- 48) Tanaka, H.; Ohshima, N.; Hidaka, H. *Mol. Pharmacol.*, **1999**, *55*, 356-363.
- 49) Berger, A. B.; Vitorino, P. M.; Bogyo, M. *Am. J. Pharmacogenomics*, **2004**, *4*, 371-381.
- 50) Verhelst, S. H. L.; Bogyo, M. *BioTechniques*, **2005**, *38*, 1-3.
- 51) Dorman, G.; Prestwich, G. D. *Trends Biotechnol.*, **2000**, *18*, 64-77.
- 52) Vezmar, M.; Deady, L. W.; Tilley, L.; Georges, E. *Biochem. Biophys. Res. Comm.*, **1997**, *241*, 104-111.
- 53) Radding, J. A.; Heidler, S. A.; Turner, W. W. *Antimicrob. Agents Chemother.*, **1998**, *42*, 1187-1194.
- 54) Foley, M.; Tilley, L. *Int. J. Parasitol.*, **1997**, *27*, 231-240.
- 55) Foley, M.; Deady, L.; Ng, K.; Cowman, A.; Tilley, L. *J. Biol. Chem.*, **1994**, *269*, 6955-6961.
- 56) Menting, J. G. T.; Tilley, L.; Deady, L. W.; Ngô, K.; Simpson, R. J.; Cowman, A. F.; Foley, M. *Mol. Biochem. Parasitol.*, **1997**, *88*, 215-224.
-

-
- 57) White, M. A. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 10001-10003.
- 58) Baker, K.; Sengupta, D.; Salazar-Jimenez, G.; Cornisa, V. W. *Anal. Biochem.*, **2003**, *315*, 134-137.
- 59) Althoff, E. A.; Cornish, V. W. *Angew. Chem., Int. Ed.*, **2002**, *13*, 2327-2330.
- 60) Becker, F.; Murthi, K.; Smith, C.; Come, J.; Costa-Roldán, N.; Kaufmann, C.; Hanke, U.; Degenhart, C.; Baumann, S.; Wallner, W.; Huber, A.; Dedier, S.; Dill, S.; Kinsman, D.; Hediger, M.; Bockovich, N.; Meier-Ewert, S.; Kluge, A. F.; Kley, N. *Chem. Biol.*, **2004**, *11*, 211-223.
- 61) Lefurgy, S.; Cornish, V. *Chem. Biol.*, **2004**, *11*, 151-153.
- 62) Licitra, E. J.; Liu, J. O. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 12817-12821.
- 63) Barik, S.; Taylor, R. E.; Chakrabarti, D. *J. Biol. Chem.*, **1997**, *272*, 26132-26138.
- 64) Gelb, M. H.; Hol, W. G. J. *Science*, **2002**, *297*, 343-344.
- 65) Modern Parasitology. Parasitic protozoa. F. E. G. Cox. Blackwell Science, 2nd edition, **1993**.
- 66) Sibley, L. D.; Håkansson, S.; Carruthers, V. B. *Curr. Biol.*, **1998**, *8*, R12-R14.
- 67) Dubremetz, J. F.; Garcia-Réguet, N.; Conseil, V.; Fourmaux, M. N. *Int. J. Parasitol.*, **1998**, *28*, 1007-1013.
- 68) Kappe, S. H. I.; Buscaglia, C. A.; Bergman, L. W.; Coppens, I.; Nussenweig, V. *Trends Parasitol.*, **2004**, *20*, 13-16.
- 69) Carruthers, V. B. *Parasitol. Int.*, **1999**, *48*, 1-10.
- 70) Cell Biology of *Toxoplasma gondii*, E. R. Pfefferkorn. Modern Parasite Biology, David J. Wyler editor, W.H. Freeman & Co., **1990**.
- 71) Gleeson, M. T. *Int. J. Parasitol.*, **2000**, *30*, 1053-1070.
- 72) Tentera, A. M.; Heckerotha, A. R.; Weiss, L. M. *Int. J. Parasitol.*, **2000**, *30*, 1217-1258.
- 73) Horvath, D. *J. Med. Chem.*, **1997**, *40*, 2412-2423.
- 74) Keeley, A.; Soldati, D. *Trends Cell Biol.*, **2004**, *14*, 528-532.
- 75) Carruthers, V. B. *Acta Trop.*, **2002**, *81*, 111-222.
- 76) Kim, K. *Acta Trop.*, **2004**, *91*, 69-81.
- 77) Dowse, T.; Soldati, D. *Curr. Opin. Microbiol.*, **2004**, *7*, 388-395.
- 78) Bonhomme, A.; Bouchot, A.; Pezzella, N.; Gomez, J.; Le Moal, H.; Pinon, J. M. *FEMS Microbiol. Rev.*, **1999**, *23*, 551-561.
- 79) Soldati, D.; Dubremetz, J. F.; Lebrum, M. *Int. J. Parasitol.*, **2001**, *31*, 1293-1302.
-

-
- 80) Opitz, C.; Soldati, D. *Mol. Microbiol.*, **2002**, *45*, 597-604.
- 81) Sibley, L. D. *Science*, **2004**, *304*, 248-253.
- 82) Soldati, D.; Foth, B. J.; Cowman, A. F. *Trends Parasitol.*, **2004**, *20*, 567-574.
- 83) Morrissette; N. S.; Sibley, L. D. *Microbiol. Mol. Biol. Rev.*, **2002**, *66*, 21-38.
- 84) Soldati, D.; Dubremetz, J. F.; Lebrun, M.; *Int. J. Parasitol.*, **2001**, *31*, 1293-1302.
- 85) Carruthers, V.B. *Parasitol. Int.*, **1999**, *48*, 1-10.
- 86) Suss-Toby, E.; Zimmerberg, J.; Ward, G. E. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 8413-8418.
- 87) Meissner, M.; Schlüter, D.; Soldati, D. *Science*, **2002**, *298*, 837-840.
- 88) Herm-Götz, A.; Weiss, S.; Stratmann, R.; Ruff, S. F.; Meyhöfer, E.; Doldati, T.; Manstein, D. J.; Soldati, D. *EMBO J.*, **2002**, *21*, 2149-2158.
- 89) Schmitz, S.; Grainger, M.; Howell, S.; Calder, L. J.; Gaeb, M.; Pinder, J. C.; Holder, A. A.; Veigel, C. *J. Mol. Biol.*, **2005**, *349*, 113-125.
- 90) Gaskins, E.; Gilk, S.; Devore, N.; Mann, T.; Ward, G. Beckers, C. *J. Cell Biol.*, **2004**, *165*, 383-393.
- 91) Patrón, A.; Mondragón, M.; González, S.; Ambrosio, J. R.; Guerrero, A. L.; Mondragón, R. *Int. J. Parasitol.*, **2005**, 1-12.
- 92) Dobrowolski, J. M.; Niesman, I. R.; Sibley, L. D. *Cell. Motil. Cytoskeleton*, **1997**, *37*, 253-262.
- 93) Schmitz, S.; Grainger, M.; Howell, S.; Calder, L. J.; Gaeb, M.; Pinder, J. C.; Holder, A. A.; Veigel, C. *J. Mol. Biol.*, **2005**, *349*, 113-125.
- 94) Poupel, O.; Tardieux, I. *Microbes Infect.*, **1999**, *1*, 653-662.
- 95) Shaw, M. K.; Tilney, L. G. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 9095-9099.
- 96) Fowler, R. E.; Margos, G.; Mitchell, G. H. *Adv. Parasitol.*, **2004**, *56*, 213-263.
- 97) Jewett, T. J.; Sibley, L. D. *Mol. Cell.*, **2003**, *11*, 885-894.
- 98) Buscaglia, C. A.; Coppens, I.; Hol, W. G.; Nussenzweig, V. *Mol. Biol. Cell*, **2003**, *14*, 4947-4957.
- 99) Holt, M. R.; Koffer, A. *Trends Cell Biol.*, **2001**, *11*, 38-46.
- 100) Yarovinsky, K.; Zhang, D.; Andersen, J. F.; Bannenberg, G. L.; Serhan, C. N.; Hayden, M. S.; Hieny, S.; Sutterwala, F. S.; Flavell, R. A.; Ghosh, S.; Sher, A. *Science*, **2005**, *308*, 1626-1629.
- 101) Carruthers, V. B.; Blackman, M. J. *Mol. Microbiol.*, **2005**, *55*, 1617-1630.
-

-
- 102) Lekutis, C.; Ferguson, D. J. P.; Grigg, M. E.; Camps M.; Boothroyd, J. C. *Int. J. Parasitol.*, **2001**, *31*, 1285-1292.
- 103) Windeck, T.; Gross, U. *Parasitol. Res.*, **1996**, *82*, 715-719.
- 104) Dzierszinski, F.; Mortuaire, M.; Cesbron-Delauw, M. F.; Tomavo, R. *Mol. Microbiol.*, **2000**, *37*, 574-582.
- 105) Mineo, J. R.; McLeod, R.; Mack, D.; Smith, E.; Khan, I. A.; Ely, K. H.; Kasper, L. H. *J. Immunol.*, **1993**, *150*, 3951-3964.
- 106) Hu, K.; Roos, D. S.; Murray, J. M. *J. Cell Biol.*, **2002**, *156*, 1039-1050.
- 107) Monteiro, V. G.; de Melo, E. J. T.; Attias, M.; de Souza, W. *J. Struct. Biol.*, **2001**, *136*, 181-189.
- 108) Carruthers, V. B.; Giddings, O. K.; Sibley, L. D. *Cell. Microbiol.*, **1999**, *1*, 225-235.
- 109) Carruthers, V. C.; Sibley L. D. *Mol. Microbiol.*, **1999**, *31*, 421-428.
- 110) Lovett, J. L.; Marchesini, N.; Moreno, S. N. J.; Sibley, L. D. *J. Biol. Chem.*, **2002**, *277*, 25870-25876.
- 111) Wetzell, D. M.; Chen, L. A.; Ruiz, F. A.; Moreno S. N. J.; Sibley, L. D. *J. Cell. Sci.*, **2004**, *117*, 5739-5748.
- 112) Huynh, M.-H.; Rabenau, K. E.; Harper, J. M.; Beatty, W. L.; Sibley, L. D.; Carruthers, V. B. *EMBO J.*, **2003**, *22*, 2082-2090.
- 113) Kim, K. *Acta Trop.*, **2004**, *91*, 69-81.
- 114) Bradley, P. J.; Ward, C.; Cheng, S. J.; Alexander, D. L.; Collier, S.; Coombs, G. H.; Dunn, J. D.; Ferguson, D. J.; Sanderson, S. J.; Wastling, J. M.; Boothroyd, J. C. *J. Biol. Chem.*, **2005**, *280*, 34245-34258.
- 115) Sinai, A. P.; Webster, P.; Joiner, K. *J. Cell Sci.*, **1997**, *110*, 2117-2128.
- 116) Sinai A. P.; Joiner, K. A. *J. Cell Biol.*, **2001**, *154*, 95-108.
- 117) Mercier, C.; Adjogble, K. D. Z.; Däubener, W.; Delauw, M.-F.-C. *Int. J. Parasitol.*, **2005**, *35*, 829-849, and references therein.
- 118) Dubremetz J. F.; Achbarou A.; Bermudes D.; Joiner K. A. *Parasitol. Res.*, **1993**, *79*, 402-408.
- 119) Carruthers V. B.; Sibley L. D. *Eur. J. Cell Biol.*, **1997**, *73*, 114-123.
- 120) Sibley, L. D., Niesman, I. R.; Parmley, S. F.; Cesbron-Delauw, M.F. *J. Cell Sci.*, **1995**, *108*, 1669-1677.
- 121) Ngô, H. M.; Yang, M.; Joiner, K. A. *Mol. Microbiol.*, **2004**, *52*, 1531-1541.
-

-
- 122) Carruthers, V. B. *Parasitol. Int.*, **1999**, *48*, 1-10.
- 123) Tomley, F. M.; Soldati, D. S. *Trends Parasitol.*, **2001**, *17*, 81-88.
- 124) Carruthers, V. B.; Sherman, G. D.; Sibley, L. D. *J. Biol. Chem.*, **2000**, *275*, 14346-14353.
- 125) Dowse, T.; Soldati, D. *Curr. Opin. Microbiol.*, **2004**, *7*, 388-396.
- 126) Brossier, F.; Jewett, T.J.; Lovett, J.L.; Sibley, L.D. *J. Biol. Chem.*, **2003**, *278*, 6229-6234.
- 127) Sadak, A.; Taghy, Z.; Fortier, B.; Dubremetz, J. F. *Mol. Biochem. Parasitol.*, **1988**, *29*, 203-211.
- 128) Binder E. M.; Kim. K. *Acta Trop.*, **2004**, *91*, 69-81.
- 129) Withers-Martinez, C.; Jean, L.; Blackman, M. J. *Mol. Microbiol.*, **2004**, *53*, 55-63.
- 130) Blackman, M. J. *Curr. Drug Targets*, **2000**, *1*, 59-83.
- 131) Blackman, M. J. *Cell. Microbiol.*, **2004**, *6*, 893-903.
- 132) Que, X.; Ngô, H.; Lawton, J.; Gray, M.; Liu, Q.; Engel, J.; Brinen, L.; Ghosh, P.; Joiner, K. A.; Reed. S. L. *J. Biol. Chem.*, **2002**, *277*, 25791-25797.
- 133) Dowse, T. J.; Soldati, D. *Trends. Parasitol.*, **2005**, *21*, 254-258.
- 134) Brossier, F.; Jewett, T. J.; Sibley, L. D.; Urban, S. *Proc. Natl. Acad. Sci. USA*, **2005**, *102*, 4146-4151.
- 135) Dowse, T. J.; Pascall, J. C.; Brown, K. D.; Soldati, D. *Int. J. Parasitol.*, **2005**, *35*, 747-756.
- 136) Morgan, R. E.; Evans, K. M.; Patterson, S.; Catti, F.; Ward, G. E.; Westwood, N. *J. Curr. Drug Targets*, **2006**, *in press*.
- 137) Carey, K. L.; Westwood, N. J.; Mitchison T. J.; Ward, G. E. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 7433-7438.
- 138) Asai, T.; Takeuchi, T.; Diffenderfer J.; Sibley, L. D. *Antimicrob. Agents Chemother.*, **2002**, *46*, 2393-2399.
- 139) Baldwin, J.; Michnoff, C. H.; Malmquist, N. A.; White, J.; Roth, M. G.; Rathod, P. K.; Phillips, M. *J. Biol. Chem.*, **2005**, *280*, 21847-21853.
- 140) Terstappen, G. C.; Reggiani, A. *Trends Pharmacol. Sci.*, **2001**, *22*, 23-26.
- 141) Chowdhury, S. F.; Lucrezia, R. D.; Guerrero, R. H.; Brun, R.; Goodman, J.; Ruiz-Perez, L. M.; Pacanoska, D. G.; Gilbert, I. H. *Bioorg. Med. Chem. Lett.*, **2001**, *11*, 977-980.
-

-
- 142) Selzer, P. M.; Chen, X.; Chan, V. J.; Cheng, M.; Kenyon, G. L.; Kuntz, I. D.; Sakanari, J. A.; Cohen, F. E.; McKerrow, J. H. *Exp. Parasitol.*, **1997**, *87*, 212-221.
- 143) Ring, C. S.; Sun, E.; McKerrow, J. H.; Lee, G. K.; Rosenthal, P. J.; Kuntz, I. D.; Cohen, F. E. *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 3583-3587.
- 144) Rastelli, G.; Pacchioni, S.; Strawaraporn, W.; Sirawaraporn, R.; Parenti, M. D.; Ferrari, A. M. *J. Med. Chem.*, **2003**, *46*, 2834-2845.
- 145) Toyoda, T.; Brobey, R. K. B.; Sano, G.; Horii, T.; Tomioka, N.; Itai, A. *Biochem. Biophys. Res. Commun.*, **1997**, *235*, 515-519.
- 146) Aronov, A. M.; Munagala, N. R.; Ortiz de Montellano, P. R.; Kuntz, I. D.; Wang, C. C. *Biochemistry*, **2000**, *39*, 4684-4691.
- 147) Zuccotto, F.; Zvelebil, M.; Brun, R.; Chowdhury, S. F.; Di Lucrezia, R.; Leal, I. Maes, L.; Ruiz-Perez, L. M.; Pacanowska, D. G.; Gilbert, I. H. *Eur. J. Med. Chem.*, **2001**, *36*, 395-405.
- 148) Freymann, D. M.; Wenck, M. A.; Engel, J. C.; Feng, J.; Focia, P. J.; Eakin, A. E. Craig III, S. P. *Chem. Biol.*, **2000**, *7*, 957-968.
- 149) Horvath, D. *J. Med. Chem.*, **1997**, *40*, 2412-2423.
- 150) Brydges, S. D.; Sherman, G. D.; Nockemann, S.; Loyens, A.; Däubener, W.; Dubremetz, J.-F.; Carruthers, V. B. *Mol. Biochem. Parasitol.*, **2000**, *111*, 51-66.
- 151) Arustamova, I. S.; Kaigorodova, E. A.; Kul'nevich, V. G.; Ugrak, B. I. *Chem. Heterocycl. Comp'd*, **1993**, *29*, 539-544.
- 152) Bruce, W. P.; Coover, H. W. Jr. *J. Am. Chem. Soc.*, **1944**, *66*, 2092-2094.
- 153) Harris, S. A.; Stiller, T. E.; Folkers, K. *J. Am. Chem. Soc.*, **1939**, *61*, 1242-1244.
- 154) Arustamova, S.; Kaigoradova, E. A.; Kul'nevich, V. G.; Ugrak, B. I. *Chem. Heterocycl. Compd (Engl transl)*, **1993**, *29*, 539-544.
- 155) Lycke, E.; Carlberg, K.; Norrby, R. *Infect. Immun.*, **1975**, 853-861.
- 156) Norrby, R.; Lycke, E. *J. Bacteriol.*, **1967**, *93*, 53-58.
- 157) Grimwood, J.; Smith, J. E. *Int. J. Parasitol.*, **1996**, *26*, 169-173.
- 158) Wenner W.; Plati J. T. *Tetrahedron*, **1946**, 751-759.
- 159) Bardhan, J. C. *J. Chem. Soc.*, **1929**, *2*, 2223-2232.
- 160) Advanced Organic Chemistry. Reactions, Mechanisms and Structure. M. B. Smith & J. March. Wiley-Inter-Science, 5th edition, **2001**.
- 161) Parchment, O. G.; Burton, N. A.; Hiller I. H.; Vincent, M. A. *J. Chem. Soc., Perkin Trans. 2*, **1993**, 861-863.
-

-
- 162) Tsuchida, N.; Yamabe, S. *J. Phys. Chem. A*, **2005**, *109*, 1974-1980.
- 163) Fu, A.; Li, H.; Du, D.; Zhu, Z. *J. Phys. Chem. A*, **2005**, *109*, 1468-1477.
- 164) Kohmoto, S.; Noguchi, T.; Masu, H.; Kishikawa, K.; Yamamoto, M.; Yamaguchi, K. *Org. Lett.*, **2004**, *6*, 683-685.
- 165) Conreux, D.; Bossharth, E.; Monterio, N.; Desbordes, P.; Balme, G. *Tetrahedron. Lett.*, **2005**, *46*, 7917-7920.
- 166) Liu, H.; Ko, S.-B.; Josin, H.; Curran, D. P. *Tetrahedron Lett.*, **1995**, *36*, 8917-8920.
- 167) Kul'nevich, V. G.; Kaigorodova, E. A.; Arustamova, I. S.; Korobchenko, L. V.; Vladyko, G. V.; Boreko, E. I. *Pharm. Chem. J. (Engl. Transl.)*, **1990**, *24*, 132-135.
- 168) Ciufolini M. A.; Browne, M. E. *Tetrahedron Lett.*, **1987**, *28*, 171-174.
- 169) Barret, A. G. M.; Lebold, S. A.; Zhang, X. *Tetrahedron Lett.*, **1989**, *30*, 7317-7320.
- 170) Tsunoda, T.; Yamamiya, Y.; Itô, S.; *Tetrahedron Lett.*, **1993**, *34*, 1639-1642.
- 171) Comins, D. L.; Jianhua, G. *Tetrahedron Lett.*, **1994**, *35*, 2819-2822.
- 172) Krchňák, V. *J. Comb. Chem.*, **2005**, *7*, 507-509.
- 173) Allan, R. D.; Johnston, G. A. R.; Kazlauskas, R.; Tran, H. W. *J. Chem. Soc., Perkin Trans. 1*, **1983**, *12*, 2983-2986.
- 174) Fu, H.; Park, J.; Pei, D.; *Biochemistry*, **2002**, *41*, 10700-10709.
- 175) Harris, S. A.; Wilson, A. N. *J. Am. Chem. Soc.*, **1941**, *63*, 2526-2527.
- 176) Blasko, G.; Varadaras, E.; Sener, B.; Freyer, A.; Asma, M. *J. Org. Chem.*, **1982**, *47*, 880-885.
- 177) Koelsch, C. F.; Prill, E. J. *J. Am. Chem. Soc.*, **1945**, *67*, 1296-1299.
- 178) Hird, N. W. *Drug Discov. Today*, **1999**, *4*, 265 – 274.
- 179) Kappe, C. O. *Angew. Chem. Int. Ed.*, **2004**, *43*, 6250-6284.
- 180) Kappe, C. O.; Larhed, M. *Angew. Chem. Int. Ed.*, **2005**, *44*, 7666-7669.
- 181) Kenseth; J. R.; Coldiron, S. J. *Curr. Opin. Chem. Biol.*, **2004**, *8*, 418-423.
- 182) Westwood, N. J.; Smith, G. F., *manuscript submitted*.
- 183) High-Resolution NMR Techniques in Organic Chemistry. T. D. W. Claridge. Tetrahedron Organic Chemistry Series, Elsevier, **1999**.
- 184) Spectroscopic Methods in Organic Chemistry. D. H. Williams, I. Fleming. McGraw-Hill Publishing Co., 5th edition, **1995**.
-

-
- 185) Schiedel, M.-S.; Briehn, C. A.; Bäuerle, P. *Angew. Chem. Int. Ed.*, **2001**, *40*, 4677-4680.
- 186) Zhu, Q.; Yoon, H.-S.; Parikh, P. B.; Chang, Y.-T.; Yao, S. Q. *Tetrahedron Lett.*, **2002**, *43*, 5083-5086.
- 187) Rosania, G. R.; Lee, J. W.; Ding, L.; Yoon, H.-S.; Chang, Y.-T.; *J. Am. Chem. Soc.*, **2003**, *125*, 1130-1131.
- 188) Theory and interpretation of fluorescence and phosphorescence, R. S. Becker Wiley Intersciences, 1969.
- 189) Fluorescence and Phosphorescence Spectroscopy: Physicochemical Principles and Practice. S. G. Schulman. Pergamon Press Ltd., **1979**.
- 190) Gurnett, A. M.; Liberator, P. A.; Dulski, P. M.; Salowel, S. P.; Donald, R. G. K.; Anderson, J. W.; Wiltsiel, J.; Diaz, C. A.; Harris, G.; Chang, B.; Darkin-Rattray, S. J.; Nare, B.; Crumley, T.; Blum, P. S.; Misura, A. S.; Tamas, T.; Sardana, M. K.; Yuan, J.; Biftu, T.; Schmatz, D. M. *J. Biol. Chem.*, **2002**, *277*, 15913-15922.
- 191) Donald, R. G. K.; Allocco, J.; Singh, S. B.; Nare, B.; Salowe, S. P.; Wiltsie, J.; Liberator, P. A. *Eukaryotic Cell*, **2002**, *1*, 317-328.
- 192) Wiesma, H. I.; Galuska, S. E.; Tomley, F. M.; Sibley, L. D.; Liberator, P. A.; Donald, R. G. K. *Int. J. Parasitol.*, **2004**, *34*, 369-380.
- 193) Foley, M.; Tilley, L. *Int. J. Parasitol.*, **1997**, *27*, 231-240.
- 194) Desneves, J.; Thorn, G.; Berman, A.; Galatis, D.; La Greca, N.; Sinding, J.; Foley, M.; Deady, L. W.; Cowman, A. F.; Tilley, L. *Mol. Biochem. Parasitol.*, **1996**, *82*, 181-194.
- 195) Fujii, N.; Haresco, J. J.; Novak, K. A. P.; Stokoe, D.; Kuntz, I. D.; Guy, R. K. *J. Am. Chem. Soc.*, **2003**, *125*, 12074-12075.
- 196) Videlock, E. J.; Chung, V. K.; Mohan, M. A.; Strok, T. M.; Austin, D. J. *J. Am. Chem. Soc.*, **2004**, *126*, 3730-3731.
- 197) Greenbaum, D. C.; Baruch, A.; Grainger, M.; Bozdech, Z.; Medzihradzsky, K. F.; Engel, J.; DeRisi, J.; Holder, A. A.; Bogoy, M. *Science*, **2002**, *298*, 2002-2006.
- 198) E. V. Blackburn, C. J. Timmons, *J. Chem. Soc.*, 1970, 172-175.
- 199) Immobilised Affinity Ligand Techniques. G. T. Hermanson, A. K. Mallia, P. K. Smith. Academic Press (New York), **1992**.
- 200) Painter, G. F.; Thuring, J. W.; Lim, Z.-Y.; Holmes, A. B.; Hawkins, P. T.; Stephens, L. R. *Chem. Commun.*, **2001**, 645-646.
-

-
- 201) Kanoh, N.; Honda, K.; Simizu, S.; Muroi, M.; Osada, H. *Angew. Chem. Int. Ed.*, **2005**, *44*, 3559-3562.
- 202) Paiva, P. M. G.; Souza, A. F.; Oliva, M. L. V.; Kennedy, J. F.; Cavalcanti, M. S. M.; Coelho, L. C. B. B.; Sampaio, C. A. M. *Bioresour. Technol.*, **2003**, *88*, 75-79.
- 203) Gray, N. S. *Curr. Opin. Neurobiol.*, **2001**, *11*, 608-614.
- 204) Marquet, A.; Frère, J.-M.; Ghuysen, J.-M.; Loffet, A. *Biochem. J.*, **1979**, *177*, 909-916.
- 205) Zhou, S. *J. Chromatogr. B*, **2003**, *797*, 63-90 and references therein.
- 206) Steen, H.; Mann, M. *Nat. Rev. Mol. Cell Biol.*, **2004**, *5*, 699-711.
- 207) Khan, A.; Taylor, S.; Su, C.; Mackey, A. J.; Boyle, J.; Cole, R.; Glover, D.; Tang, K.; Paulsen, I. T.; Berriman, M.; Boothroyd, J. C.; Pfefferkorn, E. R.; Dubley, J. P.; Ajioka, J. W.; Roos, D. S.; Wootton, J. C.; Sibley, L. D. *Nucleic Acids Res.*, **2005**, *33*, 2980-2992.
- 208) Daniels, S. B.; Katzenellenbogen, J. A. *Biochemistry*, **1986**, *25*, 1436-1444.
- 209) Chakravarty, P. K.; Krafft, G. A.; Katzenellenbogen, J. A. *J. Biol. Chem.*, **1982**, *257*, 610-612.
- 210) Talalay, P.; De Long, M. J.; Prochaska, H. J. *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 8261-8265.
- 211) Cusack, K. P.; Arnold, L. D.; Barberis, C. E.; Chen, H.; Ericsson, A. M.; Gazabulsec, G. S.; Gordon, T. D.; Grinnell, C. M.; Harsch, A.; Pellegrini, M.; Tarcsa, E. *Bioorg. Med. Chem. Lett.*, **2004**, *14*, 5503-5507.
- 212) Beck, J. J.; Stremix, F. R. *J. Nat. Prod.*, **1995**, *58*, 1047-1055.
- 213) Castelli, V. A.; Bernardi, F.; Cort, A. D.; Mandolini, L.; Rossi, I.; Schiaffino, L. *J. Org. Chem.*, **1999**, *64*, 8122-8126.
- 214) Organic Chemistry. J. Clayden, N. Greeves, S. Warren, P. Wothers. Oxford University Press, **2001**.
- 215) Forero, C.; Wasserman, M. *Mem. Inst. Oswaldo Cruz*, **2000**, *95*, 329-337.
- 216) Witke, W. *Trends Cell Biol.*, **2004**, *14*, 461-469.
- 217) Nodelman, I. M.; Bowman, G. D.; Lindberg, U.; Schutt, C. E. *J. Mol. Biol.*, **1999**, *294*, 1271-1285.
- 218) Isomura, S.; Wirsching, P.; Janda, K. D. *J. Org. Chem.*, **2001**, *66*, 4115-4121.
- 219) Doughty, M. B.; Chaurasia, C. S.; Li, K. *J. Med. Chem.*, **1993**, *36*, 272-279.
- 220) Ward, G. E.; Carey, K. L.; Westwood, N. J. *Cell. Microbiol.*, **2002**, *4*, 471-482.
-

221) Kitamura, M. *Exp. Nephrol.*, **1998**, *6*, 576-580.

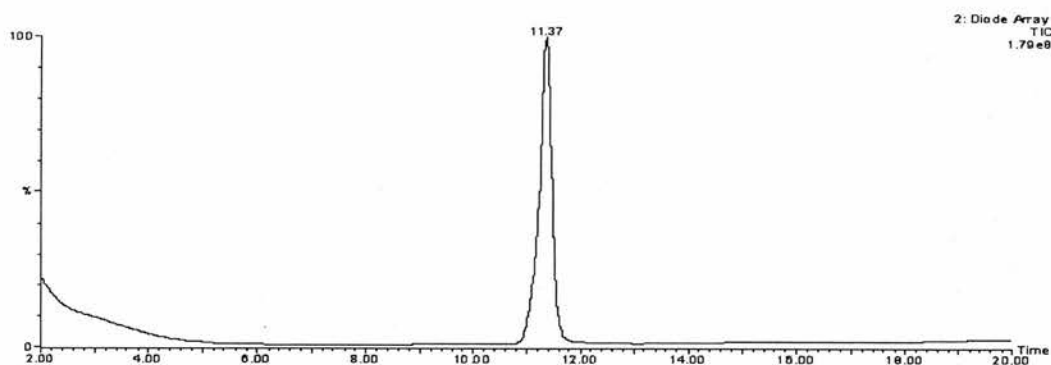
222) Messner, M.; Brecht, S.; Bujard, H.; Soldati, D. *Nucleic Acids Res.*, **2001**, *29*, E115.

Appendix 1: LCMS data

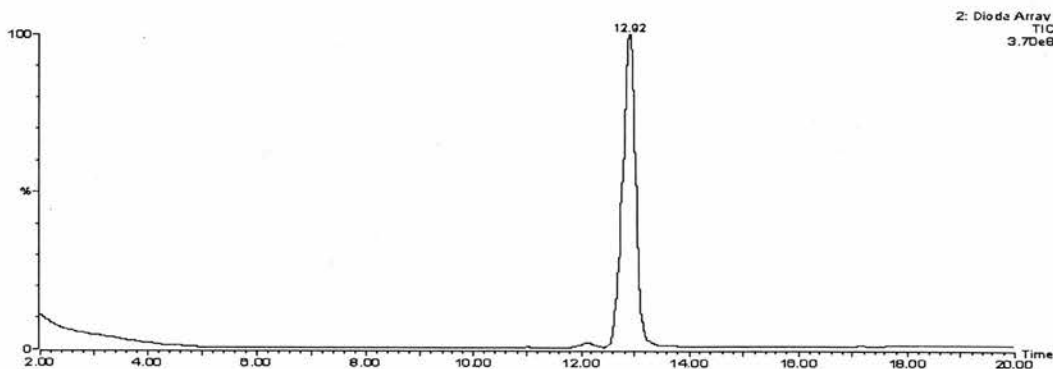
The below are the chromatograms produced by a diode array attached to an LCMS system. The chromatograms are arranged by compound number.

The separation was carried out using a Waters XTerra, RP₁₈ 5 μ M, 30 x 50 mm column with a flow rate of 0.5 ml/min and a 20 min gradient of 05-95% CH₃OH in H₂O, using a Waters 996 photodiode array detector.

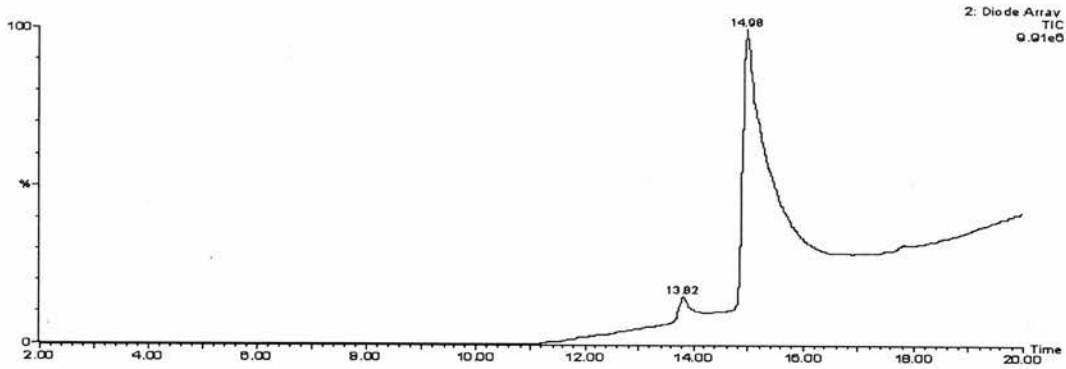
Compound 44



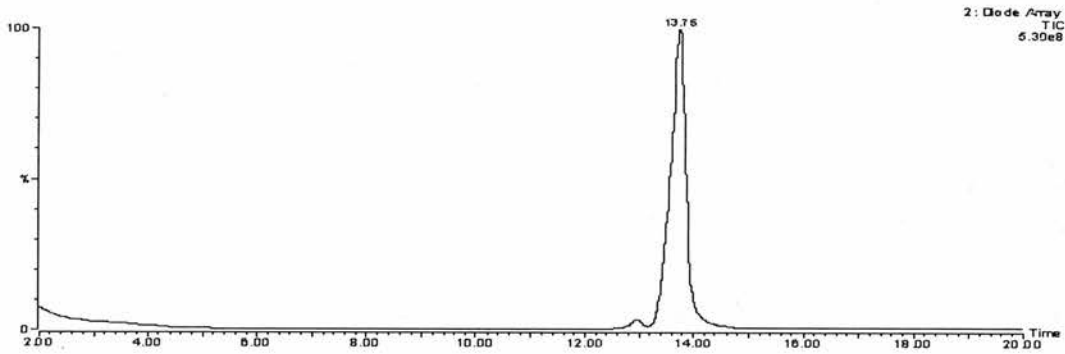
Compound 59



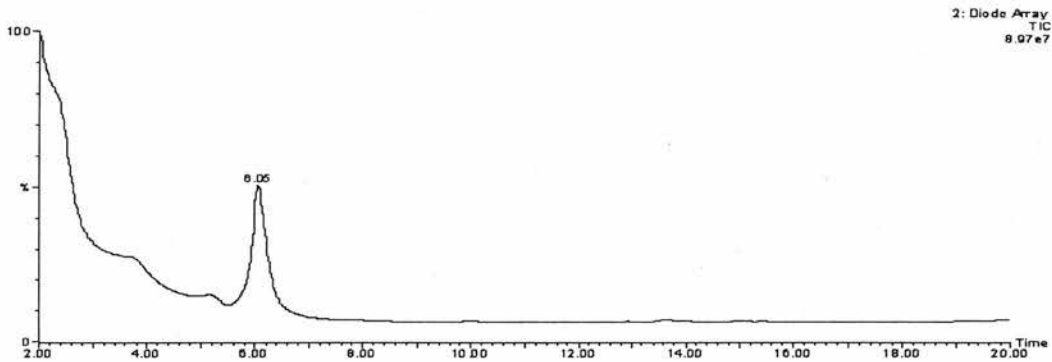
Compound 120



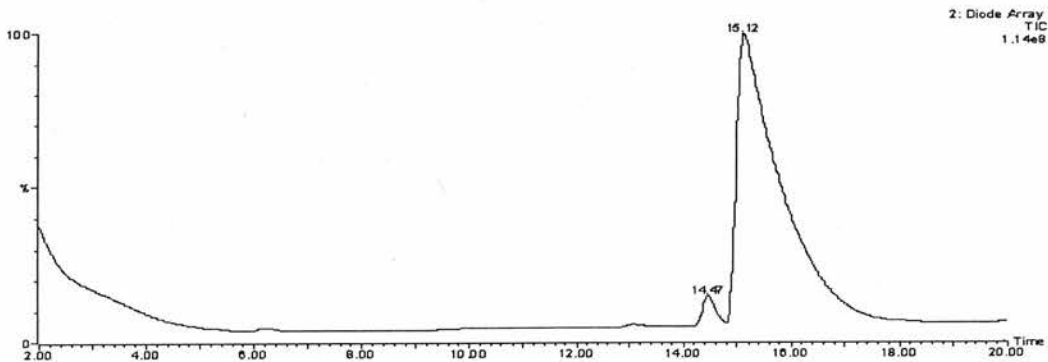
Compound 122



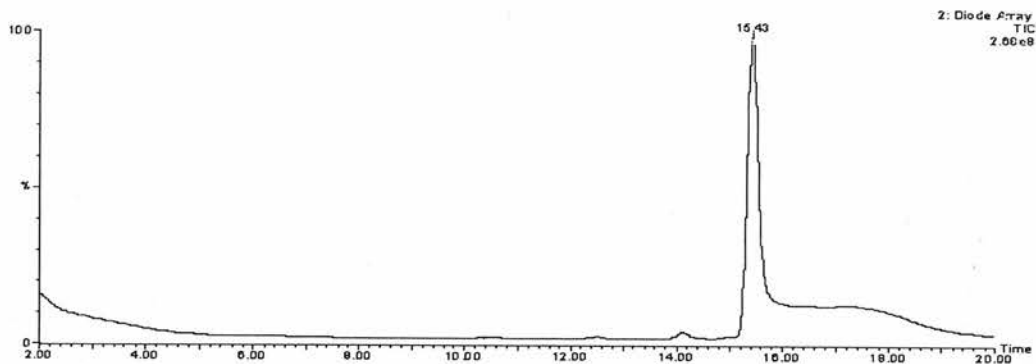
Compound 123



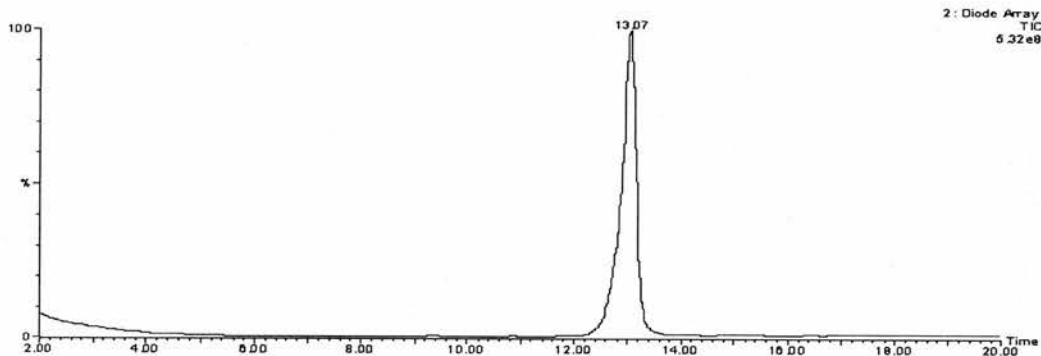
Compound 124



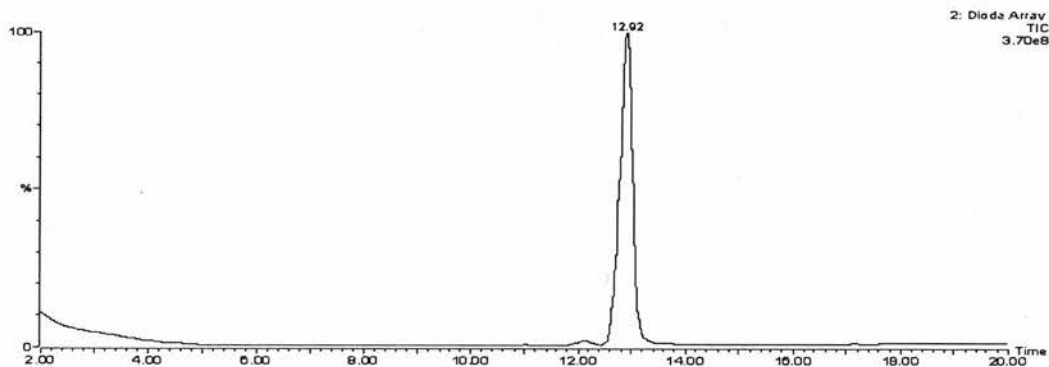
Compound 125



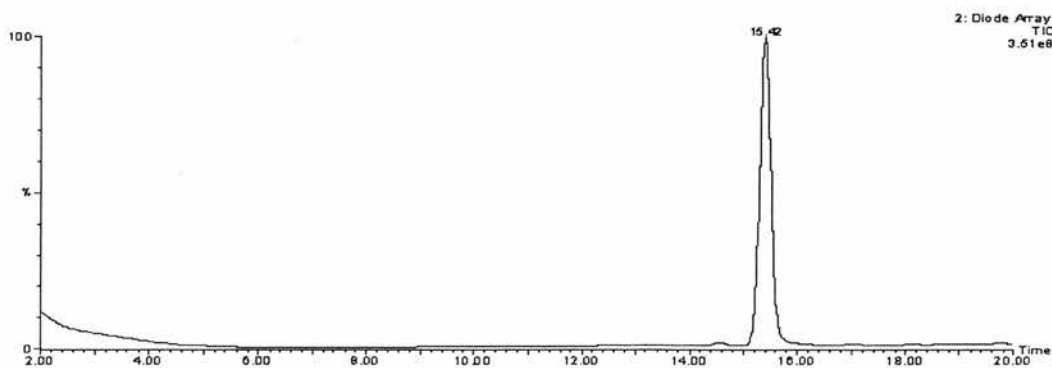
Compound 126



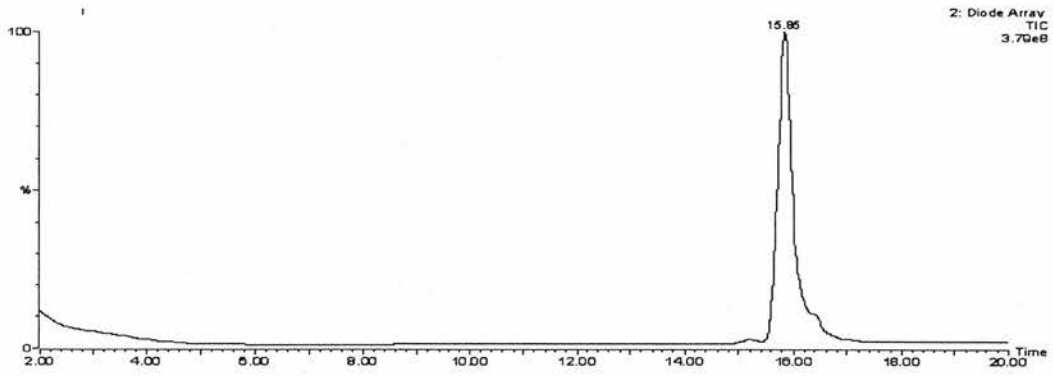
Compound 128



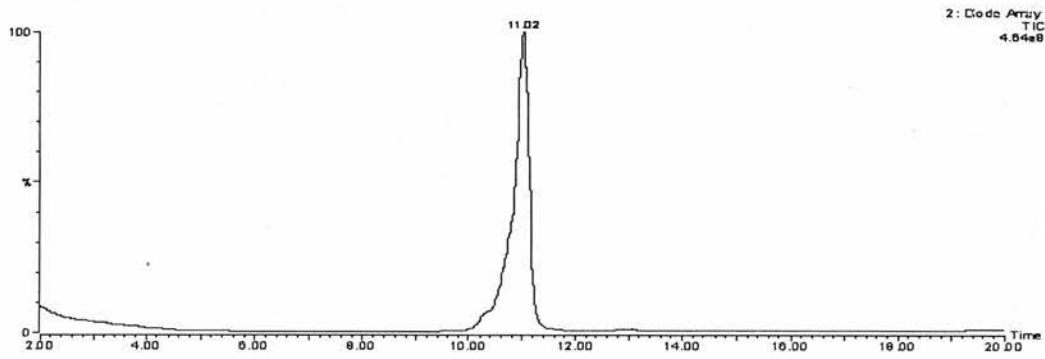
Compound 129



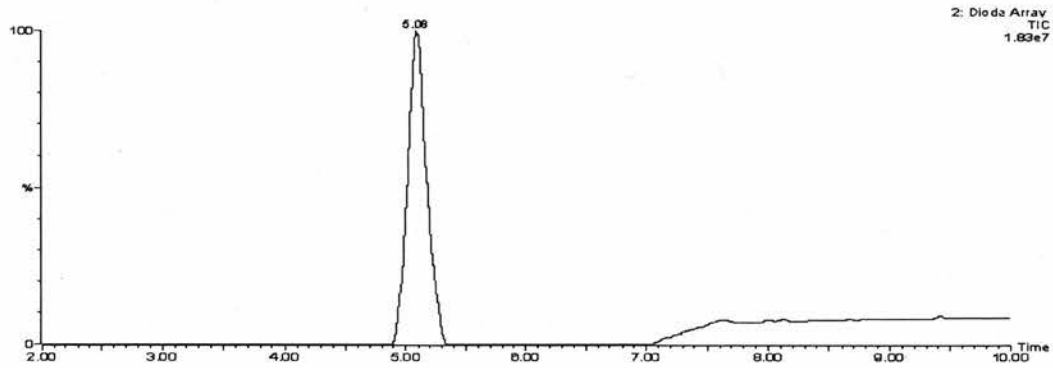
Compound 130



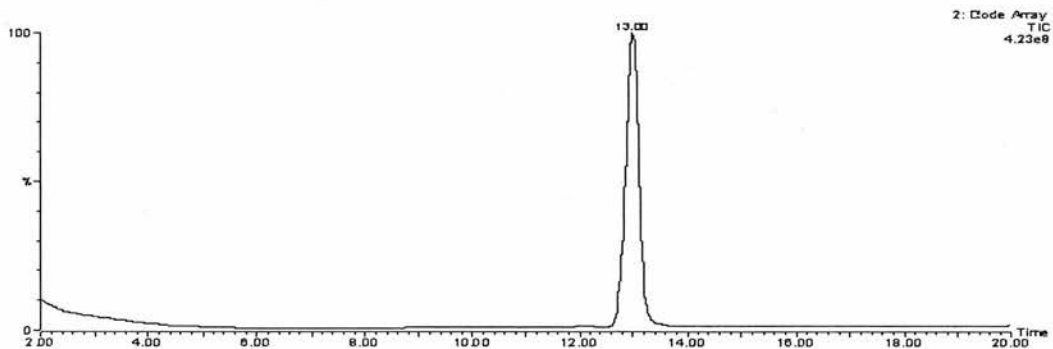
Compound 131



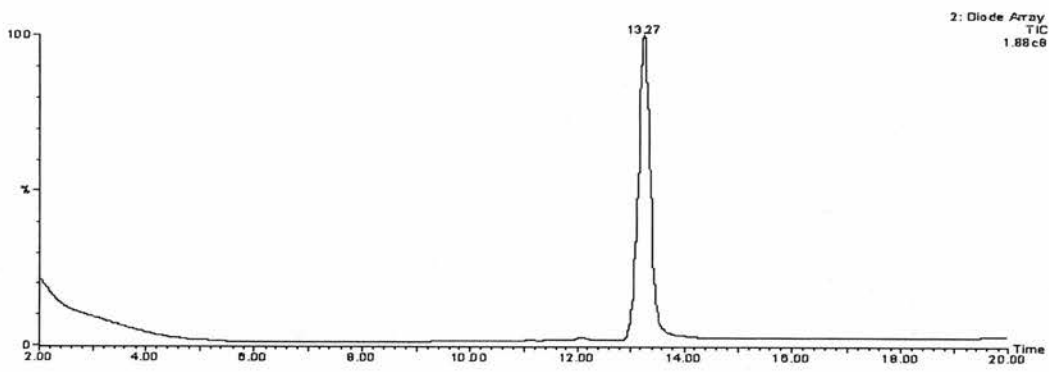
Compound 132



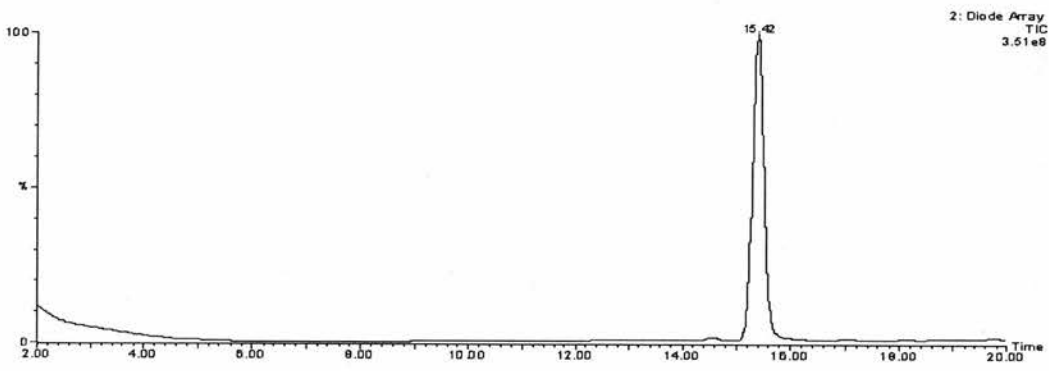
Compound 133



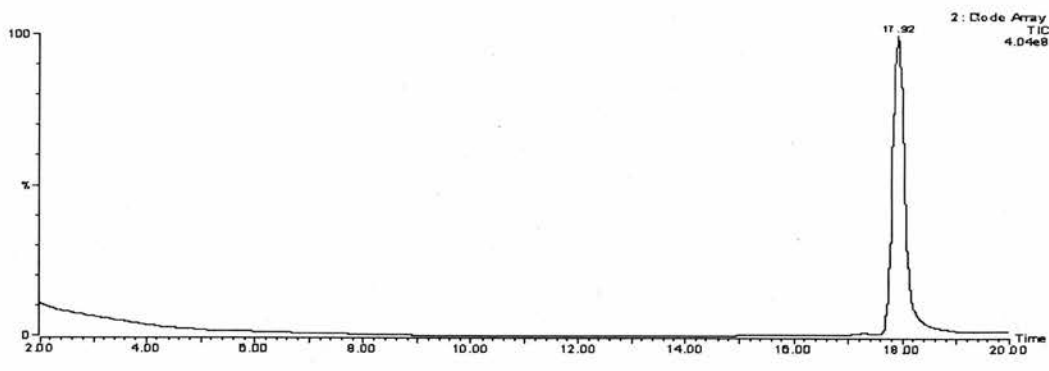
Compound 134



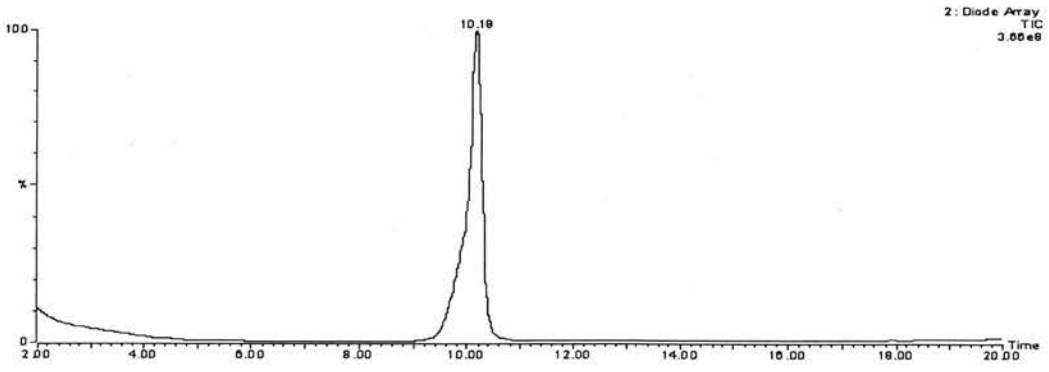
Compound 135



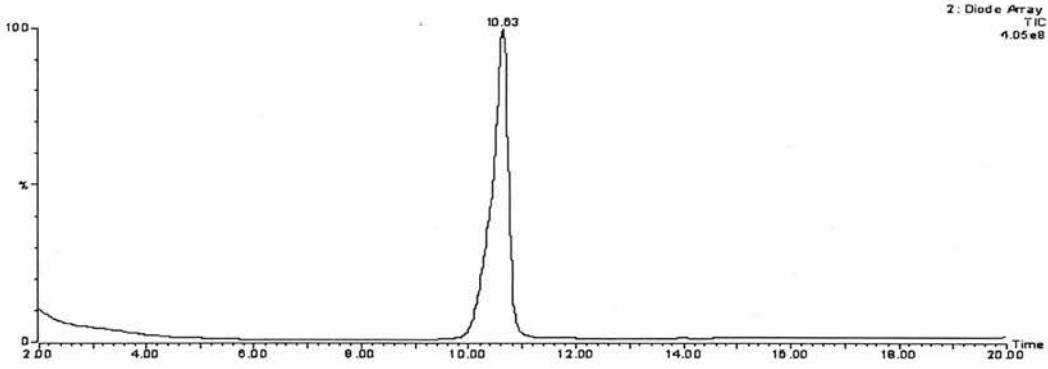
Compound 136



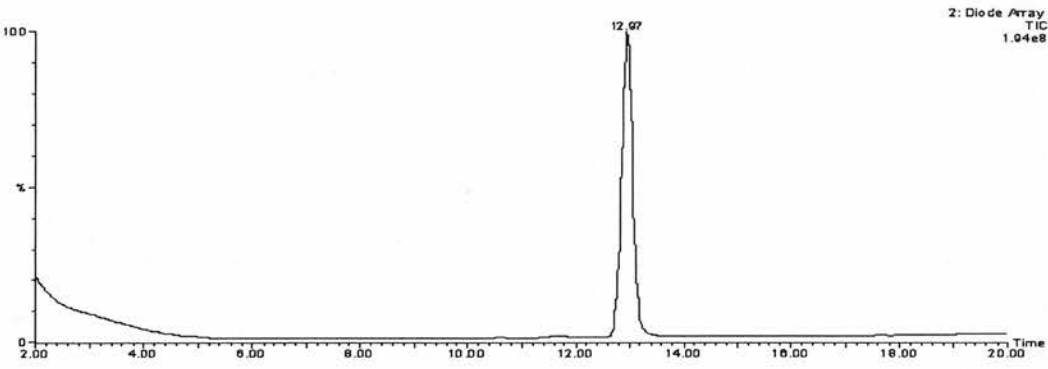
Compound 138



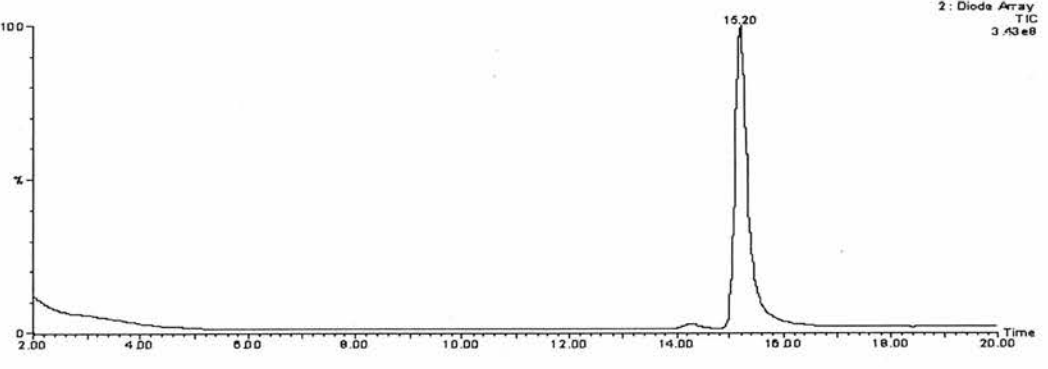
Compound 139



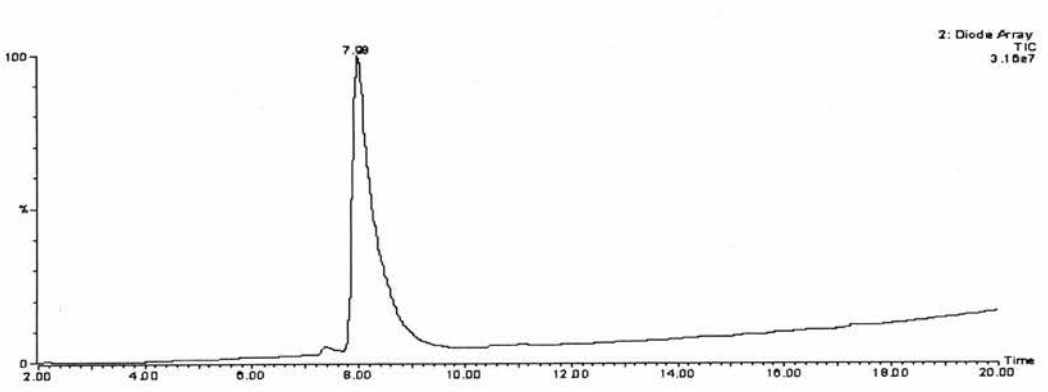
Compound 140



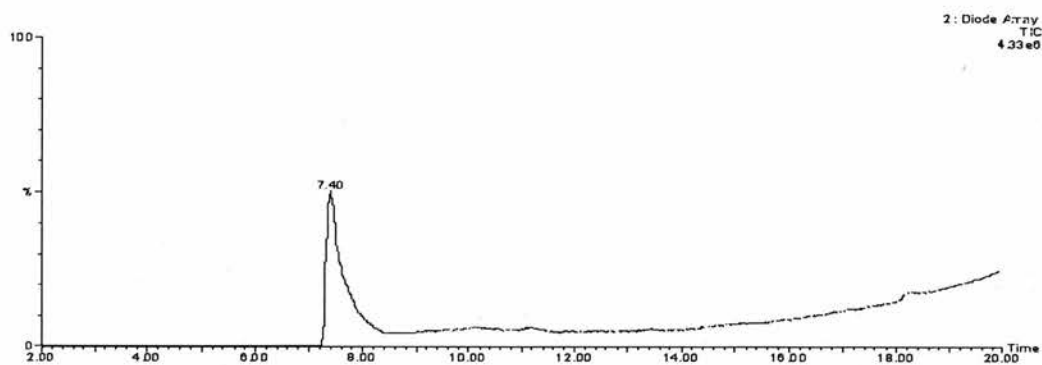
Compound 141



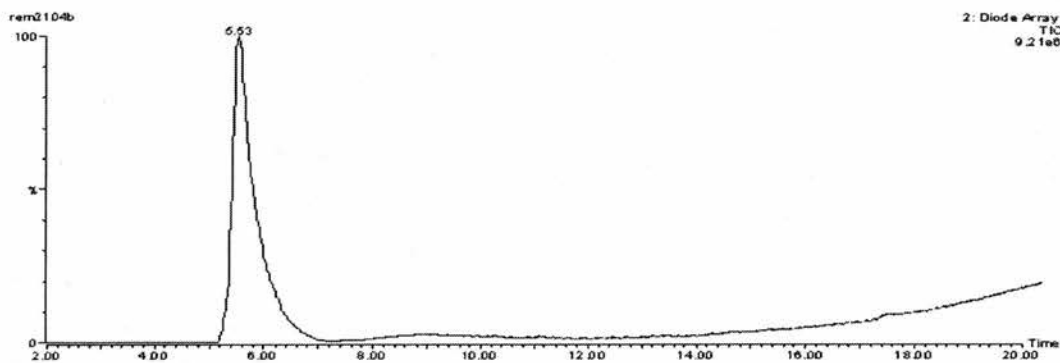
Compound 177



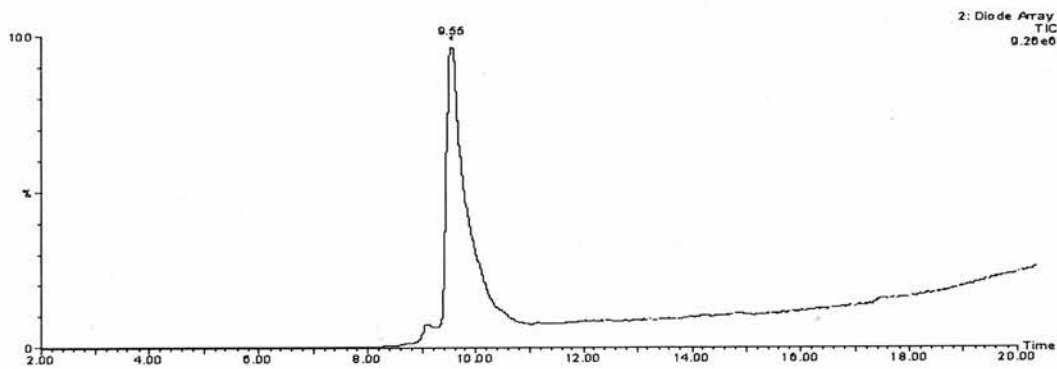
Compound 201



Compound 211

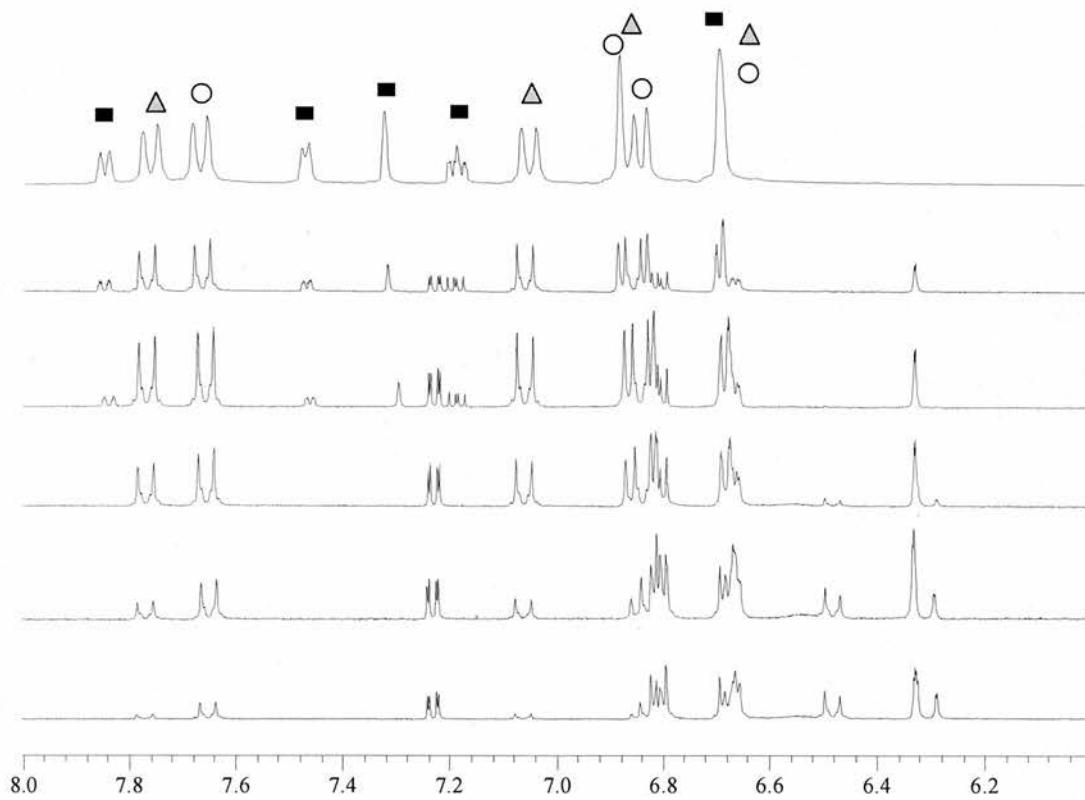


Compound 212

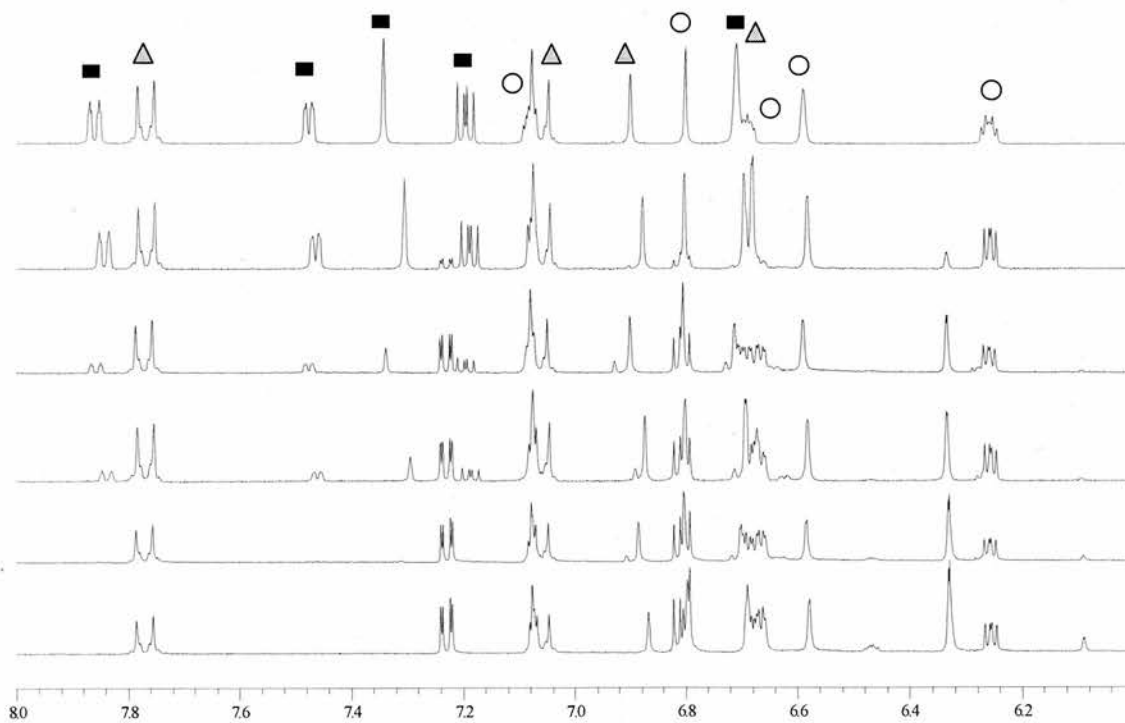


Appendix 2 : ^1H NMR competition experiment

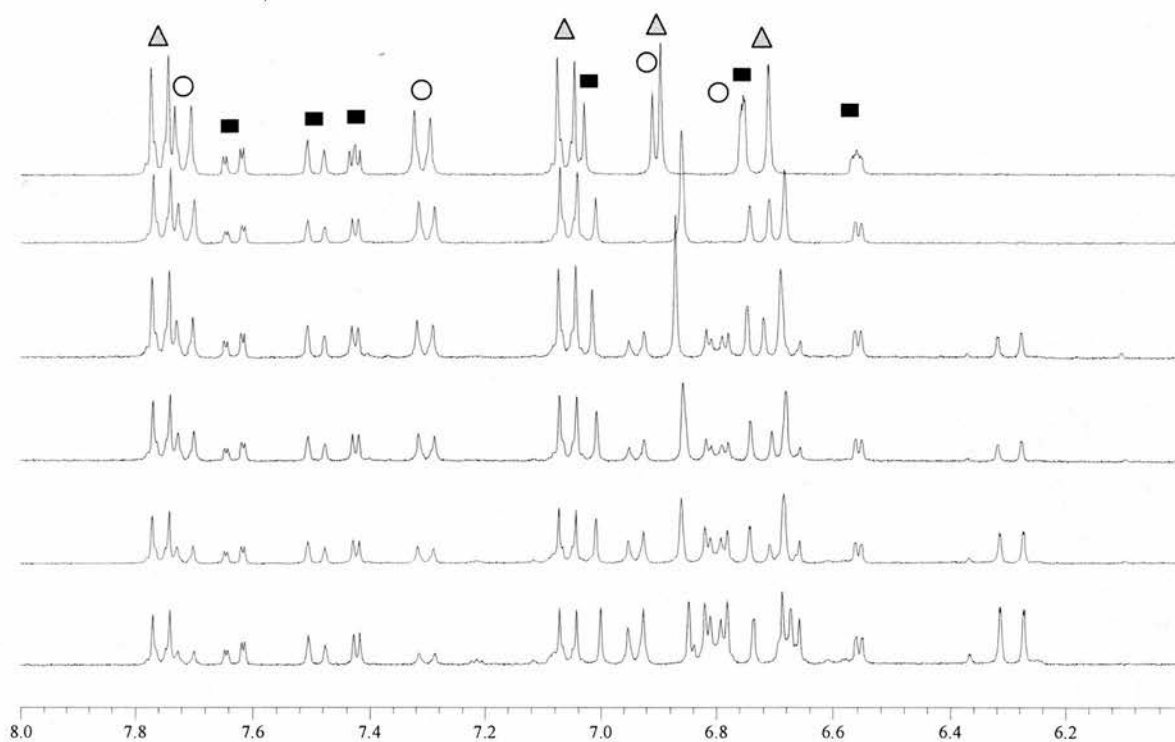
Set 2: compounds **44** (■), **131** (○), **132** (△).



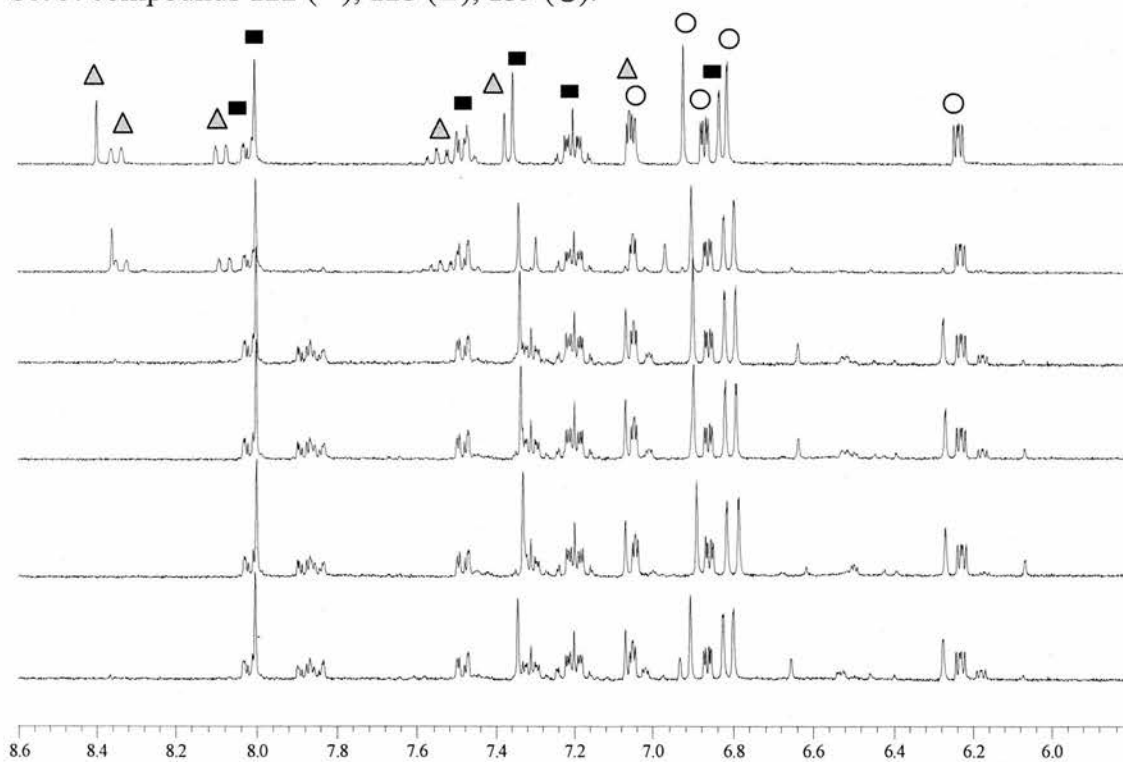
Set 3: compounds **44** (■), **132** (△), **138** (○).



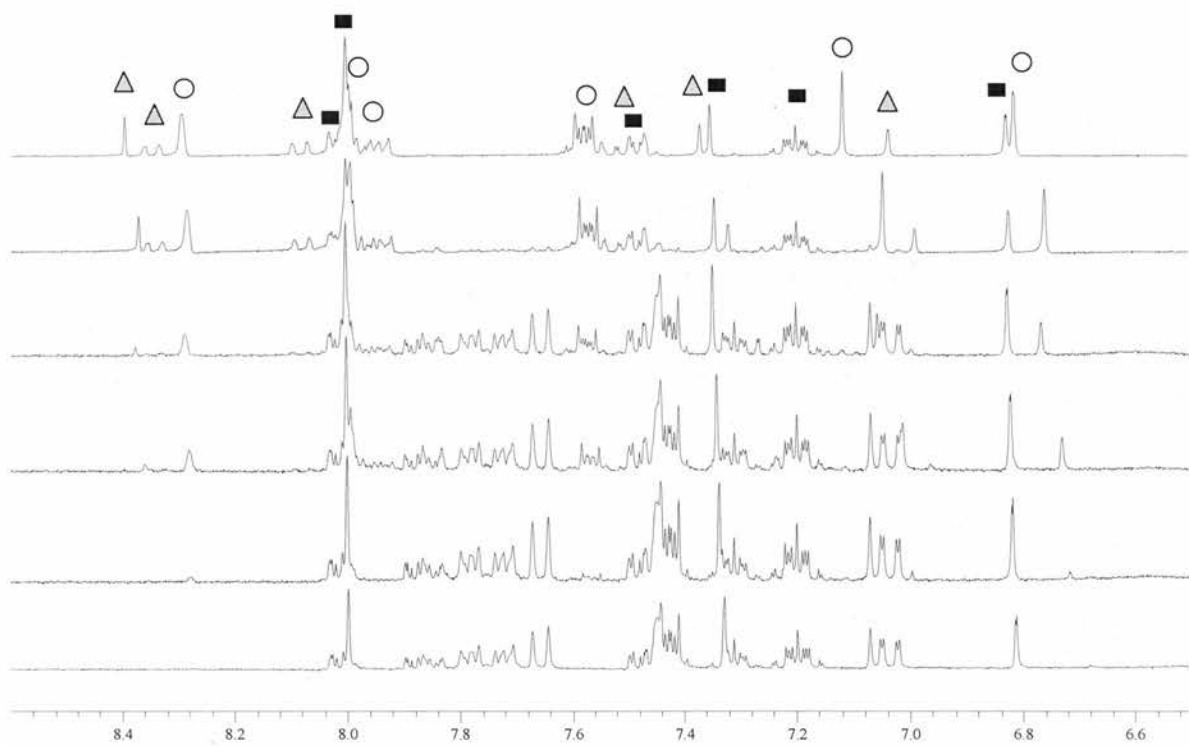
Set 4: compounds **126** (■), **135** (△), **136** (○).



Set 5: compounds **122** (■), **125** (△), **139** (○).



Set 6: compounds **122** (■), **125** (△), **129** (○).



Appendix 3: MS/MS data

The identified peptide fragments of inflammatory profilin.

Experiment 1:

(MATRIX) *(SCIENCE)* Mascot Search Results

Protein View

Match to: **Q58NA1_TOXGO** Score: **88**

Inflammatory profilin.- Toxoplasma gondii.

Found in search of X:\pe sciex data\Projects\data from Q-Star\rachel active neut.wiff

Nominal mass (M_r): **17544**; Calculated pI value: **4.40**

NCBI BLAST search of Q58NA1_TOXGO against nr

Unformatted sequence string for pasting into other applications

Taxonomy: Toxoplasma gondii

Links to retrieve other entries containing this sequence from NCBI

Entrez:

AA47289 from Toxoplasma gondii

AA33672 from Toxoplasma gondii

Variable modifications: Oxidation (M)

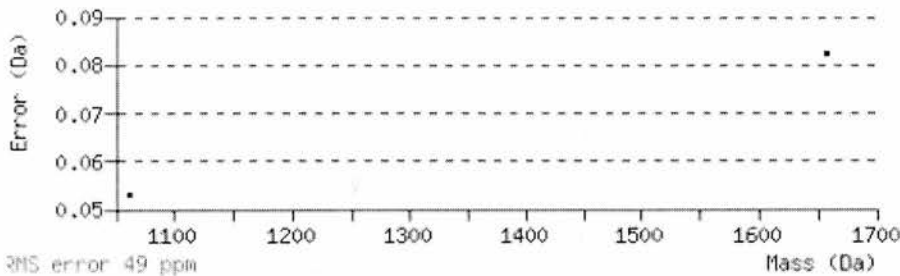
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: **15%**

Matched peptides shown in **Bold**

```
1 MSDWDPVVKE WLVDGTGYCCA GGIANAEDGV VFAAAADDDD GWSKLYKDDH
51 EEDTIGEDGN ACGKVSINEA STIKAAVDDG SAPNGVWIGG QKYKVVRRPEK
101 GFEYNDCTFD ITMCARSKGG AHLIKTPNGS IVIALYDEEK EQDKGNSRTS
151 ALAFAEYLHQ SGY
```

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
65 - 74	531.32	1060.63	1060.58	0.05	0	K.VSINEASTIK.A (<u>Ions score 41</u>)
149 - 163	829.44	1656.86	1656.78	0.08	0	R.TSALAFAEYLHQSGY.- (<u>Ions score 48</u>)



Experiment 3:

(MATRIX) *(SCIENCE)* Mascot Search Results

Protein View

Match to: Q58NA1_TOXGO Score: 87

Inflammatory profilin.- *Toxoplasma gondii*.

Found in search of X:\pe sciex data\Projects\data from Q-Star\Rachel acitve ALS.wiff

Nominal mass (M_r): **17544**; Calculated pI value: **4.40**

NCBI BLAST search of Q58NA1_TOXGO against nr

Unformatted sequence string for pasting into other applications

Taxonomy: *Toxoplasma gondii*

Links to retrieve other entries containing this sequence from NCBI

Entrez:

AAX47289 from *Toxoplasma gondii*

AAX33672 from *Toxoplasma gondii*

Variable modifications: Oxidation (M)

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P

Sequence Coverage: **26%**

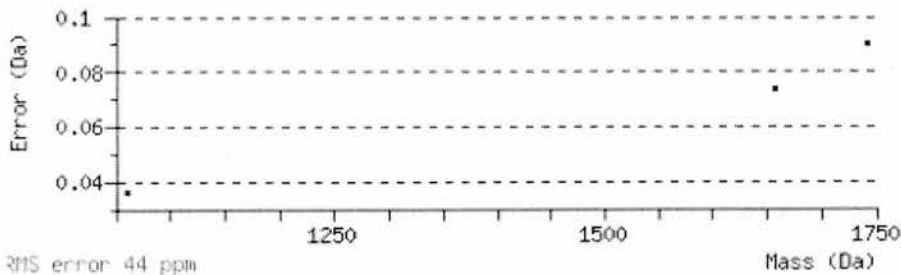
Matched peptides shown in **Bold**

1 MSDWDPVVKE WLVDVTGYCCA GGIANAEDGV VFAAAADDDD GWSKLYKDDH
51 EEDTIGEDGN ACGK**VSINEA STIKAAVDDG SAPNGVWIGG QKYKVV**RPEK
101 GFEYNDCTFD ITMCARSKGG AHLIKTPNGS IVIALYDEEK EQDKGNSR**TS**
151 **ALAFAEYLHQ SGY**

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
65 - 74	531.31	1060.61	1060.58	0.04	0	K.VSINEASTIK.A (<u>Ions score 36</u>)
75 - 92	871.47	1740.93	1740.84	0.09	0	K.AAVDDGSAPNGVWIGGQK.Y (<u>Ions</u>
149 - 163	829.43	1656.85	1656.78	0.07	0	R.TSALAFAEYLHQSGY.- (<u>Ions</u>

score 1)

score 51)



The peptide sequences identified from the trypsin digest of **195R** and not **213R**.

Experiment 1:

Q58NA1 TOXGO Inflammatory profilin.- Toxoplasma gondii.
Mass: 17544 **Score:** 88 **Queries matched:** 2

JC4214 translation elongation factor eEF-1 alpha - Ajellomyces
capsulate
Mass: 50180 **Score:** 59 **Queries matched:** 3

Q7QAL9 ANOGA ENSANGP00000011367 (Fragment).- Anopheles gambiae str.
PEST.
Score: 51 **Queries matched:** 2

Experiment 2:

JC4214 translation elongation factor eEF-1 alpha - Ajellomyces
capsulate
Mass: 50180 **Score:** 112 **Queries matched:** 2

Q6GSJ0 HUMAN Keratin 1.- Homo sapiens (Human).
Mass: 66027 **Score:** 99 **Queries matched:** 4

TRPGTR trypsin (EC 3.4.21.4) precursor - pig (tentative
sequence)
Mass: 24394 **Score:** 86 **Queries matched:** 5

Q6JUD8 9MAXI Elongation factor-1 alpha (Fragment).- Mesocyclops edax.
Mass: 41060 **Score:** 79 **Queries matched:** 2

S17253 alcohol dehydrogenase (EC 1.1.1.1) 4 precursor - yeast
(Kluyveromyces marxianus var. lactis)
Score: 73 **Queries matched:** 2

CAA44614 KLADH4 NID: - Kluyveromyces lactis
Mass: 40230 **Score:** 73 **Queries matched:** 2

I37984 keratin 9, type I, cytoskeletal - human
Mass: 61950 **Score:** 71 **Queries matched:** 2

Q9AVT7 PICAB Translation elongation factor-1 alpha (Fragment).- Picea
abies (Norway spruce) (Picea excelsa).
Mass: 49063 **Score:** 56 **Queries matched:** 2

Q6ZTI9 HUMAN Hypothetical protein FLJ44603.- Homo sapiens (Human).
Mass: 21420 **Score:** 52 **Queries matched:** 2

RNE HAEIN Ribonuclease E (EC 3.1.4.-) (RNase E).- Haemophilus
influenzae.
Mass: 106519 **Score:** 51 **Queries matched:** 2

Experiment 3:

<u>Q6GSJ0 HUMAN</u>	Keratin 1.- Homo sapiens (Human). Mass: 66027 Score: 328 Queries matched: 7
<u>I37984</u>	keratin 9, type I, cytoskeletal human Mass: 61950 Score: 110 Queries matched: 2
<u>Q58NA1 TOXGO</u>	Inflammatory profilin.- Toxoplasma gondii. Mass: 17544 Score: 87 Queries matched: 3
<u>EFHU1</u>	translation elongation factor eEF-1 alpha-1chain- human Mass: 50109 Score: 66 Queries matched: 1
<u>BAC57496</u>	AB104446 NID: - Homo sapiens Mass: 48569 Score: 49 Queries matched: 1

The total peptide sequence identifications for all the experiments:

Experiment 1, 196R:

DEBYA alcohol dehydrogenase (EC 1.1.1.1) 1 - yeast
(*Saccharomyces cerevisiae*)
Q6FQA4 CANGA *Candida glabrata* strain CBS138 chromosome I complete
sequence.- *Candida glabrata* (Yeast) (*Torulopsi*
Q9BLM8 TOXGO Hypothetical protein pdi (Protein disulfide isomerase
precursor).- *Toxoplasma gondii*.
ADH2 KLUMA Alcohol dehydrogenase II (EC 1.1.1.1).- *Kluyveromyces*
marxianus (Yeast) (*Kluyveromyces fragilis*).
K2C1 HUMAN Keratin, type II cytoskeletal 1 (Cytokeratin 1) (K1) (CK
1) (67 kDa cytokekeratin) (Hair alpha protei
Q58NA1 TOXGO Inflammatory profilin.- *Toxoplasma gondii*.
ABBOS serum albumin precursor [validated] - bovine
TRPGTR trypsin (EC 3.4.21.4) precursor - pig (tentative
sequence)
S17253 alcohol dehydrogenase (EC 1.1.1.1) 4 precursor - yeast
(*Kluyveromyces marxianus* var. *lactis*)
CAA44614 KLADH4 NID: - *Kluyveromyces lactis*
ANXA2 BOVIN Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I
heavy chain) (Chromobindin 8) (p36) (Protein I
Q5CHT5 9CRYT Elongation factor 1-alpha (EF-1-ALPHA).- *Cryptosporidium*
hominis.
Q9ZSW2 CYAPA Elongation translation factor 1 alpha.- *Cyanophora*
paradoxa.
JC4214 translation elongation factor eEF-1 alpha - *Ajellomyces*
capsulata
Q9BKE1 TOXGO Glyceraldehyde-3-phosphate dehydrogenase (Fragment).-
Toxoplasma gondii.
BAB55664 AB049736 NID: - *Toxoplasma gondii*
Q7QAL9 ANOGA ENSANGP00000011367 (Fragment).- *Anopheles gambiae* str.
PEST.
Q9P8S5 SACKL Alcohol dehydrogenase isozyme 3.- *Saccharomyces kluyveri*
(Yeast).
Q5CT27 CRYPV 40S ribosomal protein S21 (Fragment).- *Cryptosporidium*
parvum.

Experiment 1, 213R:

Q9BLM8 TOXGO Hypothetical protein pdi (Protein disulfide isomerase
precursor).- *Toxoplasma gondii*.
DEBYA alcohol dehydrogenase (EC 1.1.1.1) 1 - yeast
(*Saccharomyces cerevisiae*)
Q5UEM4 NEOCA Protein disulfide isomerase precursor (EC 5.3.4.1).-
Neospora caninum.
Q9BJ36 TOXGO ADP/ATP carrier.- *Toxoplasma gondii*.
AAC13766 TGU10429 NID: - *Toxoplasma gondii*
ANXA2 HUMAN Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I
heavy chain) (Chromobindin 8) (p36) (Protein I
Q86GI4 TOXGO Non-transmembrane antigen.- *Toxoplasma gondii*.
Q6MVV8 NEUCR Translation elongation factor eEF-1 alpha chain.-
Neurospora crassa.
Q59K68 CANAL Probable translation elongation factor EF-1 alpha.-
Candida albicans SC5314.
Q8MPF7 TOXGO Putative ribosomal protein S2.- *Toxoplasma gondii*.

Q6XQ70 SACPS Alcohol dehydrogenase 1 (EC 1.1.1.1).- *Saccharomyces pastorianus* (Lager yeast) (*Saccharomyces carls*
Q7RME1 PLAYO Actin.- *Plasmodium yoelii yoelii*.
Q9ZSW2 CYAPA Elongation translation factor 1 alpha.- *Cyanophora paradoxa*.
S43861 translation elongation factor eEF-1 alpha chain - *Podospora anserina*
Q6BMN8 DEBHA *Debaryomyces hansenii* chromosome F of strain CBS767 of *Debaryomyces hansenii* (Similar to sp|P16017
Q5CHT5 9CRYT Elongation factor 1-alpha (EF-1-ALPHA).- *Cryptosporidium hominis*.
S35772 translation elongation factor eEF-1 alpha chain - fungus (*Trichoderma reesei*)
Q5DQ94 TOXGO Receptor for activated C kinase 1.- *Toxoplasma gondii*.
Q5EVA3 9STRA Elongation factor 1 alpha (Fragment).- *Mallomonas rasilis*.
Q6PTK6 9CNID Elongation factor 1 alpha (Fragment).- *Obelia* sp. KJP-2004.
Q9HG86 PHYBL Translation elongation factor 1-alpha (Fragment).- *Phycomyces blakesleeanus*.
K2C1 HUMAN Keratin, type II cytoskeletal 1 (Cytokeratin 1) (K1) (CK 1) (67 kDa cytoke-
Q7Z901 SACCA Translation elongation factor 1-alpha (Fragment).- *Saccharomyces castellii* (Yeast).
TRPGTR trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
A44861 keratin, 67K type II epidermal - human
Q6WE46 AMOPR Actin.- *Amoeba proteus* (Amoeba).
Q7YW79 SCHJA Elongation factor 1-alpha.- *Schistosoma japonicum* (Blood fluke).
Q8H9B0 9CARY Eukaryotic elongation factor 1A.- *Suaeda japonica*.
Q8TG61 9FUNG Translation elongation factor 1-alpha (Fragment).- *Harpochytrium* sp. JEL94.
Q6XQ77 SACBA Alcohol dehydrogenase 2 (EC 1.1.1.1).- *Saccharomyces bayanus* (Yeast) (*Saccharomyces uvarum*).
Q5EUA2 PHATR Elongation factor 1 alpha (Fragment).- *Phaeodactylum tricornutum*.
Q7YTV4 9METZ Elongation factor 1 alpha.- *Trichoplax adhaerens*.
Q6FQA4 CANGA *Candida glabrata* strain CBS138 chromosome I complete sequence.- *Candida glabrata* (Yeast) (*Torulopsi*
PDIA6 MOUSE Protein disulfide-isomerase A6 precursor (EC 5.3.4.1) (Thioredoxin domain containing protein 7).- M
Q76II3 9ZYGO Translation elongation factor-1 alpha (Fragment).- *Syncephalis depressa*.
Q96979 9CILI Translation elongation factor 1-alpha (Fragment).- *Paranophrys carnivora*.
Q76II4 9ZYGO Translation elongation factor-1 alpha (Fragment).- *Dimargaris cristalligena*.
Q9BNU8 HETSP Elongation factor-1alpha (Fragment).- *Heterometrus spinifer* (Asia giant forest scorpion).
Q7Z8Y0 9ASCO Translation elongation factor 1-alpha (Fragment).- *Candida castellii*.
Q5MXP0 9NEOP Elongation factor-1 alpha (Fragment).- *Taygetis virgilia*.
Q9M516 CAPAN Translation elongation factor 1a.- *Capsicum annuum* (Bell pepper).
Q5FBV7 9STRA Beta-tubulin (Fragment).- *Protoopalina japonica*.
Q5VGR0 VERDA Translation elongation factor 1 alpha (Fragment).- *Verticillium dahliae* (*Verticillium wilt*).
Q8WTA0 9NEOP Elongation factor-1 alpha (Fragment).- *Pycnocentroides* sp. UMSP-New Zealand.

Q9GPX4 9CUCU Elongation factor 1 alpha (Fragment).- *Chaetophloeus penicillatus*.
B53522 20k cyclophilin - *Toxoplasma gondii* (fragment)
Q590R2 9HYME Elongation factor-1 alpha (Fragment).- *Clypeadon* sp.
Clsp819.
BAB55664 AB049736 NID: - *Toxoplasma gondii*
Q5ETI2 9NEOP Elongation factor 1 alpha (Fragment).- *Heliconius melpomene melpomene*.
Q5VGS5 9HYPO Translation elongation factor 1 alpha (Fragment).- *Cordyceps heteropoda*.
Q5QTD4 9MAXI Beta-actin.- *Tigriopus japonicus*.
Q5W533 9DIPT Elongation factor 1-alpha (Fragment).- *Morellia xanthoptera*.
Q9NA14 9TURB Putative elongation factor 1-alpha (Fragment).- *Symsagittifera roscofensis*.
Q9BNV1 9ARAC Elongation factor-1alpha (Fragment).- *Epactiochernes tumidus*.
KRHU0 keratin 10, type I, cytoskeletal - human
A45634 actin - *Cryptosporidium parvum*
Q95UQ2 TOXGO Subtilisin-like protein.- *Toxoplasma gondii*.
S16340 tubulin beta chain - *Toxoplasma gondii*
S21909 translation elongation factor eEF-1 alpha chain - malaria parasite (*Plasmodium falciparum*)
T43894 translation elongation factor 1 alpha [imported] - *Aspergillus oryzae*
Q8MZN7 9CUCU Elongation factor 1-alpha (Fragment).- *Xyleborus meritus*.
ADH2 KLUMA Alcohol dehydrogenase II (EC 1.1.1.1).- *Kluyveromyces marxianus* (Yeast) (*Kluyveromyces fragilis*).
Q8MZP1 9CUCU Elongation factor 1-alpha (Fragment).- *Xyleborus spathipennis*.
Q6RFC9 9CUCU Elongation factor 1 alpha (Fragment).- *Aphanarthrum affine*.
Q9NHB8 TOXGO ADP ribosylation factor 1.- *Toxoplasma gondii*.
Q870J9 9HYPO Elongation factor 1-alpha (Fragment).- *Escovopsis* sp. Esc10.
Q9GPV9 9CUCU Elongation factor 1 alpha (Fragment).- *Hylastes porculus*.
O64457 NANBA Actin.- *Nannochloris bacillaris* (Green alga).
Q5URE0 9HYME Elongation factor 1-alpha (Fragment).- *Pogonomymex huachucanus*.
Q5VGT2 9HYPO Translation elongation factor 1 alpha (Fragment).- *Balansia henningsiana*.
Q8T5L6 9ASIL Elongation factor-1 alpha (Fragment).- *Hybos* sp. NCSU-99071974.
Q967Q2 9CUCU Elongation factor 1-alpha (Fragment).- *Miocryphalus* sp. SCA13.
AAA30031 SUSACT1S1 NID: - *Strongylocentrotus purpuratus*
Q5UK87 9DINO Beta-tubulin (Fragment).- *Kryptoperidinium foliaceum*.
Q86GP6 9NEOP Elongation factor-1alpha (Fragment).- *Heteronympha merope*.
Q870K4 9HYPO Elongation factor 1-alpha (Fragment).- *Escovopsis weberi*.
Q8MY84 TOXGO Fructose-1,6-bisphosphatase.- *Toxoplasma gondii*.
Q5EPW3 CHLAC Elongation factor 1 alpha (Fragment).- *Chlosyne acastus* (Sagebrush checkerspot).
Q7YZD7 9BIVA Actin.- *Chlamys farreri*.
K1C16 HUMAN Keratin, type I cytoskeletal 16 (Cytokeratin 16) (K16) (CK 16).- *Homo sapiens* (Human).
ABBOS serum albumin precursor [validated] - bovine
ACT YARLI Actin.- *Yarrowia lipolytica* (*Candida lipolytica*).
Q910C1 RANCA Larva-specific keratin RLK.- *Rana catesbeiana* (Bull frog).
Q5SGD8 TOXGO Tgd057.- *Toxoplasma gondii*.

Q8TG64 9PEZI Actin (Fragment).- *Cercospora beticola*.
XWBO ADP,ATP carrier protein T1 - bovine
Q9P8S5 SACKL Alcohol dehydrogenase isozyme 3.- *Saccharomyces kluyveri*
 (Yeast).
Q5R7W2 PONPY Hypothetical protein DKFZp4590198.- *Pongo pygmaeus*
 (Orangutan).
S65079 actin - *Cyanidioschyzon merolae*
Q5B4R4 EMENI Hypothetical protein.- *Aspergillus nidulans* FGSC A4.
Q8ILW9 PLAF7 Actin II.- *Plasmodium falciparum* (isolate 3D7).
Q9GRE8 TOXGO Hsp70 protein.- *Toxoplasma gondii*.
Q7Z9S9 9ASCO Actin (Fragment).- *Candida humilis*.
AAG60329 AF123457 NID: - *Toxoplasma gondii*
ATAXE actin - *Entamoeba histolytica*
Q9ZQT0 GLYEC Actin (Fragment).- *Glycyrrhiza echinata* (Licorice).
Q6NX10 BRARE Hypothetical protein zgc:77591.- *Brachydanio rerio*
 (Zebrafish) (*Danio rerio*).
K2C1 MOUSE Keratin, type II cytoskeletal 1 (Cytokeratin 1) (67 kDa
 cytokeratin).- *Mus musculus* (Mouse).
Q964N7 9TURB Elongation factor 1-a (Fragment).- *Childia groenlandica*.
Q9BNV4 9ARAC Elongation factor-1alpha (Fragment).- *Chthonius*
tetrachelatus.
S17253 alcohol dehydrogenase (EC 1.1.1.1) 4 precursor - yeast
 (*Kluyveromyces marxianus* var. *lactis*)
CAA44614 KLADH4 NID: - *Kluyveromyces lactis*
ARF1 SCHPO ADP-ribosylation factor 1.- *Schizosaccharomyces pombe*
 (Fission yeast).
Q8MPF6 TOXGO Putative translation initiation factor 5A2.- *Toxoplasma*
gondii.
Q8IJ34 PLAF7 ADP/ATP transporter on adenylate translocase.- *Plasmodium*
falciparum (isolate 3D7).
Q5CNC3 9CRYT Glycogen phosphorylase 1.- *Cryptosporidium hominis*.
AAG22213 AE014297 NID: - *Drosophila melanogaster*
R3YL16 ribosomal protein S16, cytosolic - large-leaved lupine
ADT3 BOVIN ADP,ATP carrier protein, isoform T2 (ADP/ATP translocase
 3) (Adenine nucleotide translocator 3) (AN)
Q5KSP3 FUGRU Adenine nucleotide translocator s254.- *Fugu rubripes*
 (Japanese pufferfish) (*Takifugu rubripes*).
Q8T5N6 9ASIL Elongation factor-1 alpha (Fragment).- *Hoplopeza* sp.
 NCSU-95051124.
Q95VX4 9MYRI ADP-ATP translocator.- *Ethmostigmus rubripes*.
Q9BKE1 TOXGO Glyceraldehyde-3-phosphate dehydrogenase (Fragment).-
Toxoplasma gondii.
AAA30138 TOXANT28K NID: - *Toxoplasma gondii*
Q8MUM2 TOXGO Facilitative glucose transporter.- *Toxoplasma gondii*.
ADT1 MOUSE ADP,ATP carrier protein, heart/skeletal muscle isoform T1
 (ADP/ATP translocase 1) (Adenine nucleoti
Q5J1K2 ELAGV Actine.- *Elaeis guineensis* var. *tenera* (Oil palm).
P90613 TOXGO Lactate dehydrogenase (EC 1.1.1.27).- *Toxoplasma gondii*.
Q90ZF7 RANCA Keratin 8.- *Rana catesbeiana* (Bull frog).
Q7Z9U2 9SACH Actin (Fragment).- *Saccharomyces spencerorum*.
Q9VHV7 DROME CG7878-PA (LD33749p).- *Drosophila melanogaster* (Fruit
 fly).
Q7XYZ1 GRIJA Putative 40S ribosomal protein S19 (Fragment).-
Griffithsia japonica (Red alga).
AAA61223 HUMTRL NID: - *Homo sapiens*
Q8MVR4 9SPIT ADP/ATP carrier.- *Euplotes* sp.
Q84X74 CHLRE CR057 protein.- *Chlamydomonas reinhardtii*.
Q5U412 MOUSE Intergral membrane protein 1.- *Mus musculus* (Mouse).
Q697Y8 9STRA Beta tubulin (Fragment).- *Phytophthora tentaculata*.
Q7RYS8 NEUCR Hypothetical protein.- *Neurospora crassa*.
S16339 tubulin alpha chain - *Toxoplasma gondii*

Q8SRB2 ENCCU P68-LIKE PROTEIN (DEAD BOX FAMILY OF RNA HELICASES). -
 Encephalitozoon cuniculi.
HABO hemoglobin alpha chain [validated] - bovine
HACZ hemoglobin alpha chain - chimpanzee
A25359 hemoglobin alpha chain - European suslik
CAD32275 FDO487677 NID: - Felis catus
Q6TP33 9APIC Actin (Fragment).- Monocystis agilis.
Q5D4W0 PANGI Actin (Fragment).- Panax ginseng (Korean ginseng).
O65204 EUGGR Actin.- Euglena gracilis.
ARF4 HUMAN ADP-ribosylation factor 4.- Homo sapiens (Human).
Q5CFF5 9CRYT Pfsec61.- Cryptosporidium hominis.
Q5CKG0 9CRYT Casein kinase i.- Cryptosporidium hominis.
Q7N977 PHOLL Maltodextrin phosphorylase.- Photorhabdus luminescens
 (subsp. laumondii).
Q853N0 9CAUD Gp47.- Mycobacteriophage Bx1.
Q72T96 LEPIC GGDEF family protein.- Leptospira interrogans (serogroup
 Icterohaemorrhagiae / serovar Copenhageni).
I64203 aspartate-tRNA ligase (EC 6.1.1.12) - Mycoplasma
 genitalium
Q6PVZ3 CHICK Type II alpha-keratin IIC.- Gallus gallus (Chicken).
D83588 conserved hypothetical protein PA0454 [imported] -
 Pseudomonas aeruginosa (strain PAO1)
O44030 TOXGO Folate synthesis protein [INCLUDES: 2-amino-4-hydroxy-6-
 hydroxymethyldihydropteridine pyrophosphok
JC2385 protein disulfide-isomerase (EC 5.3.4.1) ER60 precursor -
 bovine
BAA03759 HUMPLCALF1 NID: - Homo sapiens
DAPB SHEON Dihydrodipicolinate reductase (EC 1.3.1.26) (DHPR).-
 Shewanella oneidensis.
Q9GSV5 TOXGO TgMIC10 precursor.- Toxoplasma gondii.
Q8H8U6 ORYSA Putative dehydroascorbate reductase.- Oryza sativa
 (japonica cultivar-group).
Q7RNC2 PLAYO Putative leucine aminopeptidase.- Plasmodium yoelii
 yoelii.
BAC90552 BA000045 NID: - Gloeobacter violaceus PCC 7421
S57262 actin - red alga (Chondrus crispus)
Q5CT27 CRYPV 40S ribosomal protein S21 (Fragment).- Cryptosporidium
 parvum.
BAA13351 D87407 NID: - Branchiostoma floridae
Q25129 HALRO ADT/ATP translocase.- Halocynthia roretzi' (Sea squirt).
O45038 SCHJA HSP70.- Schistosoma japonicum (Blood fluke).
Q9FNV2 9FLOR Elongation factor 2 (Fragment).- Botryocladia uvarioides.
Q6WRX4 BRABE ADP/ATP translocase.- Branchiostoma belcheri Q5FVC1 XENTR
 Hypothetical protein.- Xenopus tropicalis (Western clawed frog)
 (Silurana tropicalis).

Experiment 2, 196R:

Q9BLM8 TOXGO Hypothetical protein pdi (Protein disulfide isomerase
 precursor).- Toxoplasma gondii.
Q5CHT5 9CRYT Elongation factor 1-alpha (EF-1-ALPHA).- Cryptosporidium
 hominis.
Q9ZSW2 CYAPA Elongation translation factor 1 alpha.- Cyanophora
 paradoxa.
JC4214 translation elongation factor eEF-1 alpha - Ajellomyces
 capsulata
DEBYA alcohol dehydrogenase (EC 1.1.1.1) 1 - yeast
 (Saccharomyces cerevisiae)
AAC13766 TGU10429 NID: - Toxoplasma gondii
Q6GSJ0 HUMAN Keratin 1.- Homo sapiens (Human).

TRPGTR trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)

Q6JUD8 9MAXI Elongation factor-1 alpha (Fragment).- Mesocyclops edax.

Q6FQA4 CANGA Candida glabrata strain CBS138 chromosome I complete sequence.- Candida glabrata (Yeast) (Torulopsi S17253 alcohol dehydrogenase (EC 1.1.1.1) 4 precursor - yeast (Kluyveromyces marxianus var. lactis)

CAA44614 KLADH4 NID: - Kluyveromyces lactis

I37984 keratin 9, type I, cytoskeletal - human

Q9AVT7 PICAB Translation elongation factor-1 alpha (Fragment).- Picea abies (Norway spruce) (Picea excelsa).

Q95UQ2 TOXGO Subtilisin-like protein.- Toxoplasma gondii.

Q6ZTI9 HUMAN Hypothetical protein FLJ44603.- Homo sapiens (Human).

RNE HAEIN Ribonuclease E (EC 3.1.4.-) (RNase E).- Haemophilus influenzae.

Q5CT27 CRYPV 40S ribosomal protein S21 (Fragment).- Cryptosporidium parvum.

Experiment 2, 213R:

Q9BLM8 TOXGO Hypothetical protein pdi (Protein disulfide isomerase precursor).- Toxoplasma gondii.

Q8TBV2 HUMAN Annexin A2, isoform 2.- Homo sapiens (Human).

AAC13766 TGU10429 NID: - Toxoplasma gondii

DEBYA alcohol dehydrogenase (EC 1.1.1.1) 1 - yeast (Saccharomyces cerevisiae)

Q86GI4 TOXGO Non-transmembrane antigen.- Toxoplasma gondii.

Q5UEM4 NEOCA Protein disulfide isomerase precursor (EC 5.3.4.1).- Neospora caninum.

Q9BJ36 TOXGO ADP/ATP carrier.- Toxoplasma gondii.

Q8MPF7 TOXGO Putative ribosomal protein S2.- Toxoplasma gondii.

S16340 tubulin beta chain - Toxoplasma gondii

Q9GRE8 TOXGO Hsp70 protein.- Toxoplasma gondii.

Q5FBV7 9STRA Beta-tubulin (Fragment).- Protoopalina japonica.

AAG60329 AF123457 NID: - Toxoplasma gondii

Q95UQ2 TOXGO Subtilisin-like protein.- Toxoplasma gondii.

Q6XQ78 SACBA Alcohol dehydrogenase 1 (EC 1.1.1.1).- Saccharomyces bayanus (Yeast) (Saccharomyces uvarum).

Q5DQ94 TOXGO Receptor for activated C kinase 1.- Toxoplasma gondii.

Q9ZSW2 CYAPA Elongation translation factor 1 alpha.- Cyanophora paradoxa.

Q8MY84 TOXGO Fructose-1,6-bisphosphatase.- Toxoplasma gondii.

Q6WE46 AMOPR Actin.- Amoeba proteus (Amoeba).

Q5CHT5 9CRYT Elongation factor 1-alpha (EF-1-ALPHA).- Cryptosporidium hominis.

Q6XQ70 SACPS Alcohol dehydrogenase 1 (EC 1.1.1.1).- Saccharomyces pastorianus (Lager yeast) (Saccharomyces carls

UBKM tubulin beta chain - Chlamydomonas reinhardtii

Q84KM0 PERWI Beta-tubulin (Fragment).- Peridinium willei (Dinoflagellate).

Q6MVV8 NEUCR Translation elongation factor eEF-1 alpha chain.- Neurospora crassa.

Q6A8K1 CAEEL Actin protein 4, isoform c.- Caenorhabditis elegans.

Q5G927 9STRA Beta-tubulin (Fragment).- Phytophthora palmivora.

Q59K68 CANAL Probable translation elongation factor EF-1 alpha.- Candida albicans SC5314.

S30514 tubulin beta chain - Naegleria gruberi

B30309 tubulin beta chain - Euplotes crassus

1AN1E tryptase inhibitor, chain E - medicinal leech (fragments)

PDIA6 MOUSE Protein disulfide-isomerase A6 precursor (EC 5.3.4.1)
 (Thioredoxin domain containing protein 7).- M
Q6L8Q1 9PEZI Translation elongation factor 1 alpha chain.- *Rosellinia*
sp. PF1022.
S35772 translation elongation factor eEF-1 alpha chain - fungus
 (*Trichoderma reesei*)
AAK54650 AF378368 NID: - *Coccidioides immitis*
Q7Z875 CRYNV Translation elongation factor 2.- *Cryptococcus neoformans*
var. grubii (*Filobasidiella neoformans* va
Q5TT61 ANOGA ENSANGP00000023203.- *Anopheles gambiae* str. PEST.
UBZQF tubulin beta chain - malaria parasite (*Plasmodium*
falciparum)
Q6PTK6 9CNID Elongation factor 1 alpha (Fragment).- *Obelia* sp. KJP-
 2004.
Q9SAU8 WHEAT HSP70.- *Triticum aestivum* (Wheat).
AAH14775 BC014775 NID: - *Homo sapiens*
ACTA PHYPO Actin, plasmodial isoform.- *Physarum polycephalum* (Slime
 mold).
Q9STD1 ZINEL Beta-tubulin.- *Zinnia elegans* (*Zinnia*).
Q6PTR8 ASTMI Elongation factor 1 alpha (Fragment).- *Asterina miniata*
 (Bat star).
B53522 20k cyclophilin - *Toxoplasma gondii* (fragment)
Q9GQB8 9EUKA Beta-tubulin (Fragment).- *Acrasis rosea*.
S16339 tubulin alpha chain - *Toxoplasma gondii*
Q6YIC5 9APIC Beta-tubulin (Fragment).- *Babesia odocoilei*.
Q6XQ77 SACBA Alcohol dehydrogenase 2 (EC 1.1.1.1).- *Saccharomyces*
bayanus (Yeast) (*Saccharomyces uvarum*).
Q5QTD4 9MAXI Beta-actin.- *Tigriopus japonicus*.
Q86R56 9CRYT Actin (Fragment).- *Cryptosporidium* sp. 938.
Q25099 HYDMA Translation elongation factor 1 alpha.- *Hydra*
magnipapillata (*Hydra*).
A48470 translation elongation factor eEF-1 alpha chain - *Eimeria*
bovis (fragment)
Q9LU13 BLAHO Hsp70 protein.- *Blastocystis hominis*.
Q7Z901 SACCA Translation elongation factor 1-alpha (Fragment).-
Saccharomyces castellii (Yeast).
Q9HGB4 9ZYGO Translation elongation factor 1-alpha (Fragment).-
Dissophora decumbens.
ADT2 BOVIN ADP,ATP carrier protein 2 (ADP/ATP translocase 2)
 (Adenine nucleotide translocator 2) (ANT 2) (Solu
S36753 dnaK-type molecular chaperone SSA3 - yeast (*Saccharomyces*
cerevisiae)
Q94FN9 CHLS6 Actin-2 (Fragment).- *Chlorarachnion* sp. (strain CCMP 621)
 (*Pedinomonas minutissima*).
Q5EVA3 9STRA Elongation factor 1 alpha (Fragment).- *Mallomonas*
rasilis.
Q5KSP3 FUGRU Adenine nucleotide translocator s254.- *Fugu rubripes*
 (Japanese pufferfish) (*Takifugu rubripes*).
Q9BKE1 TOXGO Glyceraldehyde-3-phosphate dehydrogenase (Fragment).-
Toxoplasma gondii.
A45634 actin - *Cryptosporidium parvum*
K2C1 HUMAN Keratin, type II cytoskeletal 1 (Cytokeratin 1) (K1) (CK
 1) (67 kDa cytotokeratin) (Hair alpha protei
Q8WTA0 9NEOP Elongation factor-1 alpha (Fragment).- *Pycnocentroides* sp.
UMSP-New Zealand.
Q76FS2 ORYSA Beta-tubulin (Beta-3 tubulin) (Tubulin beta subunit).-
Oryza sativa (*japonica* cultivar-group).
Q697Y8 9STRA Beta tubulin (Fragment).- *Phytophthora tentaculata*.
Q9LD24 CHLS6 Beta-tubulin-5 (Fragment).- *Chlorarachnion* sp. (strain
 CCMP 621) (*Pedinomonas minutissima*).
Q8BF61 9ZZZZ Actin (Fragment).- uncultured organism.

Q6FQA4 CANGA *Candida glabrata* strain CBS138 chromosome I complete sequence.- *Candida glabrata* (Yeast) (Torulopsi
Q9GSJ2 9DIOP Elongation factor-1 alpha (Fragment).- *Teloglabrus milleri*.
Q5CKG0 9CRYT Casein kinase i.- *Cryptosporidium hominis*.
Q6NX10 BRARE Hypothetical protein zgc:77591.- *Brachydanio rerio* (Zebrafish) (*Danio rerio*).
Q95UK6 9DIPT Elongation factor-1 alpha (Fragment).- *Gluma keyzeri*.
Q6PBF5 XENTR Adenine nucleotide translocase.- *Xenopus tropicalis* (Western clawed frog) (*Silurana tropicalis*).
Q7Z8Q8 YEAST Translation elongation factor 1-alpha (Fragment).- *Saccharomyces cerevisiae* (Baker's yeast).
Q590R2 9HYME Elongation factor-1 alpha (Fragment).- *Clypeadon* sp. Clsp819.
Q5ETI2 9NEOP Elongation factor 1 alpha (Fragment).- *Heliconius melpomene melpomene*.
Q5QE59 9NEOP Elongation factor-1 alpha (Fragment).- *Maculinea arion* (large blue).
Q592D0 9HYME Elongation factor 1-alpha (Fragment).- *Eufriesia pulchra*.
Q5V8L7 PAXIN Actin.- *Paxillus involutus* (Naked brimcap).
Q8IJ34 PLAF7 ADP/ATP transporter on adenylate translocase.- *Plasmodium falciparum* (isolate 3D7).
Q7Z8Y0 9ASCO Translation elongation factor 1-alpha (Fragment).- *Candida castellii*.
Q9NA14 9TURB Putative elongation factor 1-alpha (Fragment).- *Symsagittifera roscofensis*.
Q5W533 9DIPT Elongation factor 1-alpha (Fragment).- *Morellia xanthoptera*.
Q9U9M0 9HEMI Elongation factor 1 alpha (Fragment).- *Stomaphis graffi*.
Q7YW79 SCHJA Elongation factor 1-alpha.- *Schistosoma japonicum* (Blood fluke).
Q8H9B0 9CARY Eukaryotic elongation factor 1A.- *Suaeda japonica*.
Q64457 NANBA Actin.- *Nannochloris bacillaris* (Green alga).
AAA30030 SUSACT05 NID: - *Strongylocentrotus purpuratus*
Q8I810 PANRE Heat shock protein 70 A2 (Fragment).- *Panagrellus redivivus*.
Q9GPV9 9CUCU Elongation factor 1 alpha (Fragment).- *Hylastes porculus*.
Q6T599 9ALVE Actin.- *Perkinsus marinus*.
Q8T7D2 9NEOP Elongation factor-1 alpha (Fragment).- *Antherina suraka*.
A48462 dense granule protein GRA4 - *Toxoplasma gondii*
Q9U540 TOXGO Heat shock protein 70 precursor.- *Toxoplasma gondii*.
Q7RDR5 PLAYO Elongation factor 2.- *Plasmodium yoelii yoelii*.
S21909 translation elongation factor eEF-1 alpha chain - malaria parasite (*Plasmodium falciparum*)
Q5MXN9 9NEOP Elongation factor-1 alpha (Fragment).- *Taygetis sosis*.
Q9NHB8 TOXGO ADP ribosylation factor 1.- *Toxoplasma gondii*.
Q8T7G7 9DIPT Elongation factor 1 alpha (Fragment).- *Lespesia datanarum*.
Q7RGU3 PLAYO GTP-binding nuclear protein ran/tc4.- *Plasmodium yoelii yoelii*.
Q6WRX4 BRABE ADP/ATP translocase.- *Branchiostoma belcheri tsingtaunense*.
Q6RFC9 9CUCU Elongation factor 1 alpha (Fragment).- *Aphanarthrum affine*.
Q95VX4 9MYRI ADP-ATP translocator.- *Ethmostigmus rubripes*.
Q8ILW9 PLAF7 Actin II.- *Plasmodium falciparum* (isolate 3D7).
Q7PDD0 PLAYO Translation elongation factor EF-1, subunit alpha.- *Plasmodium yoelii yoelii*.
ADT1 RAT ADP,ATP carrier protein 1 (ADP/ATP translocase 1) (Adenine nucleotide translocator 1) (ANT 1) (Solu

Q9GPX4 9CUCU Elongation factor 1 alpha (Fragment).- *Chaetophloeus penicillatus*.
Q8MZN7 9CUCU Elongation factor 1-alpha (Fragment).- *Xyleborus meritus*.
BAB55664 AB049736 NID: - *Toxoplasma gondii*
Q8ITR0 9SPIT Alpha-tubulin (Fragment).- *Strombidinopsis* sp.
S51682 dnaK-type molecular chaperone hsp70 - *Eimeria maxima*
(fragment)
Q8IH88 DROME AT28834p (Fragment).- *Drosophila melanogaster* (Fruit fly).
ADH2 KLUMA Alcohol dehydrogenase II (EC 1.1.1.1).- *Kluyveromyces marxianus* (Yeast) (*Kluyveromyces fragilis*).
Q7RHV1 PLAYO Adenine nucleotide translocase.- *Plasmodium yoelii yoelii*.
Q9XY78 9CRUS Elongation factor 1-alpha (Fragment).- *Cypridopsis vidua*.
Q95UK4 9DIPT Elongation factor-1 alpha (Fragment).- *Gluma nitida*.
Q5URD4 9HYME Elongation factor 1-alpha (Fragment).- *Pogonomyrmex lobatus*.
Q7YWH3 9CUCU Elongation factor 1-alpha (Fragment).- *Metamasius hemipterus* (silky cane weevil).
Q9BNV1 9ARAC Elongation factor-1alpha (Fragment).- *Epactiochernes tumidus*.
AAA61223 HUMTRL NID: - *Homo sapiens*
Q5EPW3 CHLAC Elongation factor 1 alpha (Fragment).- *Chlosyne acastus* (Sagebrush checkerspot).
Q9U9N2 9HEMI Elongation factor 1 alpha (Fragment).- *Cinara etsuhoe*.
Q9P8Z2 9ZYGO Beta-tubulin 2 (Fragment).- *Mortierella verticillata*.
A44861 keratin, 67K type II epidermal - human
Q9GPY3 9CUCU Elongation factor 1 alpha (Fragment).- *Chramesus asperatus*.
Q6NYR4 BRARE Heat shock 70kDa protein 8.- *Brachydanio rerio* (Zebrafish) (*Danio rerio*).
Q9ZQT0 GLYEC Actin (Fragment).- *Glycyrrhiza echinata* (Licorice).
Q8TG64 9PEZI Actin (Fragment).- *Cercospora beticola*.
Q7Z9S9 9ASCO Actin (Fragment).- *Candida humilis*.
Q7Z9U2 9SACH Actin (Fragment).- *Saccharomyces spencerorum*.
Q870J9 9HYPO Elongation factor 1-alpha (Fragment).- *Escovopsis* sp.
Esc10.
Q9M516 CAPAN Translation elongation factor 1a.- *Capsicum annuum* (Bell pepper).
Q7Z9S6 9SACH Actin (Fragment).- *Kluyveromyces polysporus*.
AAA30138 TOXANT28K NID: - *Toxoplasma gondii*
Q8ITQ3 9SPIT Alpha-tubulin (Fragment).- *Laboea strobila*.
Q9U4T9 TOXGO Dense granule protein GRA8.- *Toxoplasma gondii*.
Q5KFM5 CRYNE P68-like protein, putative.- *Cryptococcus neoformans* (*Filobasidiella neoformans*).
Q93147 BOTSH Heat shock protein 70.- *Botryllus schlosseri* (Star ascidian).
T43894 translation elongation factor 1 alpha [imported] - *Aspergillus oryzae*
Q5IH90 TOXGO Calcium ATPase SERCA-like.- *Toxoplasma gondii*.
Q8MPF6 TOXGO Putative translation initiation factor 5A2.- *Toxoplasma gondii*.
Q9BKD1 9EUKA Alpha-tubulin (Fragment).- *Jakoba libera*.
Q9PUG5 GADMO Beta-3 tubulin.- *Gadus morhua* (Atlantic cod).
Q25129 HALRO ADT/ATP translocase.- *Halocynthia roretzi* (Sea squirt).
S65079 actin - *Cyanidioschyzon merolae*
Q5EG80 USTMA Translation elongation factor EF1-alpha (Fragment).- *Ustilago maydis* (Smut fungus).
Q870K9 9HYPO Elongation factor 1-alpha (Fragment).- *Escovopsis* sp.
Esc20.
Q5CFC7 9CRYT Ribosomal protein S5.- *Cryptosporidium hominis*.

ARF PLAFA ADP-ribosylation factor.- Plasmodium falciparum.
ATPK HUMAN ATP synthase f chain, mitochondrial (EC 3.6.3.14).- Homo sapiens (Human).
Q8MPF5 TOXGO Putative glucose-6-phosphate-1-dehydrogenase.- Toxoplasma gondii.
Q5R7W2 PONPY Hypothetical protein DKFZp4590198.- Pongo pygmaeus (Orangutan).
Q5CNF7 9CRYT GTP-binding nuclear protein ran/tc4.- Cryptosporidium hominis.
Q6DUA7 TOXGO Hsp10.- Toxoplasma gondii.
Q5CNJ7 9CRYT Similar to RNA-dependent helicase p68 (DEAD-box protein p68) (DEAD-box protein 5).- Cryptosporidium
B84577 probable casein kinase I [imported] - Arabidopsis thaliana
ARF ASHGO ADP-ribosylation factor.- Ashbya gossypii (Yeast) (Eremothecium gossypii).
A24701 tubulin beta-3 chain - chicken
Q6P4W2 XENTR Hypothetical protein MGC76265.- Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
CAA39465 SCDB1G NID: - Saccharomyces cerevisiae
Q5J1K2 ELAGV Actine.- Elaeis guineensis var. tenera (Oil palm).
Q5B5G2 EMENI EF1A ASPOR Elongation factor 1-alpha (EF-1-alpha).- Aspergillus nidulans FGSC A4.
JC2385 protein disulfide-isomerase (EC 5.3.4.1) ER60 precursor - bovine
JC5704 protein disulfide-isomerase (EC 5.3.4.1) ER60 precursor - human
AAD30129 AF136344 NID: - Toxoplasma gondii
P90613 TOXGO Lactate dehydrogenase (EC 1.1.1.27).- Toxoplasma gondii.
AAC78495 OCU24656 NID: - Oryctolagus cuniculus
Q65204 EUGGR Actin.- Euglena gracilis.
T15206 hypothetical protein W02D3.6 - Caenorhabditis elegans
Q5I0C1 XENTR Hypothetical LOC496924.- Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
Q5CFF5 9CRYT Pfsec61.- Cryptosporidium hominis.
Q9D2K8 MOUSE Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4833436C19 product:keratin complex 2, basic, gene 1, full insert sequence.- Mus musculus (Mouse).
AAD05191 MUSKTEP2A NID: - Mus musculus
Q7PDT0 PLAYO ATP-dependent RNA helicase.- Plasmodium yoelii yoelii.
Q7RYS8 NEUCR Hypothetical protein.- Neurospora crassa.
Q86L89 9CUCU Elongation factor-1 alpha (Fragment).- Disonycha conjuncta.
Q45038 SCHJA HSP70.- Schistosoma japonicum (Blood fluke).
A41677 ADP,ATP carrier protein - Chlorella kessleri
Q7RER2 PLAYO Heat shock protein.- Plasmodium yoelii yoelii.
Q9BKD0 9EUKA Alpha-tubulin (Fragment).- Malawimonas jakobiformis.
S57262 actin - red alga (Chondrus crispus)
Q7RBS1 PLAYO Ribosomal protein S7.- Plasmodium yoelii yoelii.
Q6NUK1 HUMAN Solute carrier family 25 member 24, isoform 1.- Homo sapiens (Human).
Q86CZ8 CAEEL Hypothetical protein tag-61.- Caenorhabditis elegans.
Q8MUA7 HETGL Heat shock protein 70-C.- Heterodera glycines (Soybean cyst nematode).
T34037 heat shock 70K protein D - Caenorhabditis elegans
AAB64307 DCBTUB1 NID: - Daucus carota
Q5XXT2 9HEMI Putative mitochondrial ADP/ATP translocase.- Oncometopia nigricans.
KRHUEA keratin 6a, type II - human
ACT3 SOYBN Actin 3.- Glycine max (Soybean).

Q8I9N6 SCHGR Helicase RM62-like protein D (Fragment).- *Schistocerca gregaria* (Desert locust).
Q6IVC1 ORYSA Small GTP-binding protein.- *Oryza sativa* (japonica cultivar-group).
Q5WQM4 9STRA Enolase 1 (Fragment).- *Apodachlya brachynema*.
Q86L79 9CUCU Elongation factor-1 alpha (Fragment).- *Asphaera auripennis*.
Q7RNC2 PLAYO Putative leucine aminopeptidase.- *Plasmodium yoelii yoelii*.
Q7XYZ1 GRIJA Putative 40S ribosomal protein S19 (Fragment).- *Griffithsia japonica* (Red alga).
K2C8 MOUSE Keratin, type II cytoskeletal 8 (Cytokeratin 8) (Cytokeratin endo A).- *Mus musculus* (Mouse).
Q44030 TOXGO Folate synthesis protein [INCLUDES: 2-amino-4-hydroxy-6-hydroxymethyl-dihydropteridine pyrophosphok
Q5D4W0 PANGI Actin (Fragment).- *Panax ginseng* (Korean ginseng).
Q5CKN9 9CRYT Small GTP-binding protein sar1.- *Cryptosporidium hominis*.
CAA03813 SEQUENCE 1 FROM PATENT WO9727300.- unidentified.
Q9BNV4 9ARAC Elongation factor-1alpha (Fragment).- *Chthonius tetrachelatus*.
S21907 actin, muscle - sea squirt (*Styela clava*)
Q9UR33 9ZYGO Beta-tubulin 1 (Beta-tubulin 2) (Fragment).- *Conidiobolus coronatus*.
Q5A4E2 CANAL Hypothetical protein DED1.- *Candida albicans* SC5314.
HSP71 SCHPO Probable heat shock protein ssa1.- *Schizosaccharomyces pombe* (Fission yeast).
Q9N647 LEIMA ADP/ATP carrier, copy 2 (ATP/ADP translocase) (ADP/ATP carrier, copy 1).- *Leishmania major*.
Q8SRB2 ENCCU P68-LIKE PROTEIN (DEAD BOX FAMILY OF RNA HELICASES).- *Encephalitozoon cuniculi*.
S49303 dnaK-type molecular chaperone hsp70 - fungus (*Cladosporium herbarum*)
Q9NG25 TOXGO Toxofilin.- *Toxoplasma gondii*.
BAA13351 D87407 NID: - *Branchiostoma floridae*
Q5RFB6 PONPY Hypothetical protein DKFZp46902032.- *Pongo pygmaeus* (Orangutan).
Q90ZF7 RANCA Keratin 8.- *Rana catesbeiana* (Bull frog).
UBHU5B tubulin beta chain - human
Q5XNR0 TOXGO SEC61-gamma.- *Toxoplasma gondii*.
BAC90552 BA000045 NID: - *Gloeobacter violaceus* PCC 7421
Q5AJ91 CANAL Potential ADP-ribosylation factor.- *Candida albicans* SC5314.
Q5HZL3 XENLA LOC496366 protein.- *Xenopus laevis* (African clawed frog).
Q42067 ARATH GTP-binding protein (Fragment).- *Arabidopsis thaliana* (Mouse-ear cress).
Q7P3U0 FUSNV Export ABC transporter.- *Fusobacterium nucleatum* subsp. *vincentii* ATCC 49256.
Q6TP33 9APIC Actin (Fragment).- *Monocystis agilis*.
Q6YFA5 NEOCA Small heat shock protein.- *Neospora caninum*.
Q8GWA5 ARATH Hypothetical protein.- *Arabidopsis thaliana* (Mouse-ear cress).
Q5XLC5 TOXGO Small heat shock protein 21.- *Toxoplasma gondii*.
Q9P8S5 SACKL Alcohol dehydrogenase isozyme 3.- *Saccharomyces kluyveri* (Yeast).
Q7KSW1 DROME CG2747-PB, isoform B.- *Drosophila melanogaster* (Fruit fly).
Q7RS99 PLAYO PfMPC.- *Plasmodium yoelii yoelii*.
Q6CGD8 YARLI *Yarrowia lipolytica* chromosome A of strain CLIB99 of *Yarrowia lipolytica*.- *Yarrowia lipolytica* (Can
Q8F321 LEPIN Flagellar motor switch protein fliG.- *Leptospira interrogans*.

Q8WPN9 TOXGO Serine-threonine phosphatase 2C.- *Toxoplasma gondii*.
Q5G7K5 9STRA PPAT5.- *Hyaloperonospora parasitica*.
Q5QKN5 9AQUI Hsp70 (Fragment).- *Hydrogenobacter hydrogenophilus*.
Q643Z5 9GAMM DnaK.- *Shewanella* sp. Ac10.
Q61012 PARTE Heat shock protein 70 (Fragment).- *Paramecium*
tetraurelia.
Q9FNV2 9FLOR Elongation factor 2 (Fragment).- *Botryocladia uvarioides*.
Q5CT27 CRYPV 40S ribosomal protein S21 (Fragment).- *Cryptosporidium*
parvum.
Q56D08 PICPA Kar2p.- *Pichia pastoris* (Yeast).
AAA21808 SPIHSC70A NID: - *Spinacia oleracea*
G90539 heat shock protein DNaK [imported] - *Mycoplasma pulmonis*
(strain UAB CTIP)
Q616S0 CAEBR Hypothetical protein CBG15166.- *Caenorhabditis briggsae*.
Q9AQX2 EUGGR Alpha-tubulin.- *Euglena gracilis*.
Q7RTB6 PLAYO Dynein beta chain, ciliary.- *Plasmodium yoelii yoelii*.
Q5SNE7 ORYSA Hypothetical protein P0512G09.2-1 (Hypothetical protein
P0695A04.11- 1).- *Oryza sativa* (*japonica* cu
Q9TVW6 CAEEL Hypothetical protein gei-18.- *Caenorhabditis elegans*.
Q8L6Z3 ARATH Hypothetical protein At4g38290 (Hypothetical protein
At4g38280).- *Arabidopsis thaliana* (Mouse-ear c
T10619 hypothetical protein F21C20.120 - *Arabidopsis thaliana*
D90092 hypothetical protein orf665 [imported] - *Guillardia*
Q853N0 9CAUD Gp47.- *Mycobacteriophage Bxz1*.
Q6UNNO 9HYPO Ribosome-associated protein RAP1-like protein
(Fragment).- *Epichloe festucae*.
AAA90976 SPU02369 NID: - *Strongylocentrotus purpuratus*
D83588 conserved hypothetical protein PA0454 [imported] -
Pseudomonas aeruginosa (strain PA01)
Q5FVC1 XENTR Hypothetical protein.- *Xenopus tropicalis* (Western clawed
frog) (*Silurana tropicalis*).
Q8MTM4 TOXGO Endoplasmic reticulum retention receptor.- *Toxoplasma*
gondii.
Q6PVZ3 CHICK Type II alpha-keratin IIC.- *Gallus gallus* (Chicken).
Q5SKM6 THET8 Nicotinate phosphoribosyltransferase.- *Thermus*
thermophilus (strain HB8 / ATCC 27634 / DSM 579).
Q613Y6 CAEBR Hypothetical protein CBG16141.- *Caenorhabditis briggsae*.
Q5FN14 LACAC DNA polymerase III, beta chain (EC 2.7.7.7).-
Lactobacillus acidophilus.
Q5NYT8 AZOSE Similar to twitching motility protein PilU.- *Azoarcus* sp.
(strain EbN1).
Q8MUM2 TOXGO Facilitative glucose transporter.- *Toxoplasma gondii*.
Q8BFZ3 MOUSE *Mus musculus* 10 days neonate skin cDNA, RIKEN full-length
enriched library, clone:4732493G14 product:ACTIN, CYTOPLASMIC TYPE 5
homolog (*Mus musculus* 10 days neonate skin cDNA, RIKEN full-length
enriched library, clone:4732495G21 product:ACTIN, CYTOPLASMIC TYPE 5
homolog).- *Mus musculus* (Mouse).
ARF4 HUMAN ADP-ribosylation factor 4.- *Homo sapiens* (Human).
Q5E7Q7 VIBF1 ATP-dependent RNA helicase SrmB.- *Vibrio fischeri* (strain
ATCC 700601 / ES114).
Q8NXX1 TRIVE Nonribosomal peptide synthetase.- *Trichoderma virens*.
R3YL16 ribosomal protein S16, cytosolic - large-leaved lupine
Q8TT10 METAC Predicted protein.- *Methanosarcina acetivorans*.
AAB60213 TGU10246 NID: - *Toxoplasma gondii*
Q5RD51 PONPY Hypothetical protein DKFZp459P1426 (Fragment).- *Pongo*
pygmaeus (Orangutan).
Q5SGD8 TOXGO Tgd057.- *Toxoplasma gondii*.
S07002 actin 1 - carrot
Q5G905 PHATR Actin (Fragment).- *Phaeodactylum tricorutum*.
Q6BI20 DEBHA Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-
lyase).- *Debaryomyces hansenii* (Yeast) (*Torulaspora hansenii*).

Q86L83 9CUCU Elongation factor-1 alpha (Fragment).- *Omophoita octoguttata*.
B64129 dethiobiotin synthase homolog HI1550 - *Haemophilus influenzae* (strain Rd KW20)
LRCH1 MOUSE Leucine-rich repeats and calponin homology domain containing protein 1 (Calponin homology domain containing protein 1).- *Mus musculus* (Mouse).
Q5WYV3 LEGPL Hypothetical protein.- *Legionella pneumophila* (strain Lens).
Q8PVZ0 METMA Cell division protein.- *Methanosarcina mazei* (*Methanosarcina frisia*).
Q8Q2G5 9HIV1 Envelope glycoprotein (Fragment).- Human immunodeficiency virus 1.
Q600T0 MYCHY Hypothetical protein.- *Mycoplasma hyopneumoniae*.
AF1826 hypothetical protein all0158 [imported] - *Nostoc* sp. (strain PCC 7120)
Q9FJS8 ARATH *Arabidopsis thaliana* genomic DNA, chromosome 5, P1 clone:MIK19.- *Arabidopsis thaliana* (Mouse-ear cr
Q5SLN7 THET8 Leucine aminopeptidase.- *Thermus thermophilus* (strain HB8 / ATCC 27634 / DSM 579).
Q7G2B1 ORYSA Putative valyl tRNA synthetase.- *Oryza sativa* (japonica cultivar-group).
Q9IVC7 9HIV1 Reverse transcriptase (Fragment).- Human immunodeficiency virus 1.
Q8A3I1 BACTN Transcriptional regulator.- *Bacteroides thetaiotaomicron*.
O31093 RHILV Putative glycosyl transferase PssG.- *Rhizobium leguminosarum* (biovar *viciae*).
BAB48183 BA000012 NID: - *Mesorhizobium loti* MAFF303099

Experiment 3, 196R:

Q6GSJ0 HUMAN Keratin 1.- *Homo sapiens* (Human).
I37984 keratin 9, type I, cytoskeletal - human
Q58NA1 TOXGO Inflammatory profilin.- *Toxoplasma gondii*.
TRPGTR trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
EFHU1 translation elongation factor eEF-1 alpha-1 chain - human
Q5CHT5 9CRYT Elongation factor 1-alpha (EF-1-ALPHA).- *Cryptosporidium hominis*.
BAC57496 AB104446 NID: - *Homo sapiens*

Experiment 3, 213R:

A44861 keratin, 67K type II epidermal - human
DEBYA alcohol dehydrogenase (EC 1.1.1.1) 1 - yeast (*Saccharomyces cerevisiae*)
TRPGTR trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
Q9ZSW2 CYAPA Elongation translation factor 1 alpha.- *Cyanophora paradoxa*.
AAC13766 TGU10429 NID: - *Toxoplasma gondii*
Q5TT61 ANOGA ENSANGP00000023203.- *Anopheles gambiae* str. PEST.
Q25099 HYDMA Translation elongation factor 1 alpha.- *Hydra magnipapillata* (Hydra).
A48470 translation elongation factor eEF-1 alpha chain - *Eimeria bovis* (fragment)
Q640D9 XENLA LOC494720 protein.- *Xenopus laevis* (African clawed frog).
Q5CHT5 9CRYT Elongation factor 1-alpha (EF-1-ALPHA).- *Cryptosporidium hominis*.

Q9BJ36 TOXGO ADP/ATP carrier.- *Toxoplasma gondii*.
Q5EVA3 9STRA Elongation factor 1 alpha (Fragment).- *Mallomonas rasilis*.
Q8WTAO 9NEOP Elongation factor-1 alpha (Fragment).- *Pycnocentroides* sp. UMSP-New Zealand.
BAA11571 YSPEF1AB3 NID: - *Schizosaccharomyces pombe*
Q86MN1 PODCA Elongation factor 1-alpha.- *Podocoryne carnea*.
Q592D0 9HYME Elongation factor 1-alpha (Fragment).- *Eufriesia pulchra*.
Q5C193 SCHJA Hypothetical protein.- *Schistosoma japonicum* (Blood fluke).
Q8H9B0 9CARY Eukaryotic elongation factor 1A.- *Suaeda japonica*.
Q9GPV9 9CUCU Elongation factor 1 alpha (Fragment).- *Hylastes porculus*.
AAK54650 AF378368 NID: - *Coccidioides immitis*
Q6RIC9 9NEOP Elongation factor 1-alpha (Fragment).- *Rimisia miris*.
Q6E0N1 9ACAR Elongation factor 1 alpha (Fragment).- *Prodinychus* sp. AL7170.
Q9XY74 9CRUS Elongation factor 1-alpha (Fragment).- *Heteromysis formosa*.
Q8MZN7 9CUCU Elongation factor 1-alpha (Fragment).- *Xyleborus meritus*.
Q86GP6 9NEOP Elongation factor-1alpha (Fragment).- *Heteronympha merope*.
Q9XY78 9CRUS Elongation factor 1-alpha (Fragment).- *Cypridopsis vidua*.
Q5MXN9 9NEOP Elongation factor-1 alpha (Fragment).- *Taygetis sosis*.
Q5URD4 9HYME Elongation factor 1-alpha (Fragment).- *Pogonomymex lobatus*.
Q7YWH3 9CUCU Elongation factor 1-alpha (Fragment).- *Metamasius hemipterus* (silky cane weevil).
Q8T5M2 9ASIL Elongation factor-1 alpha (Fragment).- *Liancalus* sp. NCSU-95051152.
Q5EPW3 CHLAC Elongation factor 1 alpha (Fragment).- *Chlosyne acastus* (Sagebrush checkerspot).
Q9U9N2 9HEMI Elongation factor 1 alpha (Fragment).- *Cinara etsuhoe*.
Q7PDD0 PLAYO Translation elongation factor EF-1, subunit alpha.- *Plasmodium yoelii yoelii*.
K2C1 HUMAN Keratin, type II cytoskeletal 1 (Cytokeratin 1) (K1) (CK 1) (67 kDa cytokekeratin) (Hair alpha protei
Q5R2G2 9BASI Elongation factor-1 alpha (Fragment).- *Sporobolomyces elongatus*.
KRHU0 keratin 10, type I, cytoskeletal - human
Q6XQ70 SACPS Alcohol dehydrogenase 1 (EC 1.1.1.1).- *Saccharomyces pastorianus* (Lager yeast) (*Saccharomyces carls*
KRHUE keratin 14, type I, cytoskeletal - human
Q8T5N6 9ASIL Elongation factor-1 alpha (Fragment).- *Hoplopeza* sp. NCSU-95051124.
Q9U7M0 9CEST Elongation factor 1-a (Fragment).- *Eniochobothrium gracile*.
Q6FQA4 CANGA *Candida glabrata* strain CBS138 chromosome I complete sequence.- *Candida glabrata* (Yeast) (*Torulopsi*
Q9BLM8 TOXGO Hypothetical protein pdi (Protein disulfide isomerase precursor).- *Toxoplasma gondii*.
B53522 20k cyclophilin - *Toxoplasma gondii* (fragment)
Q5EA68 BOVIN Sec61 alpha form 1.- *Bos taurus* (Bovine).
Q6IG03 RAT Type II keratin Kb36.- *Rattus norvegicus* (Rat).
Q5I0C1 XENTR Hypothetical LOC496924.- *Xenopus tropicalis* (Western clawed frog) (*Silurana tropicalis*).
K2C1 MOUSE Keratin, type II cytoskeletal 1 (Cytokeratin 1) (67 kDa cytokekeratin).- *Mus musculus* (Mouse).
CAB22575 OSM14 PROTEIN (FRAGMENT).- *Osmunda cinnamomea* (Cinnamon fern).
Q7Z9U4 9SACH Actin (Fragment).- *Kluyveromyces piceae*.
KRHUEA keratin 6a, type II - human

Q9NC09 9MOLL Actin (Fragment).- Mastigoteuthis magna.
Q7Z9T1 9SACH Actin (Fragment).- Saccharomyces bulderi.
Q6XQ77 SACBA Alcohol dehydrogenase 2 (EC 1.1.1.1).- Saccharomyces
bayanus (Yeast) (Saccharomyces uvarum).
AAD30129 AF136344 NID: - Toxoplasma gondii
Q910C1 RANCA Larva-specific keratin RLK.- Rana catesbeiana (Bull
frog).
CAA24598 SCACT2 NID: - Saccharomyces cerevisiae
Q5UEM4 NEOCA Protein disulfide isomerase precursor (EC 5.3.4.1).-
Neospora caninum.
Q6IG00 RAT Type II keratin Kb4.- Rattus norvegicus (Rat).
Q86L89 9CUCU Elongation factor-1 alpha (Fragment).- Disonycha
conjuncta.
Q7S0L8 NEUCR Predicted protein.- Neurospora crassa.
Q86L79 9CUCU Elongation factor-1 alpha (Fragment).- Asphaera
auripennis.