Discovery of Benzamides and Triarylimidazoles Active against *Plasmodium falciparum* via Haemozoin Inhibition: High Throughput Screening, Synthesis and Structure-Activity Relationships

KATHRYN J. WICHT

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Chemistry

UNIVERSITY OF CAPE TOWN

April 2015

Supervisors: Professor Timothy J. Egan and Professor Roger Hunter

DECLARATION

I, Kathryn J. Wicht, know the meaning of plagiarism and declare that all of the work in this thesis, "**Discovery of benzamides and triarylimidazoles active against *Plasmodium falciparum* via haemozoin inhibition: high throughput screening, synthesis and structure-activity relationships**", save for that which is properly acknowledged, is my own, both in concept and execution, apart from the normal guidance of my supervisors.

I further declare that no part of the above-titled thesis has been or is being submitted for another degree at this or any other university.

I grant the University of Cape Town free license to reproduce this work, in whole or in part, for the purpose of research.

I hereby present this thesis for examination for the degree of Doctor of Philosophy (PhD) in Chemistry.

Signed at the University of Cape Town on 2nd April 2015.

Signed by candidate

signature removed

Kathryn J. Wicht

ABSTRACT

New antimalarials are desperately needed to overcome growing *P. falciparum* resistance to the current drugs. Successful quinoline-based drugs target haemozoin formation causing a cytotoxic accumulation of free haem (Fe(III)PPIX) in the parasite, a target which remains promising for future treatments. Much research has been undertaken on the quinoline antimalarials, which has led to several hypotheses of haemozoin inhibition and drug accumulation mechanisms, however, relatively few studies have been carried out for haemozoin antimalarials with alternate chemotypes. High throughput screening (HTS) can be used to identify novel scaffolds that inhibit β-haematin (βH - synthetic haemozoin) formation and which have favourable *P. falciparum* activities.

In this project, HTS has been carried out on 43,520 small, organic, drug-like compounds as part of a larger screen of 144,330 Vanderbilt University Institute of Chemical Biology (VU) chemical library compounds and 530 were found to be good inhibitors of βH relative to the chloroquine (CQ) and amodiaquine (AQ) controls. A further 171 compounds were found to inhibit parasite growth, showing improved hit rates from previous HTS efforts. Two scaffolds (A=benzamides and B=triarylimidazoles) were selected for further analysis, whereupon analogues were synthesised.

The benzamides (Scaffold 1) were synthesised from the corresponding acid chloride via conventional amide coupling. SAR analysis revealed that these planar molecules can π-π interact optimally with Fe(III)PPIX to inhibit βH, and the two most potent analogues, both pyridylbenzamides, showed *in vitro* parasite IC_{50}s of 0.7 ± 0.2 μM and 0.6 ± 0.1 μM against the CQ-sensitive NF54 strain. Amine side chains were introduced to increase basicity, an aspect that is believed to promote pH trapping in the acidic DV of the parasite. The predicted cellular accumulation ratios (CARs) of the new analogue and the quinoline antimalarials, CQ and AQ, were all ~7000 based on theoretical or measured pK_a values, whereas the pyridylbenzamides were predicted not to accumulate via pH trapping. Conversely, accumulation studies via the inoculum effect revealed experimental CARs of ~1000, ~6000 and ~100,000 for the basic derivative, non-basic pyridylbenzamides and quinoline antimalarials, CQ and AQ, respectively, demonstrating a lack of correlation

between the theoretical accumulation due to pH trapping and the experimental CARs. Fe(III)PPIX association data for the benzamides revealed relatively weak Fe(III)PPIX binding (logK of ~3.2), since the benzamide-haem interaction is exclusively via π-π interactions, compared to the quinolines which can form cation-π interactions (logK of 5.5 and 5.3 for CQ and quinidine respectively). Selected analogues were shown to inhibit haemozoin formation within *P. falciparum* cell cultures, causing unusually high levels of free haem above 50% at 2.5 times the IC_{50} compared with that for CQ of only 15%. These data suggested that haem binding may be primarily responsible for drug accumulation.

Triarylimidazole derivatives (Scaffold 2) were synthesised via a condensation with the appropriate aldehyde, benzil and ammonium acetate using literature procedures. The minimum pharmacophore for βH inhibition was identified as a triaryl-substituted imidazole with either a 2-pyridyl or a substituted 2-phenyl ring. However, SAR analysis for parasite growth inhibition showed that activity is highly dependent on the substitution pattern of the 2-phenyl ring, whereby derivatives with 3,5-methoxy-4-hydroxy substituents were significantly more potent, possibly as a result of improved solubility. Analogues from this chemotype were also shown to cause high levels of free haem, >50% at 2.5 times the IC_{50}.

The structures and activities of the synthesised derivatives together with the VU HTS data and previous GlaxoSmithKline (GSK) and St Jude Children's Research Hospital HTS data were used in training sets to create and optimise activity models based on Bayesian statistics, in order to predict the probability for inhibitory activity of other compounds as βH and parasite-growth inhibitors at a specified IC_{50} cut-off. The models for βH inhibition (100 μM IC_{50} cut-off) and parasite growth inhibition (2 μM IC_{50} cut-off) were used to identify six novel βH-inhibiting and two potent biologically-active compounds (27-fold and 100-fold enrichment respectively over random screening). It also predicted all six of the known βH antimalarials to lie within the top 2.5% of all 1510 FDA-approved drugs when ranked by the likelihood of being biologically active using Bayesian scores. Future activity results with additional compounds can be used to update the model. This continually updated model may be employed in virtual screening of other compound libraries, possibly even those generated *in silico*. This approach has potential for prioritising future libraries for more efficient screening, and for determining the most appropriate analogues for synthesis.

PUBLICATIONS AND CONFERENCE PROCEEDINGS

PUBLICATIONS

1. Sandlin, R. D.; Fong, K. Y.; Wicht, K. J.; Carrell, H. M.; Egan, T. J.; Wright, D. W. Identification of β-hematin inhibitors in a high-throughput screening effort reveals scaffolds with *in vitro* antimalarial activity, *Int. J. Parasitol.*, **2014**, 4, 316.

2. Wicht, K. J.; Combrinck, J. M.; Smith, P. J.; Egan, T. J. Bayesian models trained with HTS data for predicting β-haematin inhibition and *in vitro* antimalarial activity, *Bioorg. Med. Chem.*, **2015**, In Press, http://dx.doi.org/10.1016/j.bmc.2014.12.020.

CONFERENCE PROCEEDINGS

2011 International Conference on Biological Inorganic Chemistry (ICBIC 15), University of British Columbia, Vancouver, Canada.
Poster: Kathryn J. Wicht, Rebecca D. Sandlin, Timothy J. Egan, David W. Wright, High throughput screening for novel antimalarial drugs by a detergent mediated assay for β-haematin inhibitors.

2011 SACI Young Chemists' Symposium, University of the Western Cape, Cape Town, RSA.
Poster: Kathryn J. Wicht, Rebecca D. Sandlin, Timothy J. Egan, David W. Wright, High throughput screening for novel antimalarial drugs by a detergent mediated assay for β-haematin inhibitors.

2012 H3-D Symposium, Vineyard Hotel, Newlands, Cape Town, RSA.
Poster: Kathryn J. Wicht, Timothy J. Egan, Roger Hunter, David W. Wright, Novel antimalarial compounds from derivatives of chemotypes discovered via high throughput screening.

2013 International Conference on Biological Inorganic Chemistry (ICBIC 16), Grenoble, France.
Poster: Kathryn J. Wicht, Timothy J. Egan, Roger Hunter, David W. Wright, Strategies for improving the antimalarial activity of novel β-haematin inhibiting compounds.

2013 SACI Young Chemists' Symposium, University of Stellenbosch, Stellenbosch, RSA.
Oral Presentation: Kathryn J. Wicht, Timothy J. Egan, Roger Hunter, Strategies for improving the antimalarial activity of novel β-haematin inhibiting compounds.

2014 SACI Young Chemists' Symposium, University of Cape Town, Rondebosch, Cape Town, RSA.
Oral Presentation: Kathryn J. Wicht, Timothy J. Egan, Bayesian models trained with HTS data for predicting β-haematin inhibition and antimalarial activity.

2014 The SACI-ACS Bi-national Organic Chemistry Conference (BOCC) incorporating the 13th Franck Warren Conference, Stellenbosch, RSA.
Poster: Kathryn J. Wicht, Timothy J. Egan, Roger Hunter, Synthesis and SAR analysis of derivatives of novel antimalarial scaffolds and prediction of potentially active compounds using Bayesian models trained with HTS data.

COLLABORATIONS AND OUTSOURCED EXPERIMENTS

High throughput screening (HTS)

Screening was performed at Vanderbilt University (VU), Nashville, TN, USA in 2011 as part of a collaborative project between the research groups of T. J. Egan (University of Cape Town - UCT) and D. W. Wright (VU). A total of 43,520 compounds were screened using the NP-40 β-haematin inhibition assay by the author of this thesis and IC_{50} values were determined for approximately 700 compounds at the original cut-off of ≥50% inhibition. The data presented in Chapter Two includes both the data obtained by the author as well as the activity data acquired by R. D. Sandlin and H. M. Carrell. Data analyses were carried out by the author using the complete set of HTS results from all 144,330 library compounds.

Plasmodium falciparum (*P. falciparum*) cell culturing was carried out at VU by the author of this thesis to provide parasites for the *in vitro* SYBR green I-based fluorescence assays, which were performed on hit compounds from the primary βH inhibition screen by R. D. Sandlin.

Cell fractionation studies

Cell fractionation studies to measure free haem and haemozoin levels in *P. falciparum* cell cultures were performed on HTS hit compounds by K. Y. Fong at VU. Similar experiments on synthesised analogues were performed and analysed by J. M. Combrinck in the Division of Pharmacology, Department of Medicine at the UCT Medical School.

P. falciparum sensitivity assays and accumulation studies

Synthesised analogues of HTS hit compounds were sent to the Division of Pharmacology, Department of Medicine at the UCT Medical School for testing against the NF54 strain of the parasite. Testing against the D10 strain of the parasite was carried out by D. Taylor, J. M. Combrinck or the author of this thesis. D10 cell culturing and all cellular accumulation studies, which involved multiple D10 sensitivity assays at varying parasitaemias, were performed by the author of this thesis.

All other work in this thesis was carried out by the author.

ACKNOWLEDGEMENTS

Many people contributed to this project and/or provided me with a great deal of support. Much appreciation and numerous thanks are due to the following people:

Professors and supervisors:

- Firstly, to my supervisor, **Prof. Timothy Egan**, thank you for challenging me, sharing your extensive knowledge and teaching me to think independently, for your open-door policy, problem-solving skills and your consistent rationality, reliability and trustworthiness. I am so appreciative of all the time you dedicated to this project and for the opportunities with which you have provided me.
- My co-supervisor, **Prof. Roger Hunter**, for granting me entry into the world of organic synthesis. Thank you for adding a deeper level of mindfulness to my work and my understanding of molecular reactions. I will always remember and be inspired by your passion for organic chemistry and your original teaching style.
- **Prof. David Wright** at Vanderbilt University (VU), for the opportunity to work in an international research lab and for encouraging my thinking about the bigger issues of the planet.
- **Prof. Peter Smith** at the Division of Pharmacology, UCT Medical School, for allowing me full access to the parasite culture laboratory as well as all materials with which to carry out my experiments.

Members of the Egan Bioinorganic Research Group (The Haem Team):

- **John Woodland**, for your very valuable friendship on our eight-year-long academic journey together. Thank you for the countless motivational discussions and shared experiences which contributed significantly to my well-being.
- **Dr David Kuter**, for help with *in silico* energy calculations, for your mentorship and guidance since my honours projects, for encouraging me to join the group and for always being willing to thoroughly explain or discuss any problem, at any time.
- Past and present haem team members, in particular **Dr Melvin Ambele**, **Aneesa Omar**, **Roxanne Mohunlal**, **Fabrizio L'Abbate** and **Nikki Dare** for being responsible lab colleagues, and creating a cheerful work environment.

Members of other research groups:

- **Dr Katherine de Villiers**, for laying out the path on which I could attempt to follow, and her research group in Stellenbosch, particularly **Sharné Fitzroy** and **Ronel Müller** for collaboration on the NIH project.
- **Dr Sophie Rees-Jones**, for the wealth of organisational knowledge and lab experience and for always being happy to help me locate the chemical or equipment that I required for my synthesis.
- Past and present friends in the Chemistry Department, particularly **Taigh Anderson**, **Wade Petersen**, **James Biwi**, **Mandla Mabunda** and **Richard Payne** for all the sociable conversations over the years and for being great companions in the lab.

Other significant contributors:

- **Dr Rebecca Sandlin**, for HTS assays and cell culture training at VU and for a rewarding international collaboration.
- **Jill Combrinck**, for malaria parasite LDH assay training. Also, for always being willing to test my compounds and to share relevant data or expertise on everything parasite related.
- Members of the Division of Pharmacology, UCT Medical School, for testing my compounds on parasites, offering advice and helping me feel at home in their work-environment.
- **Peter Roberts** and **Hiten Fletcher** for collection of NMR and mass spectra.
- The Department of Chemistry at UCT, in particular **Prof. Susan Bourne** (HoD), **Elaine Rutherfoord-Jones**, **Deirdre Brooks**, **Kevin Willis**, **Prof. Neil Ravenscroft** and **Prof. Kelly Chibale**.
- The **National Institutes of Health** (NIH) for funding under grant #R01AI083145 for HTS and #R01AI110329 for all follow-up work. Further thanks to the **University of Cape Town** for additional financial assistance.

- Finally, to my Dad, **Nordan Wicht**, for giving me the best education possible and for always believing in me; my Mum, **Sharon Wicht**, for your infinite love, support, advice and strength and for always being so proud of me; and to **Sean Bethell**, for patiently sharing my PhD journey with me, and for being the loving, fun, considerate and dependable person to whom I could return home each day.

LIST OF ABBREVIATIONS

ACTs	Artemisinin-based combination therapies
AQ	Amodiaquine
Ato/Pg	Atovaquone/Proguanil
CAR	Cellular accumulation ratio
CHAPS	3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate
CD	Cinchonidine
CN	Cinchonine
CMC	Critical micelle concentration
CQ	Chloroquine
DCM	Dichloromethane
DMSO	Dimethyl sulfoxide
DV	Digestive vacuole
EtOAc	Ethyl acetate
EtOH	Ethanol
Et_3N	Triethylamine
Eq.	Equivalents or Equation
FDA	Food and Drug Administration
Fe(II)PPIX	Iron(II) protoporphyrin IX
Fe(III)PPIX	Iron(III) protoporphyrin IX
FPSA	Fractional polar surface area
GSK	GlaxoSmithKline
H2O.hyd	Mean water of hydration
Halo or HF	Halofantrine
HAP	Histo-aspartic protease
Hb	Haemoglobin
HBD	Hydrogen bond donor
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HTS	High throughput screening
IC_{50}	50% Inhibitory concentration
IR	Infrared
K_a	Association constant
LDH	Lactate dehydrogenase
MF	Mefloquine
MeOH	Methanol
MW	Molecular weight
NLB	Neutral lipid blend
NMR	Nuclear magnetic resonance
NP-40	Nonidet-P 40
OU	Okayama University
PfCRT	*Plasmodium falciparum* CQ resistance transporter
Pfmdr	*Plasmodium falciparum* multi-drug resistant gene
pRBC	Parasitised red blood cell
PSA	Polar surface area
PXRD	Powder X-ray diffraction
PC	Principle component
QC	Quinacrine
QD	Quinidine
QN	Quinine
RBC	Red blood cell
ROC	Receiver operating characteristic
RT	Room temperature
S/P	Sulfadoxine/Pyrimethamine
SARs	Structure-activity relationships
SDS	Sodium dodecyl sulfate
TCAMS	Tres Cantos antimalarial set
TEM	Transmission electron microscopy
TFA	Trifluoroacetic acid
Triton™ X-100	Polyethylene glycol tert-octylphenyl ether
TWEEN® 20	Polyoxyethylenesorbitan monolaurate
TWEEN® 80	Polyethylene glycol sorbitan monooleate
UCT	University of Cape Town
VAR	Vacuolar accumulation ratio
VU	Vanderbilt University
WHO	World Health Organisation
βH	β-Haematin
#AR	Number of aromatic rings
#HBA	Number of hydrogen bond acceptors
#HBD	Number of hydrogen bond donors
#R	Number of rings
#RB	Number of rotatable bonds
3D	Three-dimensional

TABLE OF CONTENTS

Declaration i
Abstract ii
Publications and conference proceedings iv
Collaborations and outsourced experiments v
Acknowledgements vi
List of abbreviations vii
Table of contents viii

CHAPTER ONE
INTRODUCTION 1
1.1 Historical background 1
1.2 The current global malaria disease burden 1
1.3 The malaria parasite life cycle 2
1.4 Haemoglobin degradation 5
1.5 Malaria chemotherapy 6
1.5.1 Non-haemozoin inhibiting antimalarials 7
1.5.2 Quinoline and related antimalarials 10
1.6 Drug resistance 13
1.7 Mechanism of antimalarial drug action for the quinoline and related compounds 18
1.7.1 *P. falciparum* cell culturing and drug sensitivity assays 18
1.7.2 Nucleic acid and proteins as possible targets of chloroquine 20
1.7.3 Haemozoin inhibition 20
1.7.4 Accumulation of quinoline antimalarials in the DV 33
1.8 Haemozoin and β-haematin 36
1.8.1 The structure of haemozoin and β-haematin 36
1.8.2 The mechanism of haemozoin formation 39
1.8.3 The mechanism of haemozoin inhibition 41
1.8.4 Crystal structures of drug-haem complexes 42
1.8.5 βH inhibition assays 43
1.8.6 Detecting the formation of β-haematin 47
1.9 High throughput screening in antimalarial drug development 49
1.9.1 Phenotypic vs target based approaches 50
1.9.2 Recent antimalarial HTS efforts 51
1.9.3 The scope for novel βH inhibitors discovered via HTS 53
1.10 Aims and Objectives 56
1.10.1 Aims 56
1.10.2 Objectives 56

CHAPTER TWO

HIGH THROUGHPUT SCREENING FOR NOVEL β-HAEMATIN INHIBITORS 58

2.1 Introduction 58
2.1.1 Drift, Z' and edge effects in plate-based assays 58
2.1.2 Pilot screen for βH inhibitors and antimalarials 60
2.1.3 Specific goals 60
2.2 Results and discussion 61
2.2.1 Assay validation 61
2.2.2 Compound screening for βH inhibitors 62
2.2.3 Secondary screen of βH inhibitors for *in vitro* parasite activity 71
2.2.4 Scaffold analysis 77
2.2.4.1 Haemozoin target validation via cell fractionation 79
2.2.4.2 Phenyl and pyridyl benzamides (Scaffold 1) 80
2.2.4.3 Triarylimidazoles (Scaffold 2) 85
2.3 Summary and conclusions 87

CHAPTER THREE

SYNTHESIS 89

3.1 Introduction 89
3.1.1 Scaffold 1: Benzamides 89
3.1.1.1 Amide bond formation 89
3.1.1.2 Synthetic targets for the benzamide scaffold 93
3.1.2 Scaffold 2: Triarylimidazoles 95
3.1.2.1 Literature methods for the synthesis of triarylimidazoles 95
3.1.2.2 Synthetic targets for the triarylimidazole scaffold 96
3.2 Results and discussion 98
3.2.1 Scaffold 1: Benzamides 98
3.2.1.1 Synthesis of the monobenzamides 98
3.2.1.2 Synthesis of the dibenzamides 104
3.2.1.3 Synthesis of the tribenzamide 117
3.2.2 Scaffold 2: Triarylimidazoles 118
3.2.2.1 Synthesis of triarylimidazoles with unsubstituted B and C rings 118
3.2.2.2 Synthesis of triarylimidazoles with substituted B and C rings 120
3.2.2.3 Synthesis of nitro- and amine-substituted triarylimidazoles 126
3.3 Summary and conclusions 128

CHAPTER FOUR

BENZAMIDES: STRUCTURE-ACTIVITY RELATIONSHIPS

FOR β-HAEMATIN INHIBITION AND PARASITE ACTIVITY 130

4.1 Introduction 130
4.1.1 Specific goals 131
4.2 Results and discussion 132
4.2.1 β-haematin inhibitory activity of the benzamides 132
4.2.1.1 βH inhibition assay in 96-well plates 132
4.2.1.2 Evaluation of βH inhibition for the purchased benzamides 133
4.2.1.3 Evaluation of βH inhibition for the synthesised monobenzamides 135

4.2.1.4 Factors influencing βH inhibition of mono-benzamides. 137
4.2.1.5 Evaluation of βH inhibition for the synthesised di- and tri-benzamides 140
4.2.1.6 Proposed mechanism of βH inhibition for the dibenzamides 143
4.2.2 Activity of benzamides against *P. falciparum* in culture 145
4.2.2.1 Activities of the monobenzamides 145
4.2.2.2 Activities of the di- and tri-benzamides 146
4.2.2.3 Target validation for the benzamides 150
4.2.3 Attempts to improve activity through pH trapping 153
4.2.3.1 Design of a basic analogue and pH trapping predictions 153
4.2.3.2 Target validation for derivative 7 155
4.2.3.3 The inoculum effect and cellular accumulation 156
4.2.4 Possible role of Fe(III)PPIX association in cellular accumulation 161
4.3 Summary and conclusions 167

CHAPTER FIVE
TRIARYLIMIDAZOLES: STRUCTURE-ACTIVITY RELATIONSHIPS
FOR β-HAEMATIN INHIBITION AND PARASITE ACTIVITY 168
5.1 Introduction 168
5.1.1 Specific goals 169
5.2 Results and discussion 170
5.2.1 β-Haematin inhibition activity for the triarylimidazoles 170
5.2.2 Parasite growth inhibition activity 173
5.2.3 Target validation for the triarylimidazoles 175
5.2.4 Correlation analysis for triarylimidazole SARs 176
5.3 Summary and conclusions 179

CHAPTER SIX
BAYESIAN MODELS FOR PREDICTING
β-HAEMATIN INHIBITION AND ANTIMALARIAL ACTIVITY 180
6.1 Introduction 180
6.1.1 Bayesian probability 180
6.1.1.1 Bayes' theorem 180
6.1.1.2 Bayesian classification in Discovery Studio 181
6.1.1.3 The Laplacian-adjusted probability estimate 182
6.1.2 Available antimalarial datasets 183
6.1.3 Specific goals 185
6.2 Results and discussion 186
6.2.1 Optimisation of Bayesian models 186
6.2.1.1 Model for predicting inhibition of βH formation 186
6.2.1.2 Model for predicting parasite activity 187
6.2.1.3 Comparison of molecular descriptors for βH and parasite activity 189
6.2.2 Chemical space analysis 191
6.2.3 FDA approved compounds for model validation 193
6.2.4 Prioritization of a commercial library 197
6.3 Summary and conclusions 203

CHAPTER SEVEN

CONCLUSIONS AND FUTURE WORK 205

7.1 Conclusions 205

7.2 Future work 211

CHAPTER EIGHT

EXPERIMENTAL METHODS 214

8.1 Physical methods, assays and cell culture protocols 214

8.1.1 Materials and general procedures 214

8.1.2 High throughput screening methods 215

8.1.2.1 Manual β-haematin formation assay preparation 215

8.1.2.2 Manual concentration dose response experiments 215

8.1.2.3 Verification of drift, Z' and edge effects 216

8.1.2.4 HTS Procedures 216

8.1.2.5 *Plasmodium falciparum* SYBR green I-based fluorescence assay 217

8.1.3 β-Haematin assays, titrations and physical analysis methods 217

8.1.3.1 Detergent mediated assay for β-haematin inhibition 217

8.1.3.2 Derivation of the vacuolar accumulation ratio (VAR) equation 218

8.1.3.3 Haem binding 219

8.1.3.4 Correlation analysis 220

8.1.4 *P. falciparum* cell culturing methods 220

8.1.4.1 LDH malaria parasite survival assay 220

8.1.4.2 Cell fractionation 221

8.1.4.3 The Inoculum effect 221

8.1.5 Bayesian modelling methods 223

8.1.5.1 Training set data 223

8.1.5.2 Comparing the chemical space of libraries 223

8.1.5.3 Building Bayesian models 223

8.1.5.4 Model validation using external datasets 224

8.2 Synthetic methods 225

8.2.1 Materials and general procedures 225

8.2.2 Synthesis, purification and characterisation 227

8.2.2.1 Benzamides 227

8.2.2.2 Triarylimidazoles 243

REFERENCES 250

CHAPTER ONE

INTRODUCTION

1.1 Historical background

Malaria is one of the most prevalent and oldest documented human infections, particularly in tropical regions.[1] The signs and symptoms of the disease, including headaches, chills, intermittent fevers and enlarged spleens were described in early writings from Egypt, India and China.[2] While initially believed to be caused by evil spirits, Hippocrates described the clinical symptoms of malaria in 400 BC, which were attributed to the vapour, mist and other "bad air" ascending from swamps and marshes. As a result, the disease was named "Mal'aria" by the Romans.[2,3,4] It was only 2500 years later that scientific studies could be carried out, after Charles Louis Alfonse Lavern discovered spherical parasites inside fresh red blood cells.[5] In addition, a black pigment similar to that observed more than a century before in the brain and spleen of malaria victims, was detected within the malaria parasite.[6] Mosquitoes were subsequently identified as the vector by Ronald Ross in 1897, when parasites containing the same black pigment, were found in the mosquito's gut.[7] Since then, approximately 200 mammalian, avian and reptile species of parasite, belonging to the genus *Plasmodium,* have been reported.[8,9] Of the five that infect man, *P. falciparum* is the most pathogenic, prevalent and deadly.[1,2]

1.2 The current global malaria disease burden

The latest malaria report from the World Health Organisation (WHO),[10] which summarises data from 97 malaria-endemic countries, describes the expansive deployment of artemisinin-based combination therapies (ACTs) in 2013 (*Section 1.5.1*), leading to a 47% decrease in global mortality rates from 2010 to 2013. Despite this decline in malaria-related deaths, 430,000 annual child mortality cases were recently reported from Africa alone. These deaths accounted for 74% of the 584,000 total global mortality figures. Furthermore, 198 million cases of the disease occurred globally and about 60 million children with malaria

did not receive treatment with ACTs. Biological, funding and systemic challenges, together with emerging drug resistance, continue to threaten the current treatments and the WHO encourages researchers to act with urgency. The global distribution of the disease epidemic is detailed in *Figure 1.1*.

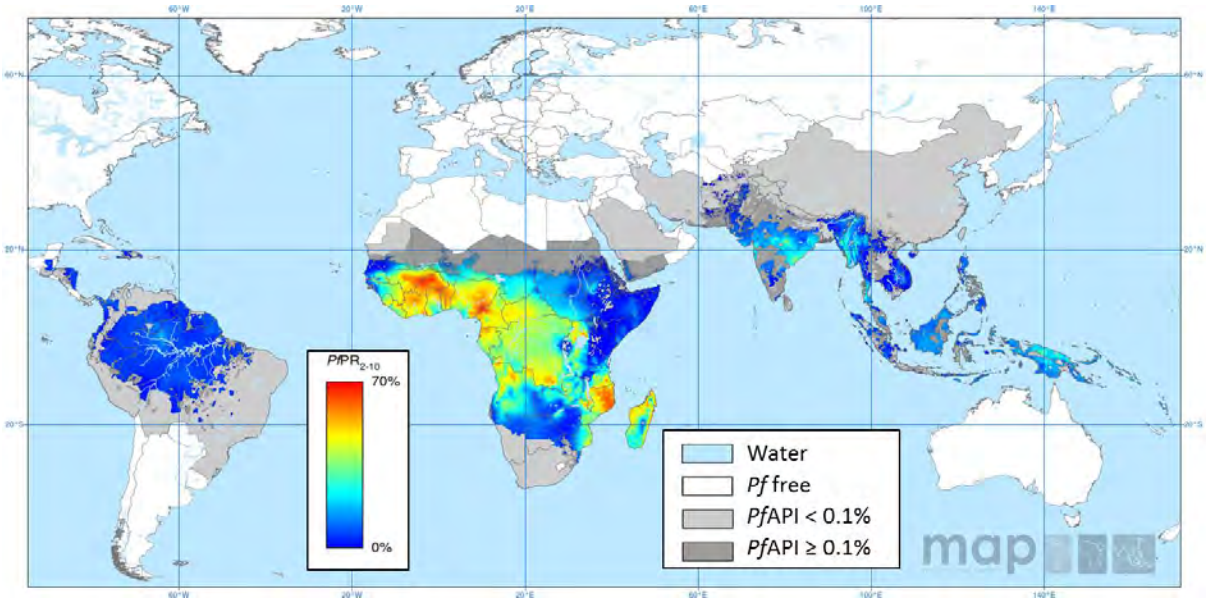


Figure 1.1 Spatial distribution of the *P. falciparum* malaria endemicity in 2010. The parasite rate is age-standardized to the two to ten year age-range (*Pf*PR2-10) for every 5×5km pixel, to account for the fact that malaria is more prevalent in children. The percentage population infected is indicated by the blue to red colour scale. Grey regions indicate the *P. falciparum* annual parasite incidence (*Pf*API). Reproduced from Gething *et al*.[11]

1.3 The malaria parasite life cycle

By 1900, the basic features of the life cycle of the parasite were recognized, but it was only in the 1950s that all three stages of human malaria parasites were comprehensively described.[2] The five species that infect humans have very similar life cycles which comprise the mosquito, liver and blood stages, as shown in *Figure 1.2* for *P. falciparum*. The bite of a female *Anopheles* mosquito inoculates the human host with about ten motile sporozoites which transit through the human dermis to access blood vessels and then, upon reaching the liver, rapidly through Kupffer cells, the liver macrophages, to invade parenchymal hepatocyte cells. Here, the liver stage of the life cycle occurs, during which the sporozoites

multiply asexually to form liver schizoints over approximately one week. This causes cell rupture and over 100,000 merozoites are released into the blood stream.[12] Rapid invasion of erythrocytes initiates the asexual blood stage, which occurs in 48 h cycles and causes the typical malaria symptoms. The blood cycle begins as rings within a parasitophorous vacuole. Further growth development into trophozoites and then blood schizonts is fuelled by host Hb degradation. At this stage, the formation of the black "malaria pigment" takes place. Following maturation of the schizont, the red blood cell (RBC) bursts open and merozoites flow into the bloodstream where they invade new erythrocytes to reinitiate the blood cycle and increase parasitemia (percent infected RBC) levels. This process continues until intervention through treatment, host response regulation or death of the victim. During this stage, gametocytogenesis also takes place, forming male and female gametocytes over approximately ten days. These have a characteristic crescent shape and circulate in the blood for up to four days until being taken up by a mosquito upon feeding. Following maturation of a gametocyte into a gamete within the gut of the mosquito, fertilisation occurs to form a motile zygote (ookinete), which penetrates the gut membrane and becomes an oocyst. Multiplication of the oocyst produces sporozoites which then travel to the mosquito's salivary glands, ready to begin the *P. falciparum* life cycle again.[2,3,13]

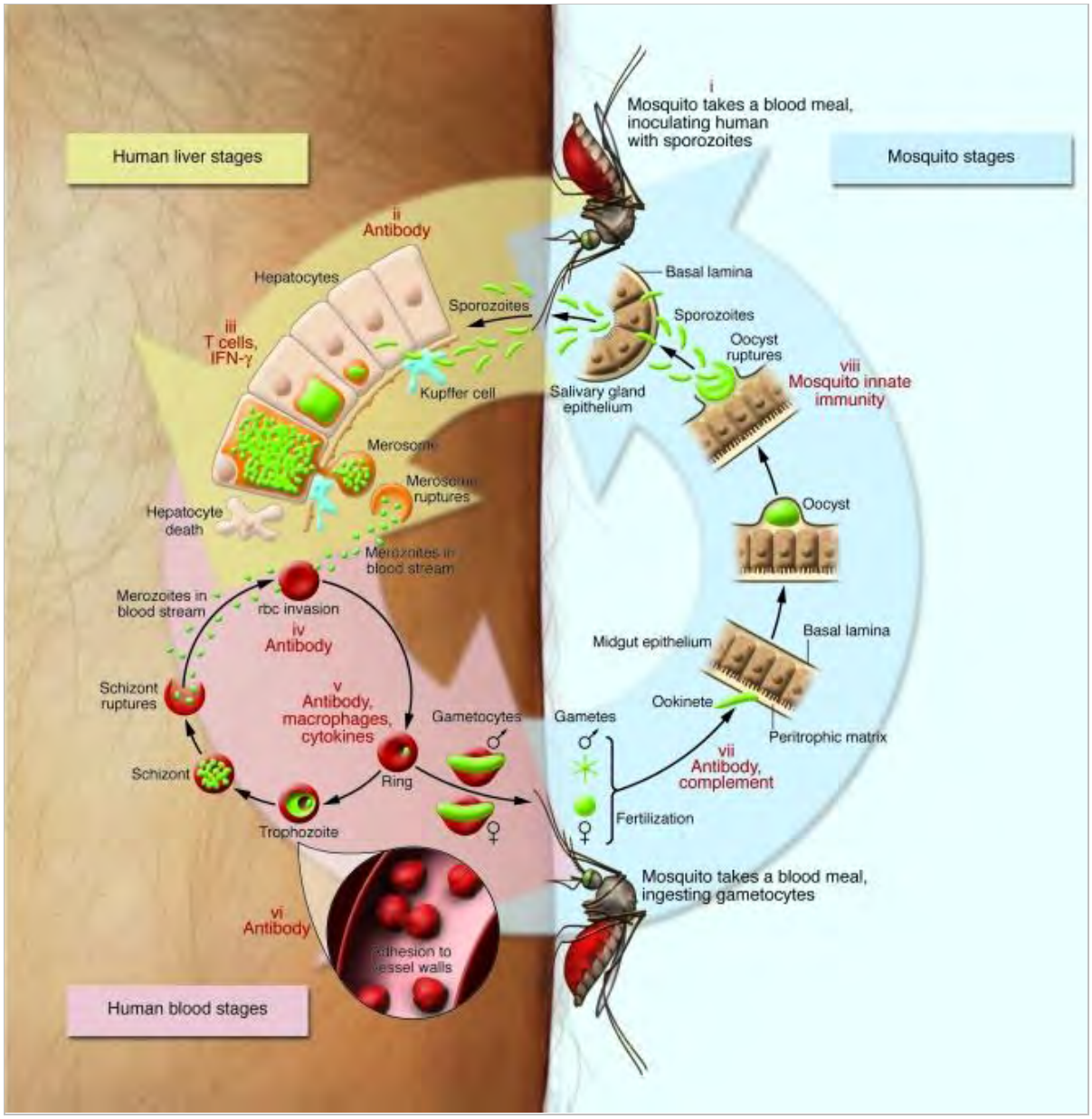


Figure 1.2 The life cycle of the *Plasmodium* malaria parasite. Copyright © 2008 American Society for Clinical Investigation, Greenwood *et al.*[13]

1.4 Haemoglobin degradation

During the blood stage of the life cycle, the malaria parasite ingests the major cytosolic host cell protein, haemoglobin (Hb). This is initiated by endocytosis via the cytostome (cell mouth), which extends between the parasite and red cell cytoplasm allowing the trophozoite to feed on Hb from the host cytoplasm. Vesicles containing Hb are then transported to the digestive vacuole (DV), an acidic secondary lysosome with a pH of 4.8-5.2.[14] Here, the Hb is degraded in order to supply the parasite with the necessary amino acids for its growth and maturation, since neither their acquirement via *de novo* synthesis nor exogenous uptake are possible.[15] In addition, the parasite requires space to grow and divide so that it does not exceed the lytic volume of the host red cell.[16] Indeed, one study showed that only ~16% of the 65% degraded host Hb was incorporated into the parasite proteins and hence, only a fraction of the acquired amino acids are used as nutrients.[17] Acid and alkaline proteases in the DV break down Hb into denatured globin, releasing a haem molecule as a by-product.[15] Further degradation of the globin by the aspartic protease enzymes plasmepsin I, II and IV and histo-aspartic protease (HAP), by the falcipain cysteine proteases and by the zinc protease, falcilysin provides peptides that are then hydrolysed to the vital amino acids, where a fraction are incorporated into parasite proteins.[18] Some studies indicate that up to 75% of erythrocyte Hb can be digested during a single blood cycle and as a result, the parasite has developed an efficient and specific pathway for Hb proteolysis.[19] In the process, the parasite faces a unique problem, since it is unable to degrade haem (Fe(II)PPIX) enzymatically, as it appears to lack a functional haem oxygenase. Instead Fe(II)PPIX is auto-oxidised by O_2 to Fe(III)PPIX, which is capable of lipid peroxidation and is therefore cytotoxic. In order to deal with the vast amounts of Fe(III)PPIX produced, a type of biomineralisation process takes place, whereby the toxic Fe(III)PPIX is converted into an inert and highly insoluble crystalline material known as haemozoin.[18,20] This process is shown in *Figure 1.3.*

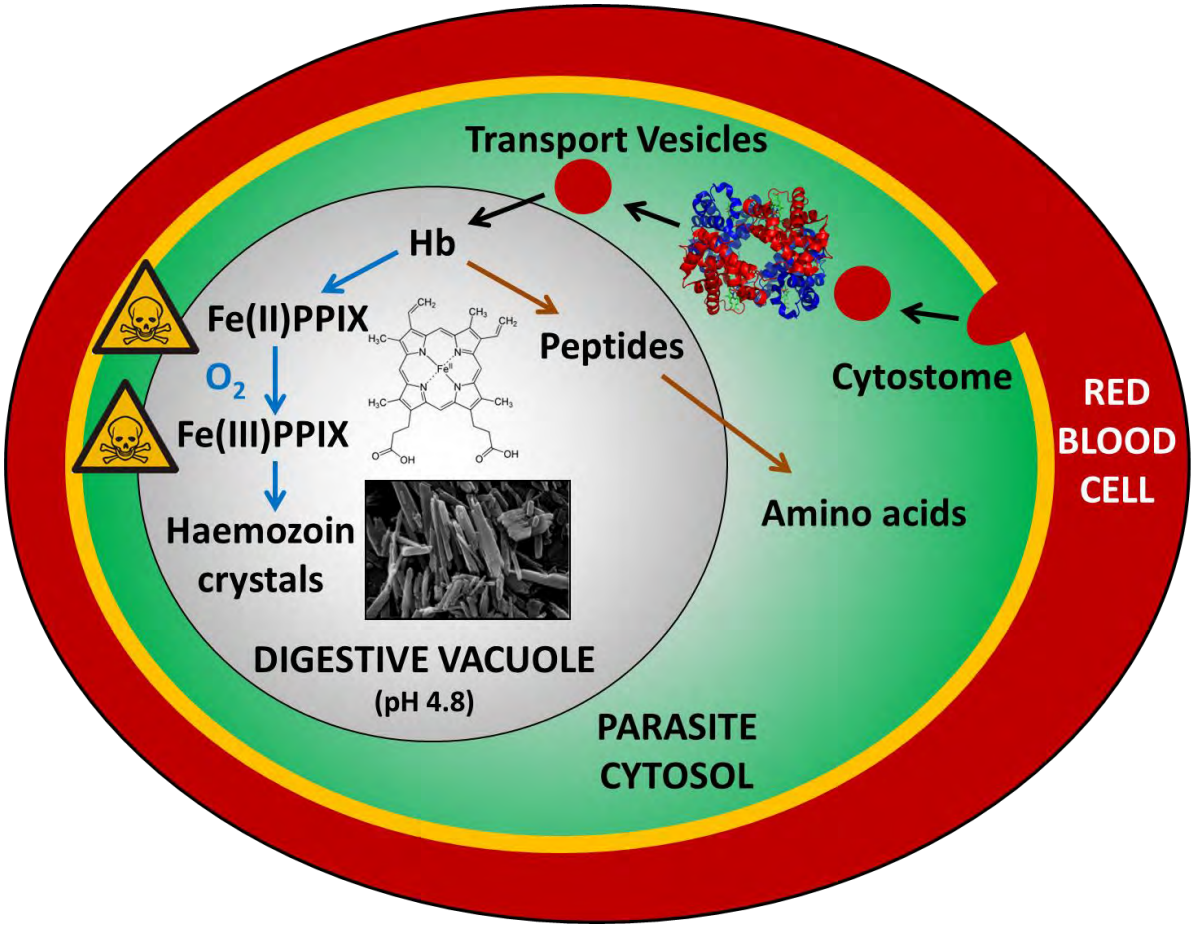


Figure 1.3 Haemoglobin (Hb) degradation in a parasitised red blood cell, resulting in amino acids for parasite growth and cytotoxic Fe(III)PPIX which is crystallised to form non-toxic haemozoin.

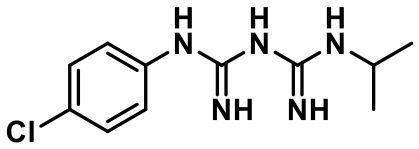
1.5 Malaria chemotherapy

Four major drug classes are currently used to treat malaria, namely, the antifolates, the artemisinin derivatives, the antimicrobials and the quinoline-related compounds. The majority of drugs from the quinoline class act during the red blood cell stage of the parasite life cycle and are believed to target the haemozoin formation process discussed in *Section 1.4.*

1.5.1 Non-haemozoin inhibiting antimalarials

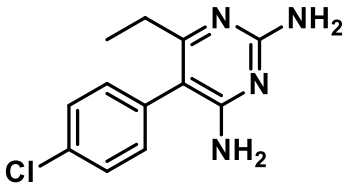
An intensive search for antimalarials in the United Kingdom during the Second World War identified proguanil (*Figure 1.4a*), an open pyrimidine derivative (biguanide), as having better activity than the well-established drug quinine (QN), although its action was slow against acute malaria and it appeared to induce drug resistance in some parasite strains.[1] Other pyrimidines, namely the 2,4-diaminopyrimidines were also shown to be parasite growth inhibitors, the most active of which was pyrimethamine (*Figure 1.4b*), developed in 1951 by Falco and Hitchings.[21] However, this compound readily induced resistance and as a result, pyrimethamine (P) was later used in combination with sulfadoxine (S), a sulfonamide antibiotic (*Figure 1.4c*), which together have a synergistic effect. These compounds target the essential folate biosynthesis pathway of the parasite by inhibiting the enzymes dihydrofolate reductase and the dihydropteroate synthesase.[22] As of 1996, proguanil (Pg) was administered in combination with atovaquone (Ato), an inhibitor of mitochondrial electron transport in the parasite (*Figure 1.4d*). Studies in the early 1990s showed that the Ato/Pg combination was significantly more effective than several different haemozoin-inhibiting drugs in Thailand, Gabon and Peru.[23] Ato, while possessing a favourable safety profile, induces an unacceptable rate of recrudescent parasitemia and resistance and is thus never given as a monotherapy.

a)



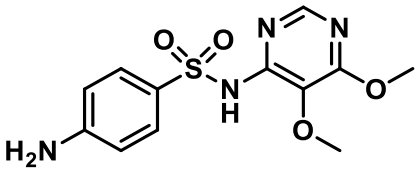
Proguanil

b)



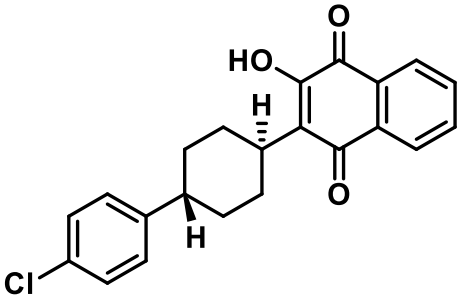
Pyrimethamine

c)



Sulfadoxine

d)



Atovaquone

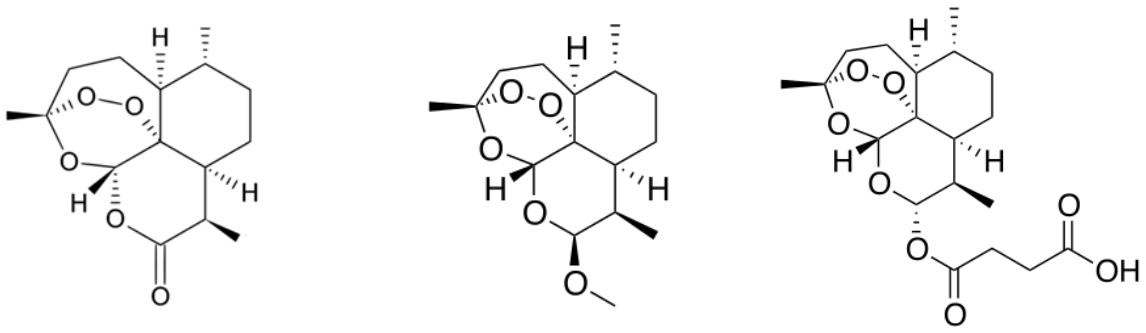
Figure 1.4 a,b,c) The antifolate drugs based on the pyrimidine chemotype which inhibit a,b) the dihydrofolate reductase enzyme and c) the the dihydropteroate synthesase enzyme and d) the antibiotic, given in combination with proguanil, which prevents parasite mitochondrial electron transport.

Currently, ACTs are the generally preferred treatment for malaria infections.[10] Artemisinin was extracted from the Chinese herb *Artemisia annua*, also known as sweet wormwood and was used as a powerful antimalarial in China for over two thousand years.[24] However, the active ingredient was only identified as artemisinin, an unstable sesquiterpine lactone with an unusual endoperoxide bridge, in the early 1970s, (*Figure 1.5a*).[25] It has been proposed that upon cleavage of this peroxide bond by reduction with Fe(II), reactive radicals which destroy the parasite, are released. As a result, this moiety has been demonstrated to be essential for the activity of artemisinin.[26] However, the mechanism of action of artemisinin remains controversial.[27] Other reports discuss the importance of carbon-centered radical formation and an electrophilic epoxide, a potent alkylating agent,[28] which specifically

alkylates malarial proteins but not those of intact erythrocytes[29]. Artemisinin also forms a covalent complex with haem and some describe the inhibition of Hb breakdown and haemozoin formation by artemisinin,[30] however the artemisinin-haem adduct has been shown to have weak parasite activity.[31] Furthermore, the release of toxic free iron upon haem degradation by artemisinin has been proposed, but increases in free iron have not been detected in artemisinin-treated parasites.[32] The artemisinin-haem adduct is therefore not believed to be related to drug activity mechanisms, especially since there is reportedly no influence by artemisinin on the amount of haemozoin formed within cultured parasites.[33] Other hypothesised drug targets include the sarco/endoplasmic reticulum Ca2+-ATPase (SERCA) pump of the parasite,[34] or the mitochondrial electron transport chain.[35] Artemisinin and its derivatives are highly selective for parasites, readily partitioning into membranes as a result of the hydrophobicity of the chemical scaffold. A 1000-fold increase in artemisinin concentration is required for comparable toxicity in mammalian cells.[32]

In the 1980s, further developments led to the synthesis of more stable derivatives such as artemether, a methyl-ether analogue (*Figure 1.5b*) and artesunate, a succinic acid derivative with improved aqueous solubility (*Figure 1.5c*). However, these were identified as pro-drugs with dihydroartemisinin as the active metabolite.[25] Artemisinin has been shown to be active against all of the asexual blood stages of the parasite, unlike the 4-aminoquinoline drugs.[36] According to the WHO, by the end of 2013, ACTs had been adopted as national policy for first-line treatment of malaria in 79 of 88 countries where *P. falciparum* is endemic. However, the quinoline-based drug, CQ, discussed in *Section 1.5.2* was still being used in nine Central American and Caribbean countries where it remains efficacious.[10]

a) b) c)



Artemisinin **Artemether** **Artesunate**

Figure 1.5 Structures of the artemisinin derivatives.

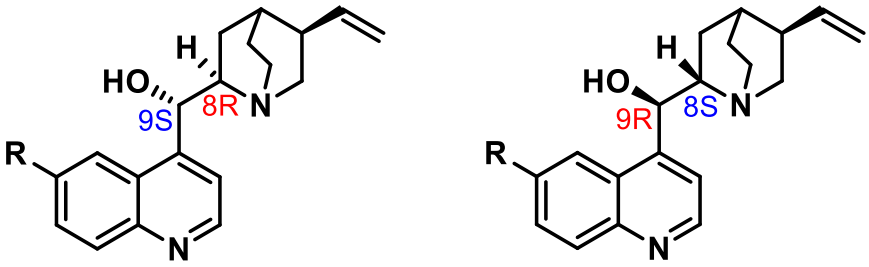
1.5.2 Quinoline and related antimalarials

For at least 2000 years, the roots, leaves and flowers of many different plants were used in an attempt to treat the fevers associated with malaria, most of them having no great effect. The first potent chemotherapy for disease control was identified in the 1630s as a result of European contact with Spanish priests, who brought the bark of "fever trees" from the Peruvian Andes. The bark was given the latin name cinchona after the Countess of Chinchon was supposedly cured of a fever following administration of an infusion of the bark. Powdered preparation of the crude cinchona bark were popular for 200 years before isolation of the two cinchona alkaloids, quinine (QN) and cinchonine (CN) in 1820 by the French chemists Pierre Pelletier and Joseph Caventou. It was not long before the other two alkaloids, quinidine (QD) and cinchonidine (CD) were isolated too and factories for manufacture of salts of QN were established. However, the popularity of the new drug led to destruction of the native chinchona forests in Peru and tree plantations became essential to meet the growing demands. By the 1930s, 10 million kilograms of cinchona bark were produced per year by Dutch plantations in Java, providing 97% of the world's QN requirements.[1,37]

The four alkaloids isolated from the cinchona bark are chemically related via their quinoline ring and with respect to symmetry at the chiral centres C-8 and C-9. As shown in *Figure 1.6*, QN (9R, 8S) and QD (9S, 8R) are diastereoisomers with a methoxy substituent at the 6-

postion of the quinoline ring. Cinchonine (CN, 9S, 8R) and cinchonidine (CD, 9R, 8S) are analogues of QD and QN respectively, containing hydrogen in place of the methoxy moiety.[37]

a) b)



R = OMe (Quinidine)
R = H (Cinchonine)

R = OMe (Quinine)
R = H (Cinchonidine)

Figure 1.6 The structures of the cinchona alkaloids (a) QD and CN with the same stereochemistry and the corresponding diastereoisomers (b) QN and CD.

The synthesis of artificial QN analogues was pioneered by German chemists during the First World War when they found themselves isolated from India and Java, the world's main suppliers of QN alkaloids. In addition, Ehrlich observed the therapeutic effect of methylene blue on malaria patients, which together with development of a technique for assessing the effectiveness of compounds, initiated the search for synthetic antimalarials.[1] The best of the synthesised compounds discovered at this early time was quinacrine (QC, *Figure 1.7a*), discovered in 1933. However, a couple of years later, the German chemist, H. Andersag synthesised a 4-aminoquinoline with comparable activity, called Resochin, while working in the Elberfeld laboratories of the Bayer I.G. Farbenindustrie. Initial reported toxicity tests indicated that Resochin was unsuitable for use in humans and its subsequent abandonment became known as the "Resochin error". This was revealed after screening of around 16,000 compounds in birds, to assess the activity of synthetic compounds against avian malarias, was carried out in the U.S. from 1943. Following toxicological and pharmacological studies in smaller animals, approximately 80 compounds were selected for testing on humans. The venture exposed a new compound, chloroquine (*Figure 1.7b*), which in 1946 was subjected

to U.S. clinical trials in 5000 individuals to study and describe the subsequent symptoms and side effects. Chloroquine (CQ) was found to be structurally identical to Resochin and superior to all previous medications, resulting in 137,250,000 tablets being purchased per year by 1962.[38] Several other synthetic quinoline-based antimalarials (*Figure 1.7c-e*) including primaquine (PQ), amodiaquine (AQ) and mefloquine (MF) had been discovered by the late 1970s. The 4-quinolinemethanol MF was identified during a screen of 250,000 compounds, which was conducted as a result of the emergence of CQ resistance in the 1960s. This screen revealed four primary scaffolds, namely the 4-quinolinemethanols and the non-quinoline-based chemotypes: 9-phenanthrenemethanols, 2,4-diaminoquinazolines and 2,4-diaminotriazines.[1] Despite efforts to improve chemotherapies, CQ turned out to be the drug of choice for over four decades and was a critical part of the world-wide malaria eradication program under the WHO.[38]

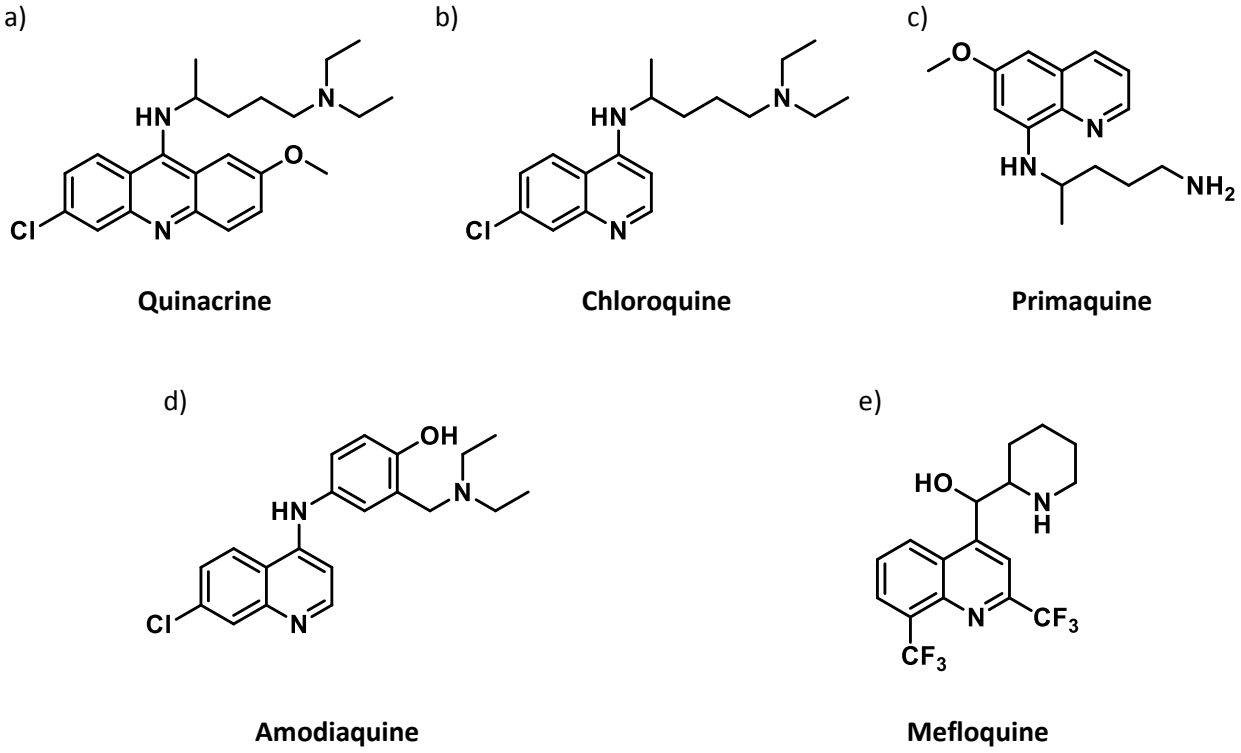


Figure 1.7 The synthetic acridine and quinoline antimalarials discovered before 1980, of which CQ was the most widely used for prophylaxis and treatment.

1.6 Drug resistance

Antimalarial drugs play a crucial role in controlling the spread and rate of *P. falciparum* infections. Between 2000 and 2013, a 47% global decline in the number of malaria-related deaths was observed, according to the WHO, along with a major expansion in the distribution of ACTs.[10] However, there appears to be a correlation between the development of resistance to antimalarials with the extent of their clinical use. As a result, there is a continuous need for novel compounds, particularly for use in combination therapies in order to combat this resistance.

The first case of *P. falciparum* CQ-resistance was reported in 1960 by American workers Young and Moore when a dose of CQ failed to treat a *P. falciparum* infection from Colombia.[1] Shortly after, further reports of unresponsive malaria infections to the usual doses of CQ appeared from Brazil and Venezuela. It is believed that evolution to the resistant phenotype required mutations in several genes, since development of resistance to CQ was relatively slow, despite the extensive use of the drug. Nevertheless, CQ resistance was soon established in south-east Asia and later spread to the other continents (*Figure 1.8*). This resulted in CQ, once the first-line malaria treatment, being declared ineffective in most parts of the world.[14,39] Even in the absence of drug pressure, the resistance phenotype persisted and appeared to be stable.[40] Although reversion to CQ-sensitivity has been observed in some areas, such as Malawi.[41]

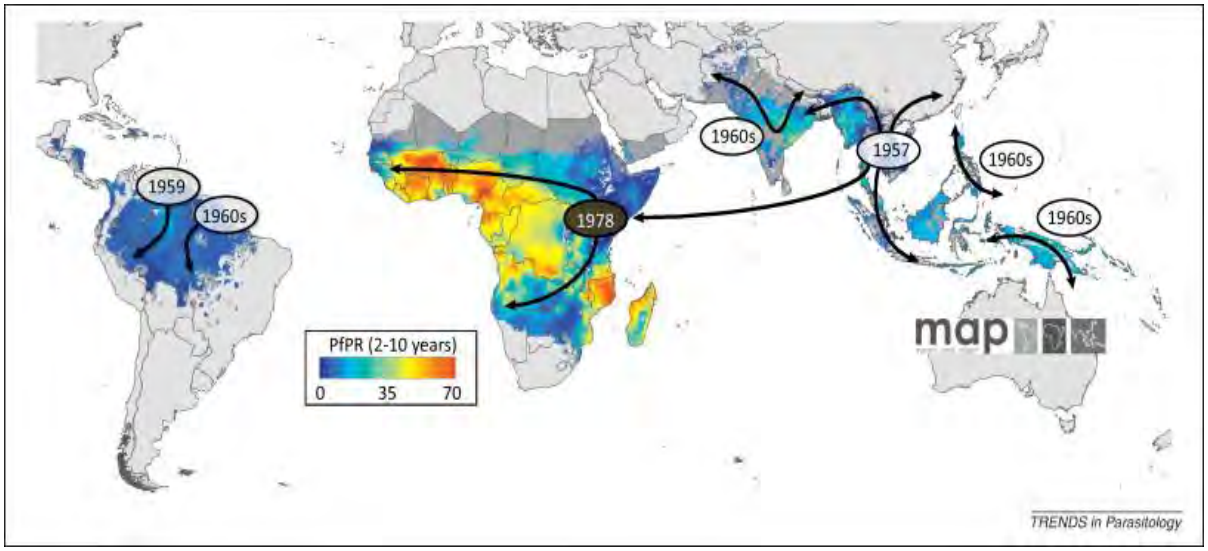


Figure 1.8 Age standardised (2-10 year age-range) *P. falciparum* endemicity map for 2010. The appearances and global spread of CQ resistance in *P. falciparum* are overlaid. White circles indicate the six independent origins of resistance which later spread and was established on the African East coast in 1978. Copyright © 2012 Elsevier, Ecker *et al.*[42]

QN, although the oldest known antimalarial, has retained efficacy considerably longer than CQ, since it has been used much less extensively.[14,43] However, its efficacy in Southeast Asia has fallen below 50% owing to its accelerated use for the treatment of CQ-resistant malaria.[44] On the other hand, resistance to the quinolinemethanol, MF, developed rapidly over just five years in Thailand due to its intensive use since it was introduced.[45,46] The mechanism of resistance to this drug has been associated with the *P. falciparum* multidrug resistance gene (*Pf*mdr1) and overexpression of its protein product Pgh-1.[47] However, some MF-resistant strains show no alternation of this gene and hence it is probable that resistance can arise by more than one mechanism.[48] Although resistance to all of the quinoline based drugs has been reported, cross resistance between the 4-aminoquinolines and the quinoline methanols does not develop readily and in some cases, resistance between these two classes is inversely correlated.[14] This suggests that parasites that are highly resistant to CQ may be more sensitive to the quinoline methanols. The appearance of resistance to the most widely-used antimalarial drugs is summarised in *Figure 1.9*. Interestingly, after the first report of resistance, QN, CQ and AQ remained effective for over a decade before resistance spread to other regions. *Figure 1.9* also highlights other

shortcomings of the known antimalarial drugs, such as halofantrine (Halo), which was not used after 1998 owing to cardiotoxicity concerns or AQ, which was temporarily disapproved as a drug owing to serious side effects. It was subsequently reinstated six years later when the benefits were believed to outweigh the risks. As of 2008, clinical resistance to artemisinins had not been demonstrated; however, this status has changed in the last seven years.

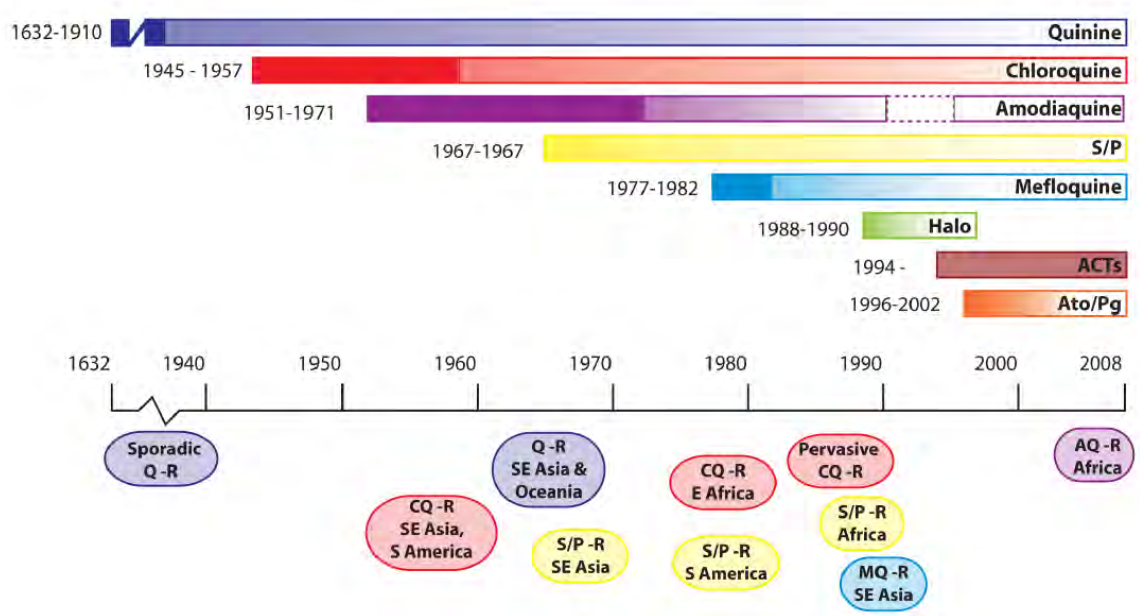


Figure 1.9 The emergence of resistance to the clinically important antimalarials. The years shown beside the drug bars represent the date the drug was first introduced to the year of first reported resistance. The circles indicate the approximate period for the spread of QN, CQ, AQ, S/P and MQ resistance through different geographical areas.

Evidence of resistance to artemisinin was first reported in 2008 along the Thai–Cambodian border after unexpected high failure rates associated with artemisinin monotherapy as well as reduced *in vitro* drug-susceptibility data were reported.[49,50] This is despite the fact that artemisinin should not be given as a monotherapy, owing to its short half-life in patients. Instead, it should be administered in combination with drugs which have a long half-life, like MF or lumefantrine.[45] However, as of 2014, the marketing of artemisinin-based monotherapies was still allowed in eight countries by 24 pharmaceutical companies, a major decline since 2007 when the WHO prescribed ACTs as first-line treatment of uncomplicated

falciparum malaria. As a result, ACTs have since been used in almost all the malaria-endemic countries.[10] Nonetheless, the WHO reports that addressing the threat of artemisinin resistance should focus on eliminating monotherapies entirely and identifying additional therapeutic resources.

For over five decades, the mechanisms of drug resistance were investigated in order to characterise the nature of physiological changes to the parasite. Particular focus has been given to CQ resistance mechanisms. The primary characteristic of CQ-resistant parasites is that they accumulate less CQ in their DV, compared with CQ-sensitive ones.[51,52] Although, there was much debate over the molecular basis for this decreased drug accumulation for many years, the key genetic mechanisms that determine resistance have now been established.[14]

The first proposal for the mechanism of decreased accumulation was as a result of reports that CQ-resistant parasites released accumulated CQ from the DV almost 50 times faster than the sensitive strains.[53] Drug efflux has been shown to be the mechanism by which less drug is accumulated in many types of cancer cells.[54] This multidrug resistance can be reversed by exposure to verapamil (*Figure 1.10*).[55]

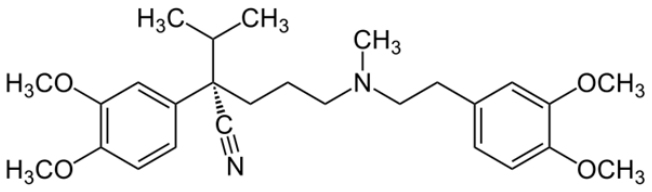


Figure 1.10 The chemical structure of verapamil.

It was believed that the mechanism of CQ resistance in *P. falciparum* is similar to that of multidrug resistance in cancer cells due to the decreased accumulation of CQ observed in resistant strains[52] and the ability of verapamil to reverse CQ-resistance.[56] Before the year 2000, several other mechanisms were proposed, including altered CQ uptake or drug efflux at the cytoplasmic membrane of the parasite; altered H^+ flux or CQ uptake at the DV membrane; increased detoxification of the CQ-haematin complex, mediated by glutathione S-transferase; and reduced access of CQ to haematin.[56,57,58,59] However, Fidock *et al.*[60] identified a DV transmembrane protein, *P. falciparum* CQ resistance transporter (*Pf*CRT) as

the primary determinant of CQ resistance. This was made possible by a detailed analysis of a genetic cross between the HB3 (CQ-sensitive) and Dd2 (CQ-resistant) strains where a highly interrupted gene, with eight point mutations compared to the wild type, was observed. Furthermore, allelic exchange experiments allowed for these mutations to be applied to a CQ-sensitive clone whereupon lower CQ activity, chemosensitisation by verapamil and reduced CQ accumulation, all the features of resistance, were observed. In particular, the K76T mutation, in which the amino acid lysine (K) is substituted by threonine (T) at position 76 was identified as being central to CQ resistance and has since been used as a molecular marker.[61,62] A study in 2008 further showed that a fluorescently tagged analogue of CQ is able to bind to *Pf*CRT.[63] The mutant *Pf*CRT protein has also been claimed to cause an increase in the acidity and volume of the DV in resistant strains Dd2 and GCO3/C3^{Dd2}, the latter of which is CQ-resistant via transfection with mutant *Pf*CRT. This DV biogenesis was demonstrated by Roepe and co-workers[64] using fluorescence imaging with spinning disk confocal microscopy, allowing for 3D z-stack images of live parasites at different time intervals.

There is also evidence to suggest that other genes contribute to CQ resistance. In particular, the *Pf*mdr1 protein causes important physiological changes in the DV.[60,65] These multigenic mutations are due to the unrestrictive, widespread and prolonged use of individual drug compounds, which put drug pressure on parasite populations.[66]

Very recently, in 2015, *Pf*CRT was shown to actively transport CQ across proteoliposome membranes from a lower to higher pH, mimicking that of efflux out of the DV.[67] The active CQ-uptake, measured after one minute, from an environment with a pH of 6 to one with a pH of 8, was shown to be enhanced in *Pf*CRT from resistant isolates 7G8 and Dd2 relative to the wild type. This was also the case for uptake by *Pf*CRT of the resistance reversal agent, verapamil. In addition, the transport of CQ could be inhibited by 25% with verapamil concentrations of approximately 100 µM and by 80% at 1000 µM verapamil. The observation agreed with that reported in 2009, where both the wild-type and resistant forms of *Pf*CRT were expressed at the surface of *Xenopus laevis* oocytes and showed remarkably different CQ uptake levels.[68] The efflux mechanism is believed to reduce CQ levels in the DV below those required to inhibit haemozoin formation.[69]

1.7 Mechanism of antimalarial drug action for the quinoline and related compounds

The persistent development of resistance to antimalarial drugs by *P. falciparum* has led to an urgent need for a better understanding of the multiple aspects of the blood-stage parasite biology as well as the mechanisms by which antimalarial drugs exert their therapeutic action. This has been made possible by the development of techniques for the *in vitro* cultivation of malaria parasites in a laboratory environment, which have become vital for determining compound activity and probing their structure-activity relationships (SARs).

1.7.1 *P. falciparum* cell culturing and drug sensitivity assays

The technique for culturing *P. falciparum* was established in 1976 by Trager and Jensen[70] and has since been adapted and refined in many versions, optimising oxygen pressure, fluid dynamics, nutrient availability, and temperature.[71] The red blood cell stage of the parasite's life cycle is now the most widely analysed, since it is during this 48 h subcycle that severe malaria symptoms occur.[72] In addition, most antimalarial drugs act to inhibit parasite growth at this stage. Access to sufficient numbers of infected cells for experimenting is made possible by massive parasite multiplication within easily obtainable host human red cells. Detecting the parasitised erythrocytes is achieved using Giemsa stain, a common technique in use since 1904[73] which binds specifically to the phosphate groups of DNA, allowing for the pink erythrocytes and dark purple parasite cells to be readily distinguished (*Figure 1.11*).

a) b)

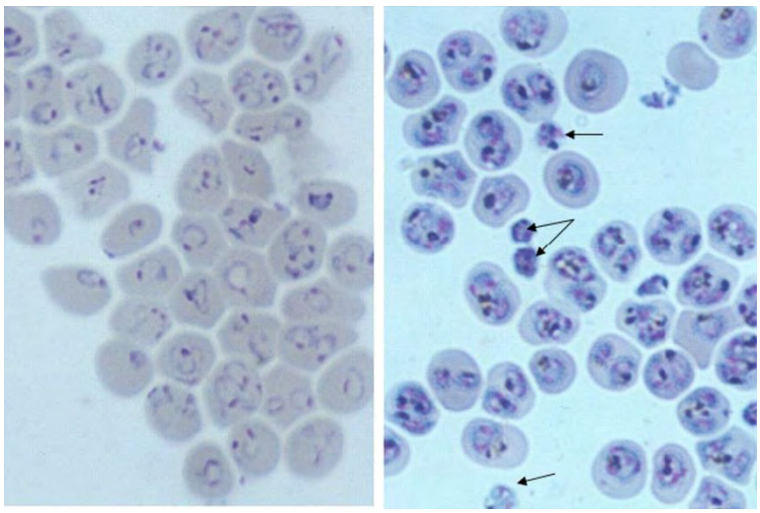


Figure 1.11 Geimsa-stained cultures of *P. falciparum*, showing different phases within the erythrocytic stage of the life cycle; a) rings and b) trophozoites and mature schizonts having burst just burst from the red blood cell (indicated by arrows). Copyright © 2008, The Society for Experimental Biology and Medicine, A. U. Orjih.[74]

Techniques for synchronising parasite cultures to the same phase are well established and often necessary for experiments sensitive to ratios of rings (*Figure 1.11a*) and trophozoites (*Figure 1.11b*).[72,75] Such a protocol is necessary for dose response experiments to determine 50% inhibitory concentrations of test compounds. Many variations of drug sensitivity assays have been developed over the past five decades, most of them involving radioactive substrates, such as [^{3}H]isoleucine, [^{3}H]hypoxanthine and [^{3}H]ethanolamine, to measure parasitic growth.[76,77,78] Although these methods were reliable, they were very expensive and required special handling and hence resulted in the development of alternate procedures. One such method is the enzymatic parasite lactate dehydrogenase (LDH) assay reported in 1993 by Makler *et al.*[79] Another, reported in 2001, detects parasites via a colorimetric highly sensitive double-site LDH enzyme-linked immunosorbent assay (ELISA) that recognises specific *P. falciparum* antigens.[80] More recently, high throughput assays involving fluorescent nucleic acid intercalating dyes, which bind specifically to DNA and allow all stages of parasites to be distinguished from the non-DNA-containing erythrocyte, have been successfully used to obtain IC_{50} values comparable to those in other assays.[81,82,83] In particular, the SYBR Green I-based fluorescence assay has been shown to be a simple, cost-

effective, relatively safe, robust and accurate method for high throughput screening (HTS) to identify inhibitors of *P. falciparum.*[84]

1.7.2 Nucleic acid and proteins as possible targets of chloroquine

Early studies on the mechanism of action of CQ revealed its interaction with DNA and toxic effect on a wide range of mammalian cells.[85] However, the ability of quinolines to inhibit DNA replication and RNA synthesis cannot account for their three orders of magnitude higher potency against *P. falciparum*. Nor do the plasmodial enzymes involved in DNA replication appear to be direct targets of CQ.[86,14] More convincing evidence as to the drug target of CQ in the malaria parasite came in 1970 from observations that the drug is only active against the erythrocytic stage, particularly while the parasite is degrading Hb.[87] It was then shown that several quinoline antimalarials directly inhibit the activity of partially purified haemoglobin-degrading proteases.[88] Inhibition of protein degradation was also demonstrated in CQ treated trophozoite-infected erythrocytes, leading to the idea that the parasite is killed by subsequent starvation as a result of its inability to feed on Hb.[89] This protease inhibition process is reversible upon removal of the drug, but the inhibition of parasite growth by quinoline antimalarials is irreversible.[90,57] This observation was critical for concluding that CQ must have more specific, irreversible effects on the parasite.

1.7.3 Haemozoin inhibition

Studies have shown that a typical red blood cell contains 310–350 mg/mL of Hb and that 25% to 75% is degraded by the parasite during a 48 h intraerythrocytic cycle.[91,92] This results in a haem concentration of approximately 20 mM if averaged throughout the red cell. However, if concentrated in the DV, which makes up 3-5% of the erythrocyte volume, a much larger haem concentration of 200-500 mM would be possible. This massive toxic waste problem faced by the parasite is overcome by the formation of haemozoin (*Section 1.4*) and is an ideal target for chemotherapy. Indeed, studies in the mid-1990s showed that quinacrine treated *P. falciparum*-infected patients contained gametocytes with reduced malaria pigment.[14] Furthermore, CQ resistant strains of *Plasmodium* were reported to

contain less pigment than the sensitive strains.[93] In 1986, Fitch *et al.*[94] proposed that CQ interferes with haemozoin formation by forming a complex with haem, after it was shown that CQ and haem interact stongly.[95] This was then supported by Slater and Cerami in 1992 when they demonstrated that haemozoin formation *in vitro* can be inhibited by high micromolar concentrations of CQ.[96] Since then there has been much debate over the precise mechanism by which the inhibition of crystal formation takes place (see *Section 1.8.2*).

Before techniques to measure differences in haemozoin and free haem levels within the parasite were developed, evidence in favour of this drug target was often obtained by attempting to find correlations between synthetic haemozoin (β-haematin) inhibition activity and parasite activity. The first study in the early 1990s which endeavoured to find a relationship between β-haematin (βH) formation and parasite activity, reported no correlation when comparing different quinoline subclasses such as CQ and QN.[96] However, Dorn *et al.*[97] criticised the trophozoite lysate based method, which this study used to promote haematin crystallisation, owing to the nonspecific binding of drugs to cytosolic proteins. Instead, they used a radio-labelled [^{14}C]-haemin assay method described in *Section 1.8.5* to carry out correlation analysis on the quinoline antimalarials. A significant correlation was identified between the βH inhibition activity and the CQ-sensitive NF54 parasite activity (*Figure 1.12*) for seven clinical antimalarial drugs, halofantrine (HF), pyronaridine (PYR), AQ, CQ, QN, QC, MQ and the experimental bisquinoline, Ro 48-6910. This was remarkable considering the structural diversity of these compounds. Furthermore, although HF was the only non-quinoline drug, it fitted the trend established by the quinolines, strongly suggesting that haemozoin formation was the target.

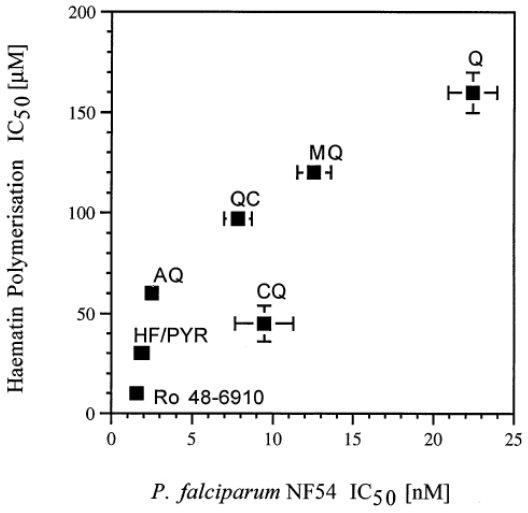


Figure 1.12 Correlation of βH inhibition (incorrectly referred to as haematin polymerisation) and *P. falciparum* NF54 activity (r^2 = 0.91, P = 0.002) for antimalarials of which seven out of eight were quinolines. Copyright © 1998 Elsevier, Dorn *et al*.[97]

There is strong evidence to suggest that the physicochemical properties of quinolines, as well as the proton gradient between the intra- and extracellular parasite environment influences compound concentration within the parasite DV.[58,98] This results in different accumulation ratios at the site of haemozoin formation, which was first linked to parasite activity in the mid-1990s.[98] Hawley *et al*.[99] showed that this accumulation parameter was an essential factor for correlating βH inhibition to parasite activity, without which, no statistically significant correlation was found ($r < 0.25$; $P > 0.28$) using multiple strains of *P. falciparum*, namely the CQ-sensitive 3D7 and HB3, and the CQ-resistant K1 and PH3. The cellular accumulation ratio (CAR) for seven quinoline antimalarials was normalised to the largest CAR within each strain and then multiplied by the corresponding parasite IC_{50}. *Figure 1.13* shows the resulting relationship between βH inhibition and the accumulation-normalised CQ-sensitive parasite activity. Interestingly, no significant correlation was found for the resistant strain (r = 0.17, P = 0.72) since drug accumulation at the specific target site is disturbed by the efflux mechanism in resistant strains, despite the observed accumulation of drug in the parasite cytosol.[99]

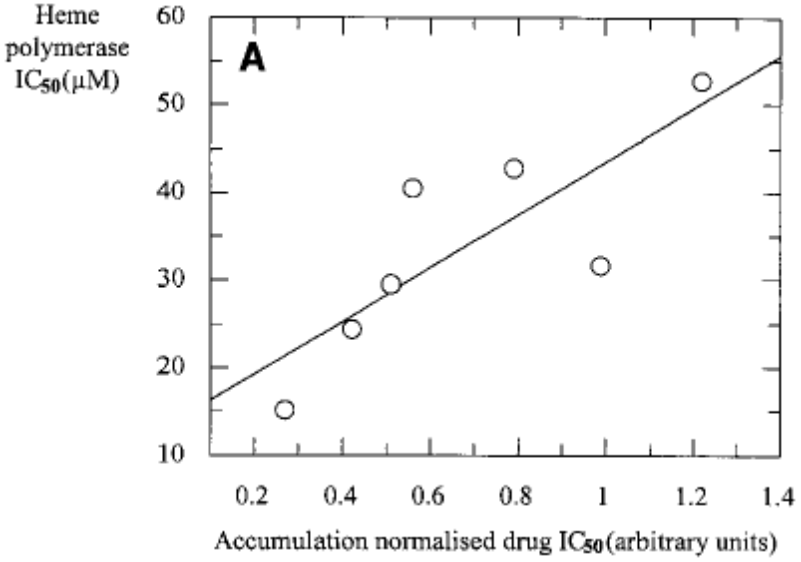


Figure 1.13 Correlation of βH inhibition (inaccurately labelled "heme polymerase") and the accumulation normalised *P. falciparum* 3D7 activity (r = 0.81, P = 0.026) for known quinoline antimalarial drugs. Copyright © 1998, American Society for Microbiology, Hawley *et al.*[99]

Another early study to relate the βH and parasite activities of synthetic CQ analogues was reported in 1999 by Vippagunta *et al*.[100] Thirteen CQ analogues varying in chain length and quinoline substituents were evaluated. In addition, binding constants (K_a), a measure of their affinity for haem, were reported. This turned out to be important for SAR analysis since a modest correlation was identified when the βH inhibition activity for each analogue was normalised to the corresponding strength of haem interaction by dividing the IC_{50} by K_a (r = 0.61, P = 0.046). They reported that knowledge of the accumulation ratios for the compounds would have likely improved the correlation statistics; however, no attempt was made to determine CARs. Nonetheless, this result was intriguing since it indicated that haem binding may have some effect on parasite activity, which may have been as a result of its influence on accumulation.

The second study on this matter that also involved CQ analogues, was reported in 2002 by Kaschula *et al.*[101] where measured βH inhibition values were derived from the assay based on the differential solubilisation of haematin in Dimethyl sulfoxide (DMSO; see *Section 1.8.5)*.[102] This study reported no statistically significant direct correlation between CQ-sensitive D10 parasite activity and βH inhibition. Rather, an impressive correlation was

observed with βH inhibition when the parasite activity was first normalised by multiplying the IC_{50} by the relative vacuolar accumulation ratio (α) for each derivative. The accumulation ratio for a specific derivative was based on its pK_a values and gave an indication of the extent of pH trapping within the DV. To calculate α, this ratio was divided by that of CQ to account for the relative amount of pH trapping each compound was capable of achieving. Since the substituents of the different derivatives displayed a range of electron withdrawing and releasing properties, the basicity of the quinoline nitrogen varied significantly between derivatives, resulting in a 10-fold difference in pH trapping ability. Consequently, a good correlation between βH inhibition and parasite activity, normalised to account for pH trapping in the DV, was observed with an r^2 of 0.83 and a P value of 0.0043 (*Figure 1.14*).

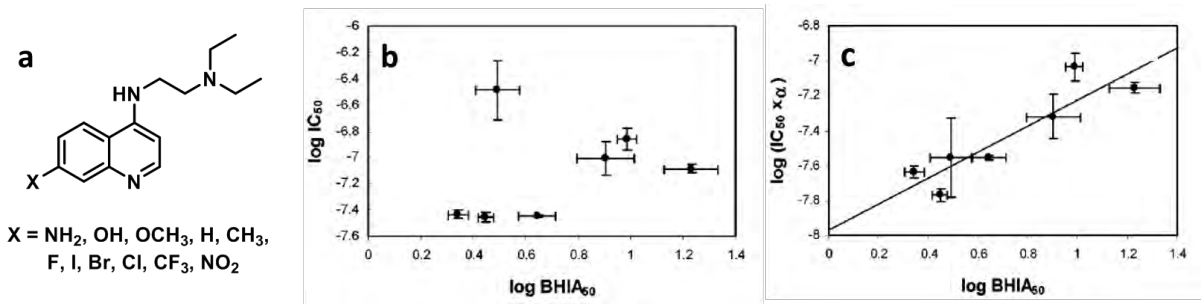


Figure 1.14 a) The synthesised quinoline derivatives and b) structure-activity relationship (SAR) trends showing no correlation between the βH inhibition ($BHIA_{50}$) and the D10 parasite activity (IC_{50}) and c) an excellent correlation for βH inhibition with the normalised parasite activity (IC_{50} x α). All correlations were carried out with logged values. Reprinted (adapted) with permission from Kaschula *et al.*[101] Copyright © 2002, American Chemical Society.

Advances in the βH inhibition assay procedure enabled superior correlations (*Figure 1.15*) to be found for these compounds using the biological activity data supplied by Kaschula *et al.*[101] This came after the development of the ferrihemochrome method by Egan and co-workers.[103] (see *Section 1.8.6*). Although no direct correlation between βH inhibition and parasite IC_{50} was previously identified, the improved free haem detection technique gave βH IC_{50} values (Phiβ) which did correlate with parasite activity (r^2 = 0.68 and p = 0.0005).

Furthermore, multiple correlation analysis revealed a relationship whereby a linear combination of the βH activity and the vacuolar accumulation ratio (VAR) were capable of predicting parasite activity at the 99% confidence level.

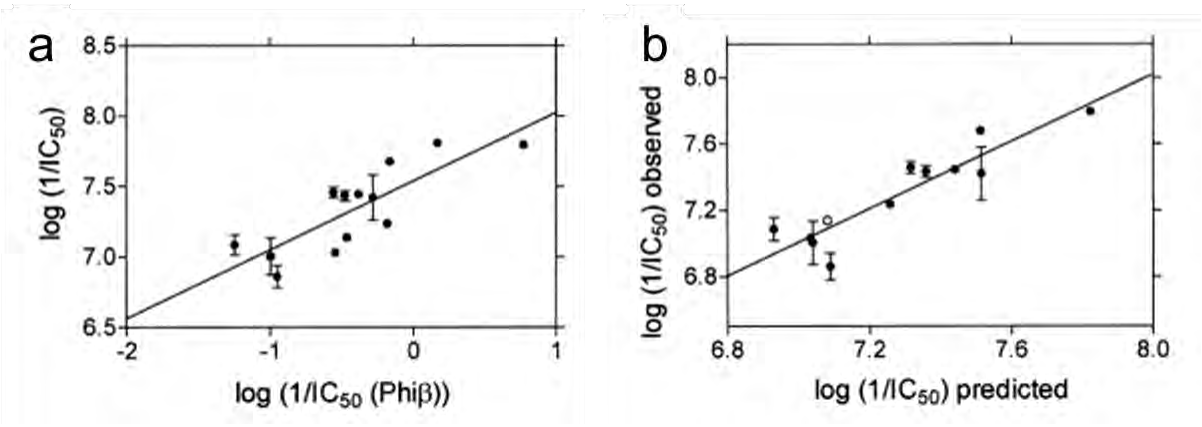


Figure 1.15 Correlations for the Kaschula *et al.*[101] quinoline derivatives using the ferrihemochrome βH inhibition assay method.[103] The parasite activity was related to βH inhibition via a) direct correlation or b) multiple correlation analysis with $\log(1/IC_{50}) = 0.61*\log(1/IC_{50}(Phi\beta)) + 0.17*\log VAR + 6.98$ ($r^2 = 0.84$, F = 21.3; >Fcrit = 8.65 at the 99% confidence level), where VAR is the vacuolar accumulation ratio. Copyright © 2005, Elsevier, Ncokazi, *et al*.[103]

Using the same assay method for βH inhibition, some years later the same group synthesised a series of 4-amino-7-chloroquinolines with dibenzylmethylamine (dibemethin) side chains, which were also shown to exhibit a correlation with biological activity against the parasite.[104] Here, several interesting trends were observed, although no statistically significant correlations were identified directly relating βH inhibition to either the parasite activity or the accumulation normalised activity. An additional parameter, *pos*, was defined to account for the position (*ortho*, *meta* or *para*) of the aminoquinoline group on the dibemethin phenyl ring. This enabled a good multiple linear correlation of log IC_{50} with the accumulation ratio, log βHIC_{50}, and *pos* at the 95% confidence level.

Similar trends have also been observed for non-quinoline compounds. In 2002, several phenothiazines were shown to be good βH inhibitors, although only a subset of them displayed potent antimalarial activities. In addition, the βH IC_{50} values were not able to predict the relative parasite growth inhibition activities on the basis of different levels of

uptake, reported to be a combination of haem binding and pH trapping.[105] A series of xanthones with side chains ranging from two to eight carbon atoms in length were reported to show a correlation between the haem binding constants and D6 (CQ-sensitive) parasite activity (*Figure 1.16*).[106] This result was intriguing in light of the results reported by Vippagunta *et al.,*[100] which required the division of βH inhibition values by haem binding constants to achieve a significant correlation with the parasite IC_{50}. These studies suggested a possible connection, the particulars of which are unknown, between these two features, namely haem binding and *in vitro* parasite activity.

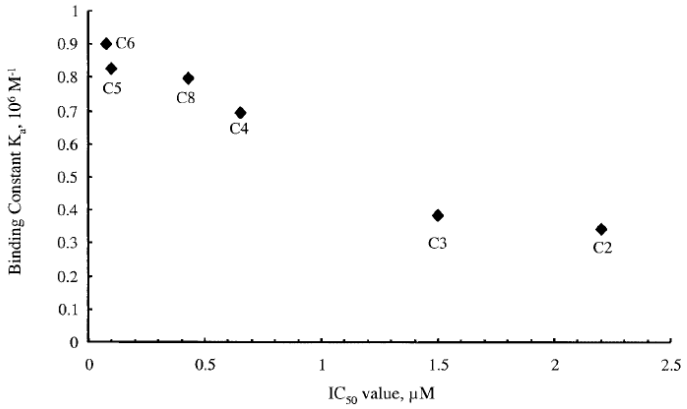


Figure 1.16 Correlation of CQ-sensitive D6 antimalarial potency (IC_{50}) versus haem binding affinity (K_a) for the 3,6-bis-ω-*N*,*N*-diethylaminoalkoxyxanthones. Copyright © 2002, American Society for Microbiology, Kelly *et al*.[106]

Recently, numerous reports detailing correlations between βH inhibition and parasite activity have been published for neocryptolepines and isocryptolepines. Lu *et al.*[107] reported a very weak correlation between the log of NF54 activity with the βH activity (*Figure 1.17a*), showing that the biological activity of a compound depends on more factors than its ability to inhibit haemozoin, if this is indeed the target. In fact, the parasite activity could be well correlated using multiple physicochemical parameters such as the mean water of hydration (H_2O.hyd), polar surface area (PSA) and number of hydrogen bond donating atoms (HBD) by $\log(NF54IC_{50}) = -0.299(H_2O.hyd) + 0.028(PSA) - 0.652(\ln(HBD)) + 3.108$ ($r^2 = 0.904$, $P < 0.0001$), which was statistically significant at the 99% confidence level (*Figure 1.17b*). Further analysis of this equation revealed that a balance between the hydration and polarity

of a molecule must be established for good biological activity. This was also the case in a study by Wang *et al.*[108] for *N*-6-Me-neocrytolepines. A statistically significant correlation was observed for a four-parameter model involving solubility (Sol), H_2O.Hyd, logP and the square of the dipole moment (μ^2) as $\log(NF54IC_{50}) = 0.302(Sol) - 0.188(H_2O.hyd) + 0.446(\log P) + 0.002(\mu^2) - 4.670$, ($r^2 = 0.940$, $P < 0.0001$), statistically significant at the 99% confidence level (*Figure 1.17c*). Although this specific correlation did not involve βH inhibition, a good correlation between the log of the βH inhibition and the log of the NF54 parasite activity was observed when only the five compounds with basic amino side chains were considered (*Figure 1.17d*). This result indicated that activity against the parasite is strongly influenced by the ability of a compound to inhibit haemozoin formation, at least for those that are able to accumulate in the DV through pH trapping. Yet another study, this time involving isocryptolepine analogues, showed correlations between βH inhibition and NF54 activity, albeit weak (*Figure 1.17e*).[109] In addition, the statistics could be improved by incorporation of physicochemical parameters, namely solubility (sol.) and percentage hydrophilic surface area (%hydrophil.) so that $\log(NF54IC_{50}) = 0.193(Sol.) - 0.027(\%Hydrophil.) - 10.069(1/\beta HIC_{50}) - 2.208$, $r^2 = 0.61$, $P < 0.0001$, statistically significant at the 99% confidence level (*Figure 1.17f*). This result was in agreement with the trend found previously by Wang *et al.,*[108] again reflecting the balance of properties relating to hydration and polarity for optimum parasite activity. Upon hybridisation of the isocryptolepines with artesunate, no correlations between βH inhibition and parasite activity could be found.[110] In fact, the hybrid derivative with the lowest βH activity exhibited the highest antimalarial activity. This result further confirmed the disconnection between the IC_{50} values from the extracellular βH assay and intracellular parasite growth assay.

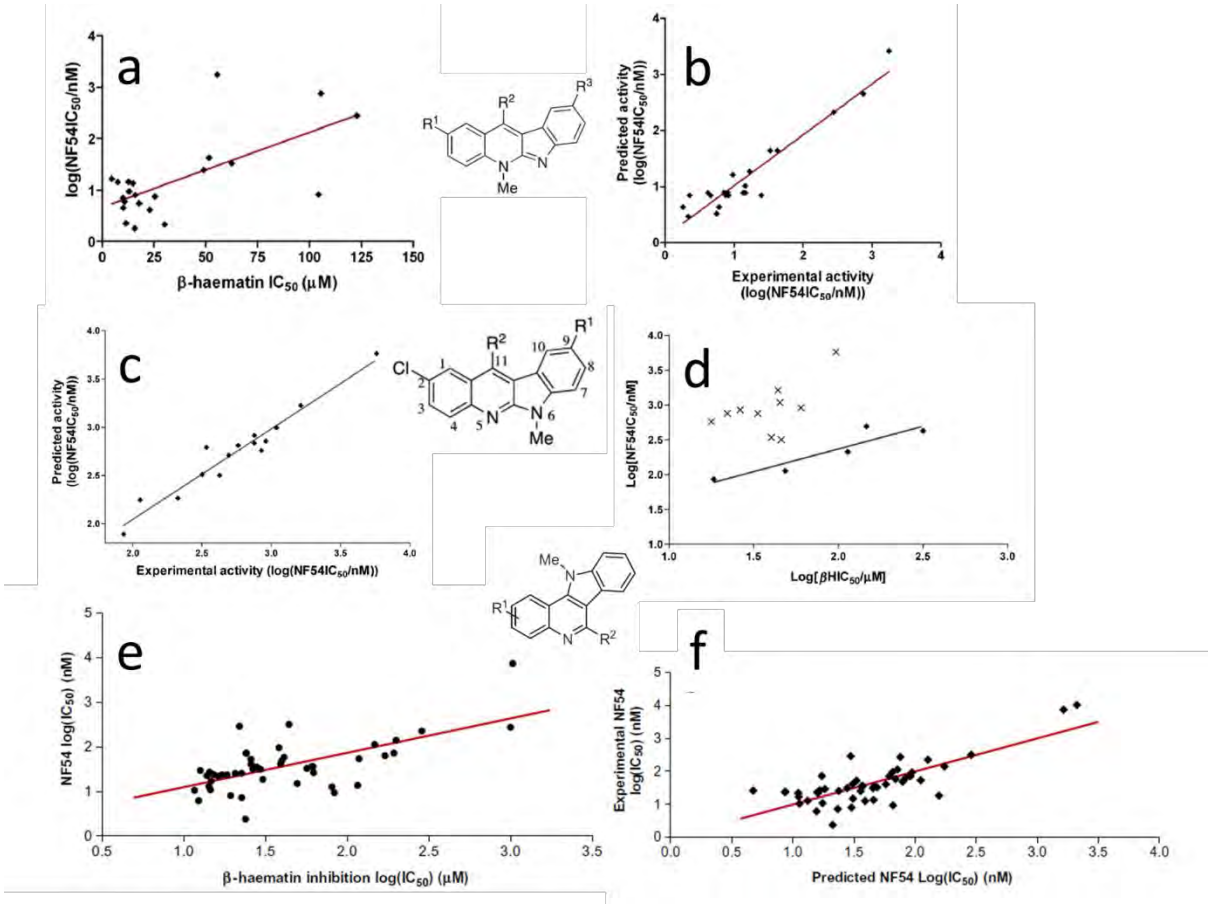


Figure 1.17 Correlations reported for neocryptolepines and isocryptolepines. **(a)** a very weak correlation between the log of NF54 activity with the βH activity (r^2 = 0.43, P = 0.001) and **(b)** multiple correlation analysis to give log(NF54IC_{50}) = -0.299(H_2O.hyd) + 0.028(PSA) - 0.652(ln(HBD)) + 3.108; (r^2 = 0.904, P < 0.0001) statistically significant at the 99% confidence level. Reproduced from Lu *et al.*[107] Copyright © 2013, Elsevier. **(c)** Multiple correlation analysis to give log(NF54IC_{50}) = 0.302(Sol) − 0.188(H_2O.hyd) + 0.446(log P) + 0.002(μ^2) − 4.670; (r^2 = 0.940, p < 0.0001) statistically significant at the 99% confidence level and **(d)** a good correlation between the βH inhibition and the NF54 parasite activity only observed for compounds with basic amino side chains (filled circles). Reproduced from Wang *et al.*[108] Copyright © 2013, The Pharmaceutical Society of Japan. **(e)** A weak correlation between βH inhibition and NF54 activity (r^2 = 0.41, P < 0.0001) and **(f)** multiple correlation analysis to give log(NF54IC_{50}) = 0.193(Sol.) - 0.027(%Hydrophil.) - 10.069(1/βHIC_{50}) - 2.208; (r^2 = 0.61, P < 0.0001) statistically significant at the 99% confidence level. Reproduced from Wang *et al.*[109] Copyright © 2014, Elsevier.

In 2006, the effect of CQ on haemozoin levels within live erythrocytic parasites was observed using spinning disk confocal microscopy by Roepe and co-workers.[111] By measuring pixel density in 3D differential interference contrast transmittance z-stack images, haemozoin production at various time intervals and concentrations of CQ were measured. While CQ did not affect the initial rate of haemozoin formation at 15 to 22 h, it was found to restrict the normal haemozoin levels reached at 27 to 29 h in a dose dependent manner for both CQ-sensitive HB3 and CQ-resistant Dd2 strains at the relevant IC_{50} (*Figure 1.18*). The addition of a resistance reversal agent, verapamil, resulted in haemozoin inhibition at the lower HB3 IC_{50} in the Dd2 strain. Furthermore, preformed haemozoin was not destabilised, even at high CQ concentrations, for either strain.

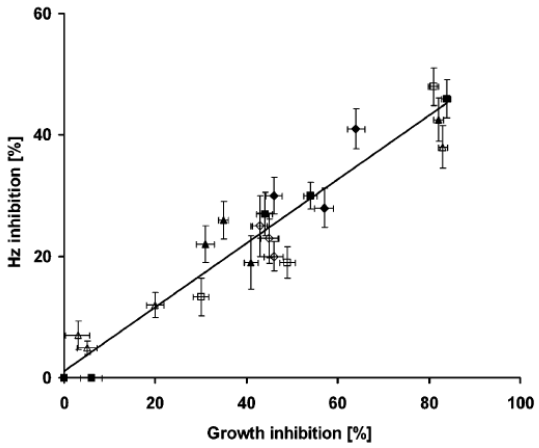
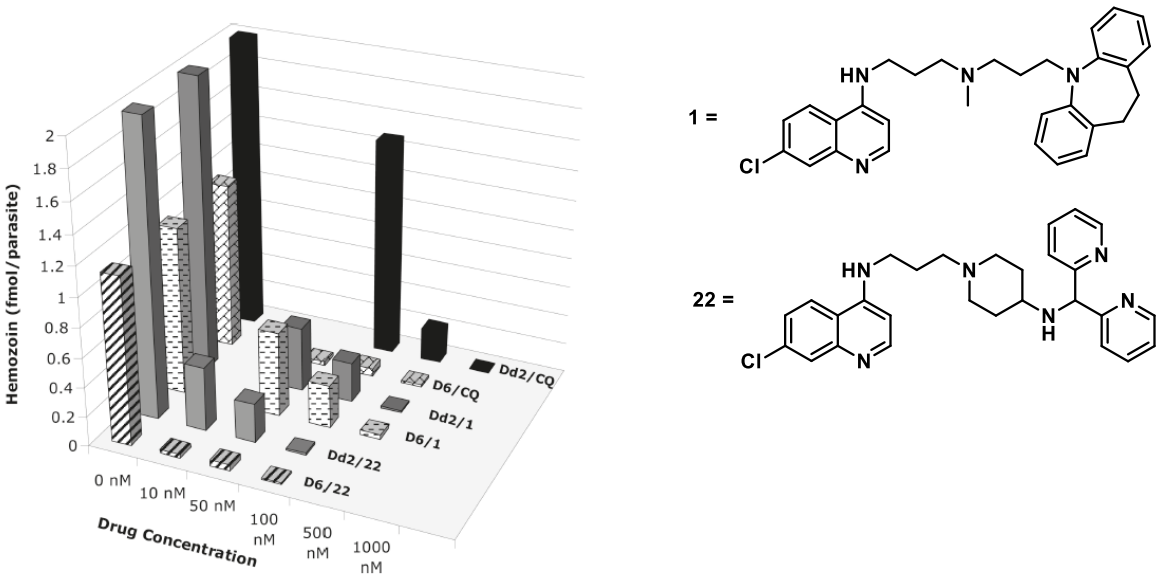


Figure 1.18 A linear relationship between the percent haemozoin inhibition and the percent parasite growth inhibition for the CQ-resistant Dd2 (black markers) and CQ-sensitive HB3 (white markers) strains without verapamil (squares), with verapamil (triangles) or with a combination of CQ and verapamil (diamonds). Reprinted with permission from Gligorijevic *et al.*[111] Copyright © 2006, American Chemical Society.

In 2010, a series of CQ analogues containing resistance reversal-like moieties were synthesised and shown to overcome CQ resistance.[112] Upon testing βH inhibition IC_{50} values, a correlation with parasite growth inhibition was observed with r^2= 0.66. More suggestively, two potent compounds, **1** and **22** (*Figure 1.19*), were shown to cause a decrease in haemozoin formation in parasites. The haemozoin was isolated from lysed parasites by centrifugation and washing and then quantified via dissolution in 0.2 M sodium hydroxide

solution and UV-vis absorbance measurements at 400 nm. After normalisation to account for the haematocrit and parasitaemia, a per parasite concentration of haemozoin was plotted versus the concentration of drug (*Figure 1.19a*), clearly revealing a concentration dependent reduction in haemozoin. This was also demonstrated using microscopy, where decreased haemozoin in the DV could be visualised (*Figure 1.19b*).

a)



b)

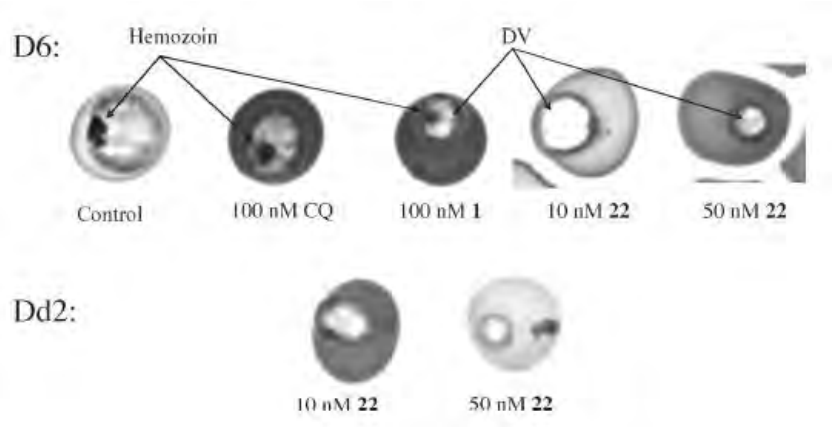


Figure 1.19 a) Haemozoin production in D6 and Dd2 parasites with varying concentrations of CQ and resistance reversal analogues **1** and **22**; b) microscope images showing a reduction of haemozoin in the DV in the presence of CQ, **1** and **22** relative to the control for the D6 strain (top) and little or no haemozoin for the Dd2 strain in the presence of **22** (bottom). Reprinted with permission from Burgess *et al.*[112] Copyright © 2006, American Chemical Society.

Studies that demonstrate a reduction of haemozoin in *P. falciparum* whole cells with test compounds, while supporting haem detoxification via haemozoin formation as the mechanism of action, do not conclusively validate the target, as diminished quantities of haemozoin may result from reduced Hb uptake or digestion, which would also result in less free haem. A more convincing experiment was reported more recently by Combrinck *et al*.,[113] where cell fractionation was used to directly measure both the amount of haemozoin and free haem in cultured *P. falciparum* whole cells. They reported a dose-dependent decrease in haemozoin and a simultaneous increase in toxic free haem in the presence of known quinoline antimalarial drugs. Furthermore, the dose-response curve for the increase in free haem as well as that for parasite survival crossed over at their respective 50% inhibitory concentration (IC_{50}) values, confirming a correlation between these two events. They also showed that the antifolate pyrimethamine and its combination with sulfadoxine do not inhibit cellular haemozoin formation, as expected. In particular, CQ showed the largest effect in the free haem to haemozoin ratio at three times its IC_{50} (*Figure 1.20*). This study provided direct evidence in support of haemozoin inhibition as the mechanism of action of CQ.

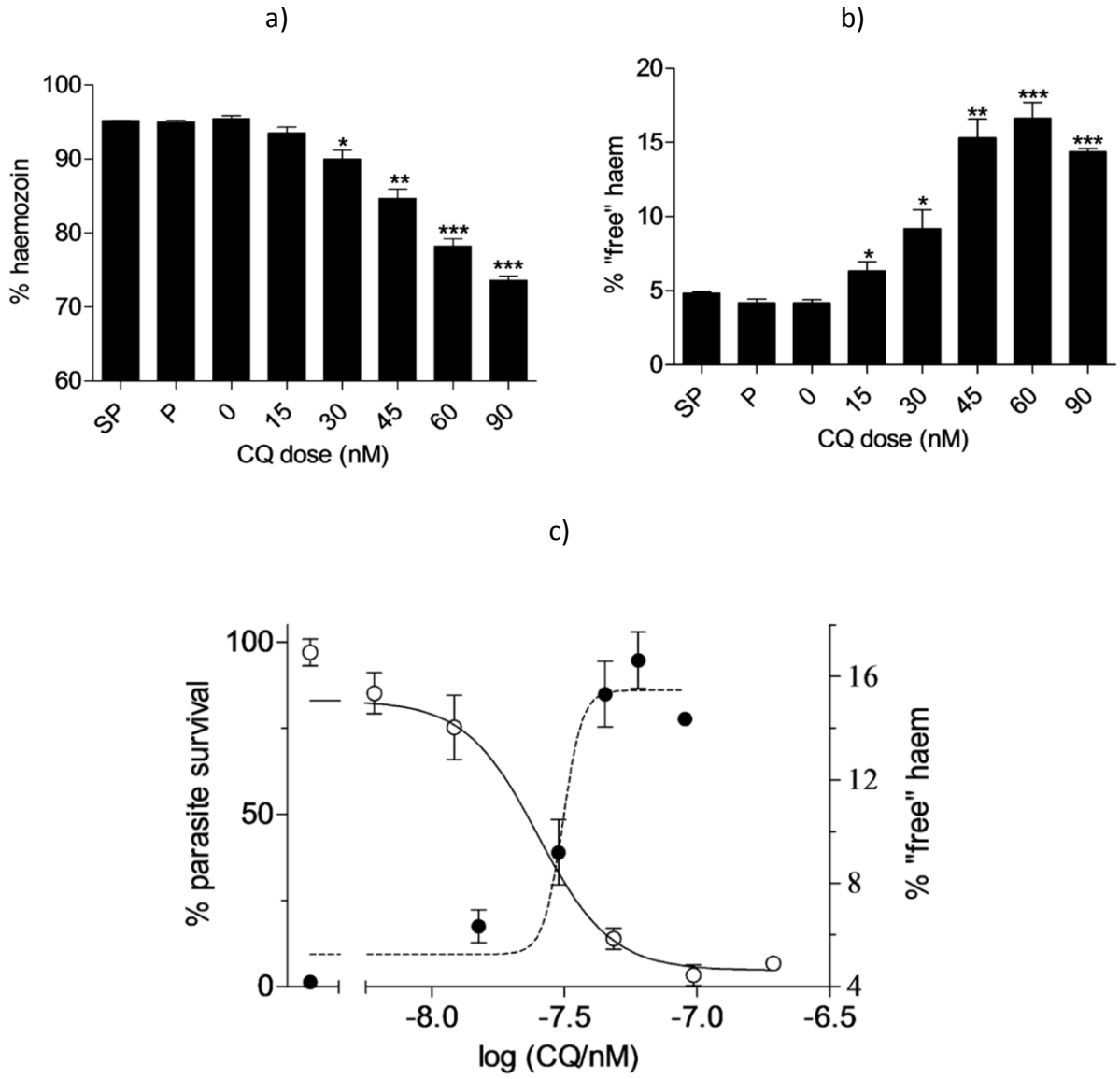


Figure 1.20 The effect of CQ on a) % haemozoin, b) % free haem and c) % parasite survival overlaid with % free haem. SP specifies the negative control, sulfadoxine/pyrimethamine, P represents pyrimethamine and asterisks indicate statistical significance relative to the control (2-tailed t-test): *P < 0.05; **P < 0.01; ***P < 0.001, n = 3. Copyright © 2013, American Chemical Society, Combrinck *et al.*[113]

1.7.4 Accumulation of quinoline antimalarials in the DV

In the 1970s, many papers were published revealing the extraordinary ability of malaria parasites to accumulate CQ,[114] AQ,[115] QN[116] and QC.[117] In 1974 Fitch *et al.*[52] showed that the accumulation of CQ is sensitive to the pH gradient between the DV and the extravacuolar environment by varying the pH of the latter. A decade later, Yayon *et al*.[51] determined the internal pH of the subcellular compartments of *P. falciparum* within the RBC. The DV, where sequestering of toxic free haem takes place, was determined using fluorescent probes to have an acidic pH of ~5. A year later they reported high pH-dependent accumulation levels of CQ in the DV, investigated by flow dialysis, in which diffusion across a membrane can be measured[118] as well as the tracing of CQ via fluorescence or a ^{14}C labelled analogue.[119] Millimolar concentrations of CQ were predicted to exist in the DV at therapeutic doses, accounting for the high activity of the drug. In the 1990s, more sophisticated methods allowed the proton gradients to be varied, by decreasing the extravacuolar pH with the acidifying agent NH_4Cl or via inhibitors of vacuolar proton pumps.[120,121] Accumulation measurements under these conditions showed that lowering the external pH results in less accumulation and hence these studies concluded that CQ accumulates in the acidic DV as a result of its weak base properties. This pH-dependent accumulation can be explained via the phenomenon of pH trapping, whereby a neutral species passively diffuses into the acidic DV and is protonated to form a less lipophilic (lower logD) charged species, which is then unable to cross the membrane. CQ, which has two protonatable sites, namely the quinoline nitrogen (pK_{a1} = 8.4) and diethylamino moiety (pK_{a2} = 10.1), exists almost entirely as the diprotonated species (CQH_2^{2+}) at a pH of 5.[122] Neutral CQ can diffuse into the DV where it will become protonated and trapped, with a lower intravacuolar concentration of neutral species present. Since equilibrium requires equal concentrations of the neutral species (CQ) both intravacuolarly and externally, more CQ will diffuse in, until concentrations of the neutral species are equal. The great majority of the available drug will exist in the DV as the CQH_2^{2+} species (*Figure 1.21*).[123]

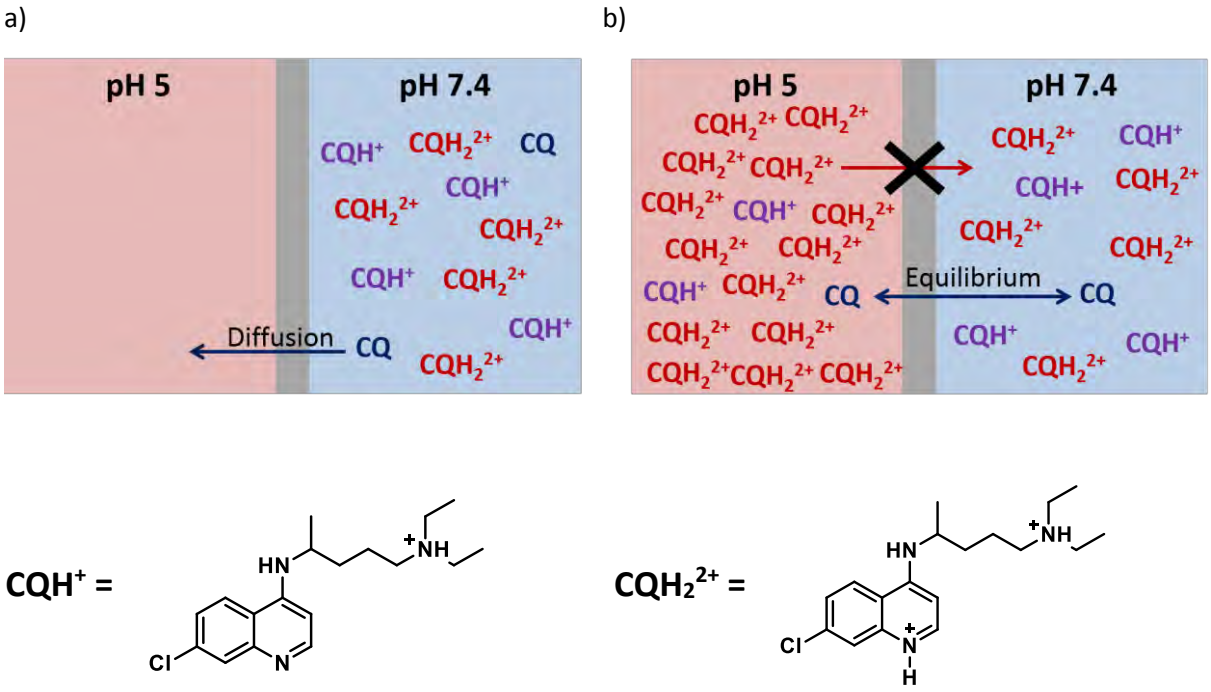


Figure 1.21 Representation of CQ pH-based accumulation (non-saturable uptake) in the acidic DV (red) from the neutral extravacuolar environment (blue). In a) neutral drug diffuses from high to low concentration, after which equilibrium is established in b) with the diprotonated species (CQH_2^{2+}) existing in a large excess, trapped in the DV.

The principle of pH trapping has also been reported to account for the accumulation of other quinoline antimalarials since they all possess two basic atoms.[124] However, in 1996 Hawley *et al.*[98] measured the accumulation of AQ analogues with varied side chains, possessing a range of pK_a values and lipophilicities. They found that experimental accumulation ratios were much greater than those calculated theoretically based on drug pK_a, compartmental pH, and Henderson-Hasselbach considerations. Later experiments by the same group showed that a saturable uptake of CQ, possibly due to haem binding, is responsible for the large accumulation observed.[59] This was identified after the availability of free haem was blocked via a specific inhibitor of Hb digestion and a concentration dependent reduction in CQ-binding sites was demonstrated. This saturable uptake completely accounted for the reduced CQ activity and accumulation in CQ-resistant strains as well as the reversal of resistance by verapamil (VP) as shown in *Figure 1.22*. In addition, erythrocyte ghost cells, loaded with Fe(III)PPPIX, could model this saturable CQ binding and was dependent on the Fe(III)PPIX concentration.[125]

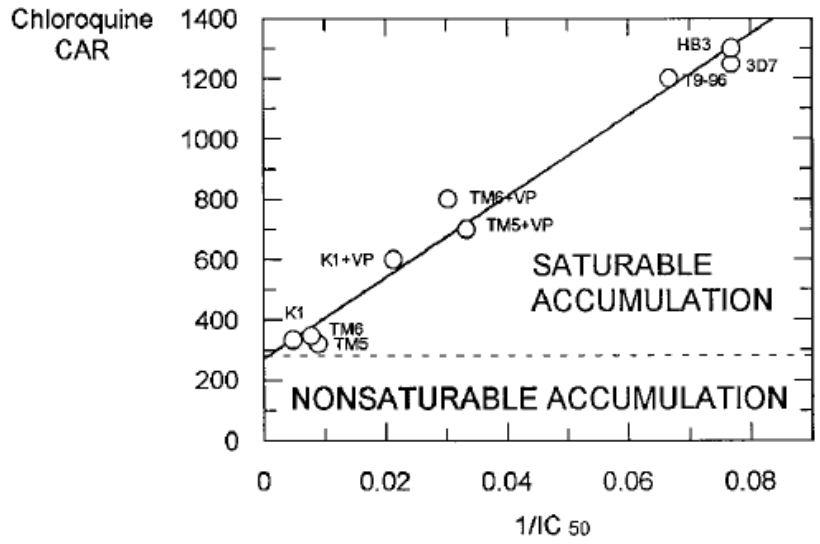


Figure 1.22 The CAR of CQ versus the inverse IC_{50} in three CQ-sensitive strains (HB3, 3D7 and T9-96) and three CQ-resistant strains (K1, TMS5 and TMS6) with and without verapamil (VP). Copyright © 1998-2010, Molecular Pharmacology Online, Bray *et al.*[59]

Although the non-saturable binding from pH trapping only accounts for 5-12% of the total observed accumulation, a review in 2001 by Egan noted that without this factor, the CAR of CQ would be negligible.[123] This is due to the fact that, in the absence of pH trapping, the concentration of CQ in the DV would be equivalent to the external concentration, which for CQ at its IC_{50} is ~20 nM. At this low concentration, an insignificant amount of CQ-Fe(III)PPIX complex is able to form based on the measured Fe(III)PPIX binding constant (K) of CQ of 3.3 x 10^5 M^{-1} and a free haem concentration ([Fe(III)PPIX]) of 20 µM (*Eq. 1.1*). Therefore, Fe(III)PPIX binding, which accounts for the majority of the cellular accumulation, requires pH trapping to achieve a sufficient drug concentration to form a Fe(III)PPIX-CQ complex in the DV.[123]

$$K = \frac{[CQ{-}Fe(III)PPIX]}{[Fe(III)PPIX][CQ]} \qquad \textbf{Eq. 1.1}$$

1.8 Haemozoin and β-haematin

Although malaria pigment, now known as haemozoin, has been studied for centuries, the exact mechanism by which it forms and by which its formation is inhibited, remains unknown. Given that many successful antimalarials act by inhibiting this process, there has been much research interest surrounding ferriprotoporphyrin IX (Fe(III)PPIX), the structure of haemozoin and techniques for producing haemozoin extracellularly, over the past several decades.

1.8.1 The structure of haemozoin and β-haematin

In 1717, an Italian physician, G. M. Lancisi reported a brown discolouration of the liver, spleen and brain of patients infected with malaria.[126] However, it was more than a century later that this dark brown pigment was connected to malaria infections, after being found by the German anatomist, J. H. Meckel, in the blood and spleen of a mentally ill patient.[127,128] The structural identity of malaria pigment remained unknown for over six decades. Originally proposed to be melanin, the pigment was identified as being of haematin origin in 1911 by W.H. Brown.[129] However, it was believed that it could not be pure haematin, but rather a Fe(III)PPIX-protein complex.[130] Until the late 1980s, the general belief was that the parasite detoxified haem into a crystalline complex with a specially synthesised protein, as the purest preparations of malaria pigment contained a glycine-rich polypeptide.[131]

In 1987, isolation of pure haemozoin from mouse erythrocytes infected with *Plasmodium berghei* NYU-2 parasites was achieved by Fitch and co-workers, via extensive washing.[132] The necessary removal of both protein and lipid contaminants was carried out by breakdown of the crude material with a non-specific protease and extraction in chloroform/ MeOH respectively. The pure material was much less soluble and contained a lower percentage of oxygen than that of haematin purchased from Sigma. Soon after, Slater *et al.*[133] revealed further structural details via characterisation of haemozoin using a range of spectroscopic techniques and by comparison with synthesised haemozoin, βH. An absorbance peak at about 650 nm was observed in both the synthesised and parasite-extracted haemozoin, characteristic of an aqueous suspension of the intact substance which

had been reported previously by Fitch *et al*.[132] Furthermore, the elemental composition data for βH was similar to haemozoin and also consistent with the previous report. Solubilising haemozoin in 0.1 M NaOH resulted in UV-vis absorbance and mass spectrum (MS) data, which were identical to haematin, confirming that haemozoin was closely related to haem, without a protein component. However, the chemical structure was elucidated when Fourier-transform infrared (IR) spectroscopy showed that the chemical nature of the haem carboxylate groups was significantly different for haematin and haemozoin. Specifically, characteristic IR bands from the C=O and C-O stretching frequencies, specifically assigned to a carboxylate coordinated to a ferric ion with characteristics very similar to unidentate acetate coordination, were observed in the haemozoin spectrum but not the haematin spectrum. This suggested that haemozoin involved direct coordination between a propionate of one haem and the iron (Fe) centre of another. Extended X-ray absorption fine structure (EXAFS) spectroscopy, which determines the local environment of a metal centre, was another technique used to characterise haemozoin. The data showed that the atom type and atom distances in the local environment of the iron centres in haematin and haemozoin differed significantly. The combined data strongly suggested that an iron-carboxylate linkage between two adjacent haem molecules allowed monomers to aggregate into an insoluble product, although this was originally thought to form a polymer (*Figure 1.23a*).[133] In fact, this was not the only misconception regarding the structure of haemozoin in the early 1990s. Electron spin resonance (ESR) spectroscopy, which analyses spin state and charge, had led to the mistaken conclusion that the central atom in haemozoin was a low-spin ferric iron (Fe(III)), and was only corrected when Bohle *et al.*[134] convincingly demonstrated that it is the high-spin Fe(III) species that is present in haemozoin. This was achieved using electron paramagnetic resonance (EPR) and Mőssbauer spectroscopy measurements, which were in agreement. The same group also showed, via powder X-ray diffraction (PXRD) patterns, that haemozoin is a microcrystalline material belonging to a centrosymmetric space group and is identical to βH.[135] This crystallographic data suggested that the haem units form antiparallel polymer chains, which hydrogen bond via propionate groups (*Figure 1.23b*). However, at the turn of the century, this proposal was also invalidated when Pagola *et al.*[136] analysed PXRD data obtained with synchrotron radiation by Rietveldt refinement enabling the crystal structure of βH to be solved. Instead of interacting polymer chains, a dimer pair was identified in a single unit cell. Two dimer pairs

form a hydrogen bond via the propionate group of one Fe(III)PPIX molecule from each dimer. These interacting units then stack via π-π interactions, resulting in 3-dimensional βH crystal (*Figure 1.23c*). Furthermore, given the chemical, spectroscopic[133] and crystallographic[135] evidence that βH is identical to haemozoin, these updated structural data were justifiably also valid for intracellular haemozoin.

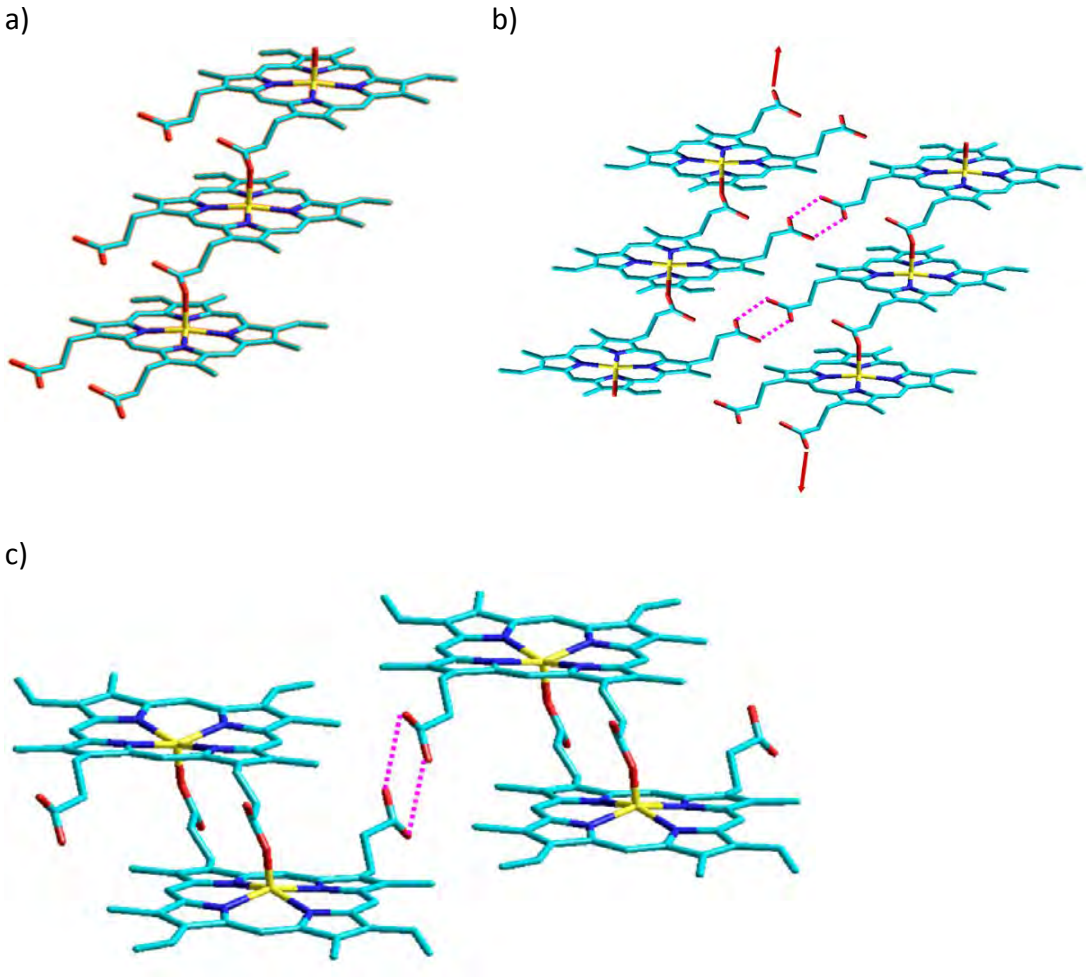


Figure 1.23 Evolution of the proposed structures of haemozoin showing a) a polymer b) the centrosymmetric polymer chains connected via hydrogen bonds and c) interacting dimers.

1.8.2 The mechanism of haemozoin formation

It is well accepted that the formation of haemozoin takes place in the acidic DV in order for the malaria parasite to avoid the toxicity caused by oxidative stress from liberated free haem. However, the exact mechanism by which haemozoin is formed remains elusive despite the proposal of several hypotheses including enzyme catalysis,[96] protein mediated formation,[137,138] spontaneous formation,[139] autocatalysis[140] or lipid mediated formation.[141-143] Haemozoin crystal formation mediated by lipids has recently received the most attention owing to strong evidence for this process. A recent study tested the feasibility of haematin crystallization from solutions that mimic either the aqueous phase or the lipid structures in the DV and found both aqueous and amphiphilic organic solvents to be crucial for formation of βH crystals with similar morphology to that of haemozoin.[144] Furthermore, haemozoin crystals, contained in nanosphere lipid droplets, have been observed during the trophozoite stage of the parasite using transmission electron microscopy (TEM) as shown in *Figure 1.24a* and *1.24b*.[145] Lipid associated with haemozoin was found to be made up of several fatty acyl glycerides including monostearic, monopalmitic, dipalmitic, dioleic and dilinoleic glycerols which when extracted, individually mediate βH formation. Hydrogen bonding between the Fe(III)PPIX propionate groups is favoured in a lipophilic environment and this has been proposed as the reason why the crystallisation process occurs at these hydrophobic sites. Recently, emulsions of monopalmitoylglycerol (MPG) and of a neutral lipid blend (NLB), consisting of MPG, monostearoylglycerol (MSG), dipalmitoylglycerol (DPG), dioleoylglycerol (DOG) and dilineoylglycerol (DLG) in a 4:2:1:1:1 ratio, have been shown to form droplets of varying size which control the βH crystal size during βH formation.[146] The specific composition of the NLB was chosen to mimic that associated with haemozoin isolated from *P. falciparum.* Under conditions of temperature, pH and ionic composition that mimic the parasite DV, kinetic studies demonstrated the efficiency of the NLB in mediating this process.[147,148,149] Crystals were found at the surface of the lipid droplets with the {100} crystal face in contact with the lipid, which supported the hypothesis that the neutral lipid surface acts as a template for βH crystal growth (*Figure 1.24c*).

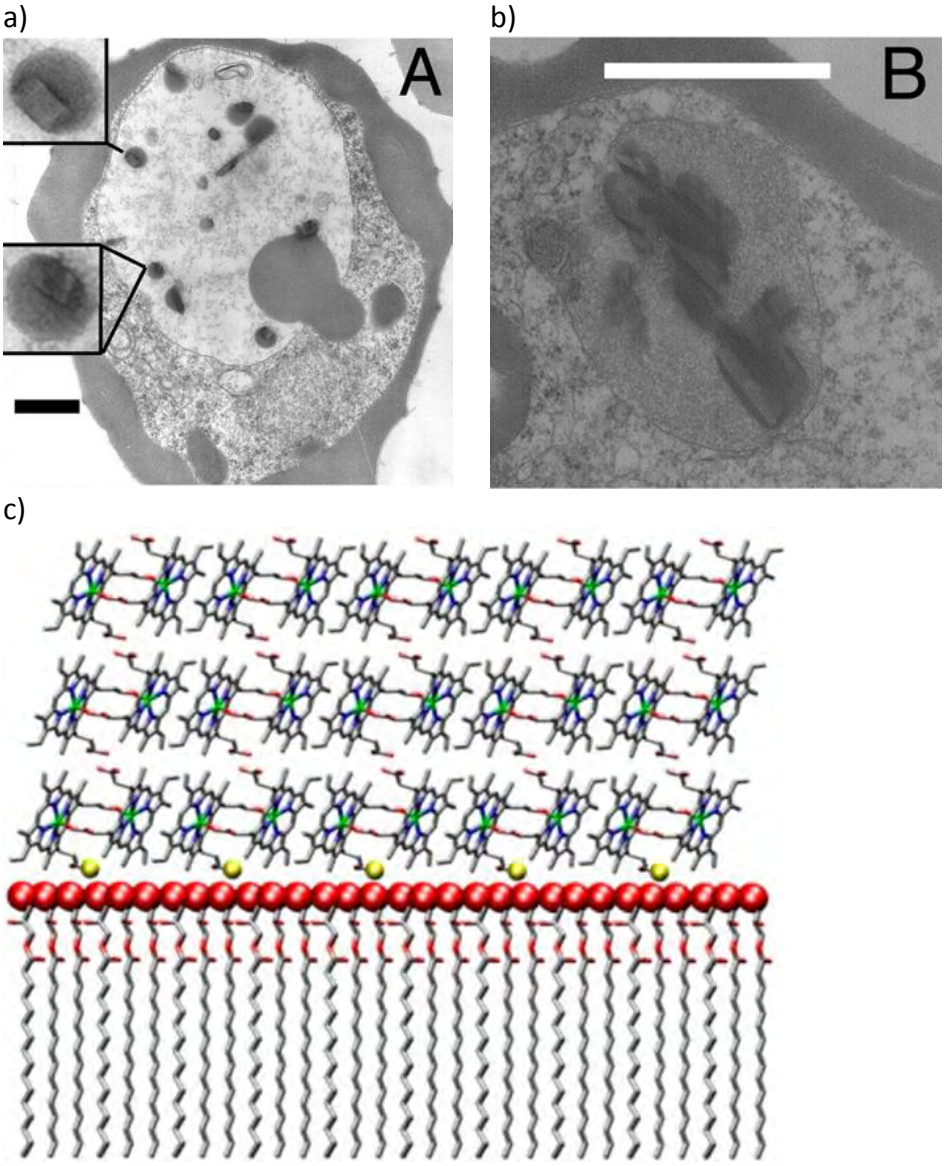


Figure 1.24 a) TEM of an early stage trophozoite showing haemozoin contained within neutral lipid nanospheres and b) Mature trophozoite containing larger crystals inside neutral lipid spheres within the parasite DV. Copyright © 2007, The Biochemical Society, Pisciotta *et al.*[145] c) The surface of a MPG lipid droplet as a template for βH crystal growth. Reprinted (adapted) with permission from Ambele *et al.*[146] Copyright © 2013, American Chemical Society.

1.8.3 The mechanism of haemozoin inhibition

Despite much research interest into the haemozoin inhibition mechanism, the specific way in which haemozoin-inhibiting drugs cause an increase in free haem and a decrease in haemozoin formation remains uncertain.[150] Elucidation of the inhibition process, especially for the quinoline antimalarials, would greatly assist with rational drug design. The inhibition of synthetic haemozoin (βH) by CQ and other 4-aminoquinolines has been studied extensively. This is despite the widespread resistance to these drugs, since there is no indication of modifications to haemozoin formation as the drug target.[68] Furthermore, the target on which they act is uniquely attractive since the biochemistry within the parasite is sufficiently different to that of the human host.[18] In 1992, Slater and Cerami proposed that CQ and QN inhibit a haem polymerase enzyme, blocking the formation of haemozoin.[96] However in the absence of trophozoite lysate, CQ was also able to inhibit βH formation and isolation of the haem polymerase enzyme was never achieved.[18,139] It was thus established that haemozoin crystal growth is a physio-chemical process and not enzyme-mediated, although proteins and/or lipids may be involved with initiation of the process *in vitro*.[140,151-153] In 1995 this led to the conclusion that CQ inhibited haemozoin formation by binding to monomeric haematin.[140] However, shortly after, it was conversely reported that quinoline antimalarials bind primarily to the μ-oxo dimer and not the monomeric form of haematin, leading to the hypothesis that a shift in equilibrium reduces the monomeric haematin available for haemozoin formation.[97] The formation of the μ-oxo dimer was confirmed very recently in 2014, when it was shown that CQ induces its formation at both the physiologically relevant pH 7.5 and at pH 5, to mimic that of the DV.[150] Furthermore, a complex with 2:4 CQ:Fe(III)PPIX stoichiometry in aqueous solution was observed using magnetic susceptibility, UV-visible absorption, magnetic circular dichroism (MCD) and diffusion techniques. Although this μ-oxo dimer complex likely forms as a result of the free haem concentration increase when haemozoin is inhibited, there is uncertainty regarding whether it is responsible for the inhibition process. In fact, Pagola *et al.*[136] suggested that the interaction of CQ with haematin cannot account for haemozoin inhibition since haematin is present at considerably higher concentrations than CQ in the DV. Studies which are consistent with this suggestion, report that either quinoline-haem complexes can be incorporated into haemozoin to terminate growth[137] or that the drug can intercalate

between the planar iron porphyrin groups on the smallest visible and hence fastest growing crystal face to inhibit its growth. The second hypothesis was proposed by Leiserowitz and co-workers.[154] based on the powder-crystal structure of βH.[136] This computational study modelled the structure of the haemozoin crystal, allowing CQ to be manually docked into the highly corrugated structure of the fastest-growing face, where it presented a convincing fit with the surface. This study was also able to explain previously unexplained findings, such as the need for the 7-chloro and the 4-amino groups for the βH inhibition activity of CQ. This extends to other quinoline antimalarials as well, since the diastereomer of QN, 9-epiquinine, and primaquine, which are both βH inactive were shown either to dock with the haemozoin crystal surface only by adopting a strained conformation or to be incapable of docking at all, accounting for their lack of inhibition activity. Very recently, another study demonstrated four surface binding sites on the haemozoin crystal which could accommodate a haemozoin inhibiting molecule, using atomic force microscopy.[144] They further showed that CQ adsorbs onto the terraces between haemozoin growth steps on the {100} crystal face. As a result, the mechanism whereby an inhibitor binds to the haemozoin crystal surface to restrict solute addition currently appears to be the most successful hypothesis.[113,155]

1.8.4 Crystal structures of drug-haem complexes

Although the reduction in haemozoin in the presence of an inhibitor is now believed to be as a result of crystal surface adsorption, as opposed to drug-haem complexation in solution, several attempts have been made to evaluate the specific interactions between Fe(III)PPIX and the quinoline antimalarials.[156] In 2008, de Villiers *et al.*[157] reported the first crystal structure for a complex of Fe(III)PPIX with an antimalarial drug, in this case, HF (*Figure 1.25a*). The structure, obtained by single crystal X-ray diffraction, revealed coordination of the hydroxyl group of HF to the Fe(III) centre and a π-π interaction between the phenanthrene ring and the porphyrin (*Figure 1.25b*). This allowed the authors to proposed structures for the cinchona alkaloids, QN, QD and their disastereoisomers, 9-epiquinine and 9-epiquinidine in order to explain their relative βH activities. This was justified by the similarity of the UV-vis absorbance spectra of Fe(III)PPIX complexed with HF or QD, indicating comparable interactions. Their suggested structures were validated in 2012, when Gildenhuys *et al.*[158] reported single crystal structures of QN and QD complexed with

(Fe(III)PPIX). The observed interactions which gave rise to activity encouraged the authors to suggest a pharmacophore scaffold for the aryl methanol compounds which contains three components; a secondary alcohol capable of Fe(III) coordination, an inflexible tertiary amine for hydrogen bonding and a fused aromatic ring system for π-π stacking interactions (*Figure 1.25c*). Knowledge of these key features for strong haem binding have potential to aid in rational design of novel antimalarials.

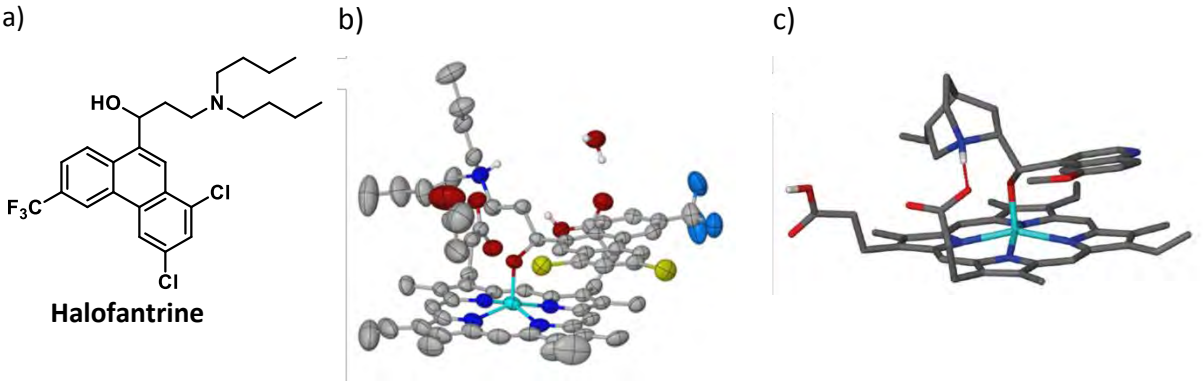


Figure 1.25 a) The structure of HF, b) the single crystal structure of Fe(III)PPIX-HF, Copyright © 2008 Elsevier Inc, de Villiers *et al.*[157] and the c) the single crystal structure of Fe(III)PPIX-QD, Copyright © 2012, American Chemical Society, de Villiers *et al.*[158]

1.8.5 βH inhibition assays

The continuous search for new antimalarials requires improved methods for identifying active compounds. In particular, target-based drug discovery demands assays which are efficient and which mimic the biological environment appropriately. Several assays for inhibition of βH have been developed. The first, reported by Egan *et al.*[139] in 1994 was a qualitative assay in which infrared spectroscopy was used to confirm the presence of free or crystallised haematin based on characteristic peaks. Here, the βH was formed in solution containing a high-concentration of acetate buffer, however, the amount of βH formed could not be quantified.[139] Another assay also used this method for crystal formation, but detection was via the differential solubilisation of haematin and haemozoin in DMSO and the assay required a long 18 h incubation period.[102] Dorn *et al.*[140] reported the formation of βH in lower concentrations of acetate buffer and used radio-labelled haematin which could

be incorporated into preformed haemozoin and then quantitatively determined using scintillation counting. Kurosawa *et al.*[159] and Sullivan *et al.*[137] also employed radio-labelled haematin, however this method suffers many disadvantages such as requiring skills, special equipment and resources that are limited in most laboratories as well as being expensive and subject to legal restrictions in some countries. Other methods to initiate crystal formation involved trophozoite lysate,[96] purified βH[140] and a combination of acetonitrile extracts of trophozoite lysate and βH.[140,141,159]

Until 1999, these assay methods for βH inhibition did not account for the possible role of lipids in haemozoin formation.[141,142] Thereafter, assays utilising lipids and detergents for the formation of βH were reported, the first by Fitch *et al.*[143] since replicating the acidic and hydrophobic environment of the DV *in vitro* is less demanding with the these mediators.[160] Initially, detection of crystallisation was made possible via the differential solubility of haematin in aqueous sodium dodecyl sulfate (SDS) solution.[143,105] While the neutral lipid blends had been shown to promote crystal formation *in vitro,* they were comparatively expensive and involved time-intensive preparation.[161] As a result, mimetics of lipid mediators, in the form of surfactants, were employed to promote formation of βH in an efficient manner, after the individual components of the blend were shown to mediate βH formation to some extent.[145]

"Surfactant" is an abbreviation of surface-active agent. Examples include detergents, dispersants, foaming agents and emulsifiers. They are able to lower the minimum energy required to create an interface (interfacial free energy) within a system of immiscible phases as they adsorb onto the interface of that system, decreasing the surface tension. Amphiphilic surfactants characteristically consist of both a hydrophobic tail group and a hydrophilic head group which can either be anionic, cationic, non-ionic or zwitterionic. A zwitterionic detergent is neutral overall owing to oppositely charged chemical groups (eg. +1 and -1) which form the head group. In water, the system acts by minimising contact between the hydrophilic water molecules and the hydrophobic tail groups of the surfactant. The result is a layer of specifically oriented surfactant molecules at the surface of the water, with the hydrophilic groups pointing into the aqueous phase and the hydrophobic tail directed away.[162] Surfactants can also form spherical aggregates known as micelles above the critical micelle concentration (CMC) which varies with surfactant type, temperature,

pressure and the presence of other substances.[163] The surface of the micelle consists of the hydrophilic head groups in contact with the aqueous phase, while the tail groups make up the internal structure of the sphere where they are shielded from water molecules.

In 2007, Huy *et al.*[164] used the detergent TWEEN® 20 to mediate βH formation in their inhibition of haem crystallisation assay, which was designed for HTS. This was an important simplification of the assay as it reduced the incubation time to 4 h and was much cheaper than the neutral lipid blend while producing relatively similar quantities of βH under the same conditions. The shortcoming for this assay was in the IC_{50} values of known βH inhibitors that were ~10 fold higher for CQ when compared with the values for the neutral lipid blend mediated assay. This result demonstrated that TWEEN® 20 was not the most accurate lipid blend mimic.

Optimisation of the detergent-mediated assay was reported by Carter *et al*.[160] in 2010 in which the Nonidet P-40 detergent (NP-40) was determined to be the closest imitator of the NLB, resulting in the largest yield of βH crystals. In addition, an assay for haemozoin inhibition with the known inhibitors CQ and AQ gave comparable IC_{50} values when using the NLB and NP-40. These authors studied a variety of lipophilic mediators, differing in chemical structure and ionic character, in order to optimise the platform for βH formation. Specifically, the detergents investigated included Triton™ X-100, NP-40, TWEEN® 20 and TWEEN® 80, which are non-ionic and the zwitterionic and anionic detergents, CHAPS and SDS respectively. Using ratios of the CMCs for each detergent, the appropriate initial assay concentration could be approximated by comparison with the TWEEN® 20 concentration employed by Huy *et al*.[164] Firstly, the detergents were tested for their ability to nucleate crystal growth and the βH products characterised by infrared spectroscopy (characteristic stretches at 1664 cm^{-1} and 1211 cm^{-1}) and powder X-ray diffraction. Those that successfully produced βH which matched that of the Bohle dehydrohalogenation βH product, namely TWEEN® 20, TWEEN® 80 and NP-40 (≥69% yields), were subjected to concentration dose response experiments using known haemozoin inhibiting antimalarial drugs.[165] The anionic and zwitterionic detergents were ineffective mediators since they did not mimic the hydrophobic nature of the lipids as well as the non-ionic detergents. Overall NP-40 was found to be the most suitable detergent since it gave a similar βH yield to the lipid blend and comparable IC_{50} values for AQ and CQ as shown in *Table 1.1*. The NP-40 concentration

was varied and the amounts of βH produced under assay conditions were plotted (*Figure 1.26*), showing an optimised concentration of 30.6 μM. In contrast, the neutral lipid blend was highly efficient at promoting βH formation over a wide concentration range, demonstrating the importance of optimising the lipophilic mediator concentration.

Table 1.1 Data collected by Carter *et al.*[160] for lipophilic mediators of βH formation. NP-40 as a mediator was shown to give a good βH yield and IC_{50}s of the standards compared well to that of the neutral lipid blend assay.

Lipophilic Mediator	Mediator Conc. (μM)	βH Yield (%)	Amodiaquine IC_{50} (μM)	Chloroquine IC_{50} (μM)
Neutral Lipid Blend	50	75	23.07	85.26
NP-40	**30.6**	**74**	**25.73**	**50.99**
TWEEN® 20	9.77	71	316.3	262
TWEEN® 80	14.9	69	201.8	195.7
SDS	1140	10	225.4	245.6
Triton™ X-100	58.6	7.4	N/A	N/A
CHAPS	1466	7.0	N/A	N/A

a) b)

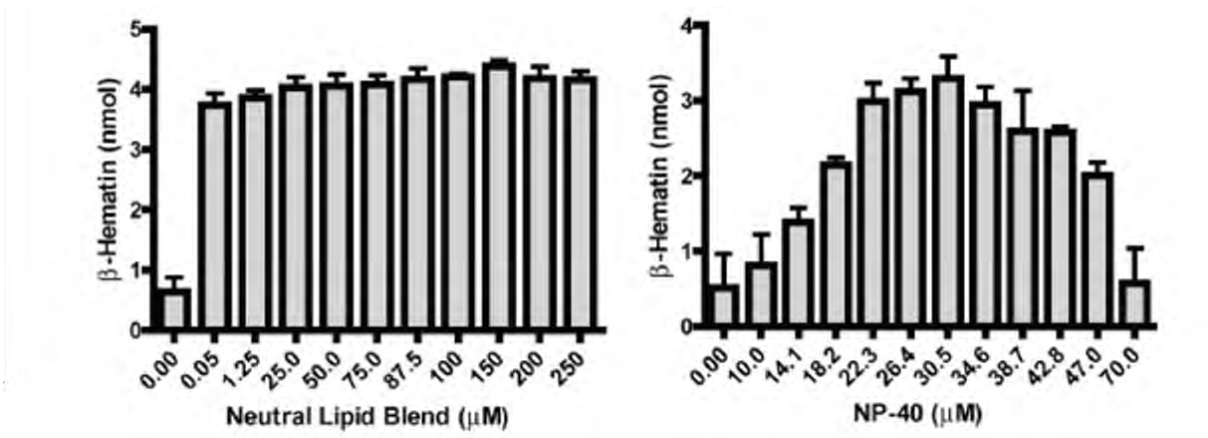


Figure 1.26 a) The NLB forms βH with similarly high yields over a range of concentrations and b) the NP-40 detergent forms the maximum amount of βH at 30.6 μM. Copyright © 2014 Bentham Science, Cater *et al.*[160]

The development of the NP-40 assay was intended for use in HTS in order to identify novel βH inhibiting compounds. In such a project, both synthetic and natural products including organic and inorganic molecules are typically screened. It is therefore a requirement that the βH inhibition assay can be applied to compounds with a wide range of physical properties. These properties include logP, PSA, hydrophobic surface area and hydrogen bond donor/acceptor parameters which all contribute to the solubility of the compound. Since not all compounds are water soluble, a study was carried out in order to evaluate assay performance with possible interfering solvents at varying concentrations. Several chelating agents, thiol groups, pigments and media extracts were also assessed to increase confidence that a "hit" in the HTS was not a result of a common interferent.

1.8.6 Detecting the formation of β-haematin

Identifying compounds which inhibit the formation of βH requires a method for detecting either the amount of βH crystals formed or the amount of non-crystallised free haematin remaining. There are several techniques for characterising βH, although many are not appropriate for HTS. One such method is IR spectroscopy,[166] which not only requires 48 h of drying time but also does not quantitatively measure βH formed, as is the case with elemental analysis and X-ray powder diffraction, other detection techniques used in the 1990s.[166] In 1996, Sullivan *et al.*[137] reported a method for measuring the amount of βH formed. The process required several repetitions to purify the βH crystals via centrifuging and washing with 2.5% sodium dodecyl sulfate (SDS) buffered with 0.1 M sodium bicarbonate (pH 9.1). The purified βH then had to be dissolved in 0.1 M NaOH and the UV-vis absorbance read at 400 nm after which the extinction coefficient of 10^5 could be used to determine the haematin concentration. This was a very time consuming and large-scale process making it unsuitable for HTS. Later in 2000, Parapini and co-workers[102] published a technique for quantifying haematin, which took advantage the ability of DMSO to dissolve haematin but not βH. The absorption of the haematin solution was then measured against a standard curve for concentration analysis. The disadvantage of this technique was the time consuming overnight reaction and double centrifugation steps. A solution to the problem was developed four years later by Ncokazi and Egan[103] using an established method for formation of a ferrihemochrome complex.[167] The method involved the addition of pyridine

which forms a low-spin complex with haematin but not with βH. They monitored the absorbance spectrum of the pyridine-Fe(III)PPIX complex at 405 nm and found that the Soret band increased in intensity, sharpened and shifted from 389 to 404 nm (*Figure 1.27a*). This indicated a change in the speciation of the haematin which is initially in an aggregated state, since it forms a dimer in aqueous medium and then becomes monomeric with the addition of pyridine.[168] This monomerisation is useful since absorbance is directly proportional to concentration for monomeric porphyrins, and indeed Beer's law was obeyed for concentrations up to 30 μM haematin at 405 nm (*Figure 1.27b*). In contrast, aggregated porphyrins tend to deviate from Beer's law because the monomer is favoured at low concentrations and aggregation is favoured at higher concentrations.[169] The appearance of the weak and broad Soret and Q-bands in *Figure 1.27a* (dimer form) is due to the splitting of the absorbance bands. This splitting arises during excitation when interactions occur between the transition dipole moments of the two porphyrins that form the dimer. When pyridine is added the Q-band intensifies and has an absorbance maximum at 527 nm. This is characteristic of a low spin complex,[170] suggesting a ligand substitution reaction has taken place, in which the axial water ligand is replaced by two pyridine molecules.

a) b)

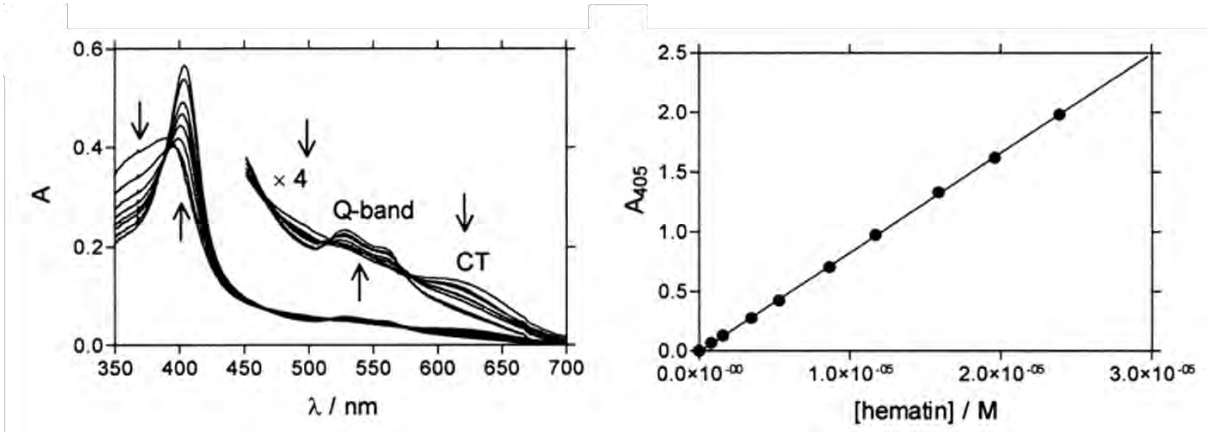


Figure 1.27 a) Haematin in aqueous solution with increasing pyridine concentration (shown by arrows). CT = Charge Transfer Band and b) Plot of the Soret band of hematin at 405 nm in 5% (v/v) aqueous pyridine, pH 7.5 (20 mM HEPES) showing Beer's law is obeyed over the concentration range (r^2 = 0.9998).[103] Copyright © 2005 Elsevier, Ncokazi *et al.*[103]

The increase in intensity of the Soret band upon addition of pyridine is favourable for the application of detecting free haem because the sensitivity of the haem-pyridine complex is improved relative to aqueous haematin. Another advantage of this method is that the pyridine is likely to displace other coordinated substances which would alter the molar extinction coefficient of Fe(III)PPIX making it impossible to precisely quantify the free haematin.[103]

The conditions, such as concentration and pH, under which pyridine coordinates with haematin were optimised. At large concentrations of pyridine and at high pH, pyridine is able to react with haematin as well as βH since the crystal is less stable under basic conditions. However, at neutral pH and lower concentrations, the energy required to disrupt the crystal and break the hydrogen bonds of the propionate groups is greater than the energy gained from pyridine-Fe coordination. Thus pyridine does not react with βH under these conditions. At neutral pH the haematin ligand is H_2O, a weak field ligand which is easily replaced by pyridine, a strong field ligand, resulting in the reaction of pyridine with non-crystallised haem. The ideal conditions were found to be 5% (v/v) pyridine at pH 7.4 (20 mM HEPES buffer) at room temperature (RT). The reaction is fast and the colour develops instantly, maintaining intensity for up to 24 h, suggesting that pyridine does not disrupt βH over time.

1.9 High throughput screening in antimalarial drug development

The many inherent difficulties in discovering and developing novel antimicrobial drugs have resulted in no new chemical classes of antimalarials being introduced into clinical practise since 1996. In addition, resource commitment towards antimalarial research has been relatively lacking in both public and private institutions.[171] The shortage of novel, publicly accessible antimalarial scaffolds, in particular those with varying mechanisms of action has had a negative impact on drug discovery efforts.[172] An open-source model for research and development (R&D) has been adapted by organisations known as public-private partnerships, creating a cost effective way of finding new therapies for neglected diseases, with promising results. Open-source genomic and metabolomic information is constantly

getting richer with terabytes of data being added to the pool while chemical and structural data are much more limited. Availability of open-source tools were, until recently, abundant in bioinformatics but were lacking in chemistry owing to the dominance of commercial software. Fortunately, initiatives such as the launch of PubChem in 2004 and tools such as eMolecules and Jmol are revolutionising the way in which chemical information is shared. Furthermore, the latest computerised tools are able to efficiently mine data and search the literature, and *in silico* modelling software allows for toxicity prediction, visualisation of binding sites and metabolic network mapping.[173] As a result of this progress, tools for accessing and analysing HTS data previously available only to industry, can now be obtained by non-profit scientists. However, the scarcity of data remains a problem for open-source R&D. HTS provides a way of rapidly increasing the structure-activity data available for medicinal chemists as a starting point for drug discovery. This is particularly important for antimalarial drug development, since antimalarial therapies, past and present, are based on a limited number of chemotypes (e.g. the quinolones, cycloguanils, tetracylines and sesquiterpene lactones) which often leads to cross-resistance.[174] Owing to the relatively large amount of research on quinoline-based antimalarials, there is an urgent need for structural diversity in compounds active against *P. falciparum.* HTS provides a way of finding novel, diverse scaffolds which would otherwise not be investigated as potential antimalarials.

1.9.1 Phenotypic vs target based approaches

Advances in the construction of compound libraries as well as the increasing knowledge of drug targets have encouraged parallel advances in HTS. This has led to a new discipline focused on assay design, given that most regularly used assays are unsuitable for industrial scale screening. Assays should be optimised for speed, efficiency, signal detection and low reagent consumption. With these improvements, screening in relatively high throughput has been made possible with automated liquid handling and signal detection systems.[175]

For the past century, two forms of screening have generally been employed for early stage drug development, namely phenotypic screens and target-based screens. Phenotypic screens measure the effect of a compound within a cell or tissue, without prior knowledge

of the mechanism of therapeutic action or the number of druggable targets. In 2011, a study by Swinney and Anthony reported that more first-in-class small molecules were found via phenotype-based screens between 1999 and 2008 than other screening methods.[176] However, determining the target or in some cases multiple targets of the identified hits, has typically been a difficult, time-consuming or even an impossible process.[177] In the 1980s, advances in molecular biology allowed screens against defined targets to be developed. Target-based screening has been the more popular over the past few decades.[178] This is particularly favourable in the case where large-scale screening against a specific target, using a biologically relevant assay, is more efficient and less expensive than testing hundreds of thousands of compounds on the whole-cell. It is also possible to combine the techniques by employing target-based before phenotypic screening. In this case, once target-based screening has been carried out to identify a much smaller fraction of the total library as hits, whole-cell testing can then be conducted to take into account other factors such as cell permeability and target-site accumulation.

1.9.2 Recent antimalarial HTS efforts

Target-based HTS of diverse libraries relies on the knowledge of specific *P. falciparum* inhibiting pathways for appropriate assays to be developed. In 2004, GlaxoSmithKline (GSK) screened compounds against *Plasmodium* LDH and identified azole derivatives, which were selective for *Plasmodium* LDH over human LDH. Unfortunately *in vitro* activity could not be linked to enzyme inhibition and they concluded that LDH was not druggable.[179] Haemozoin inhibition is an alternative, valid pathway since this process has not been directly disrupted by the structure-specific efflux mechanism which causes drug resistance in the parasite (see *Section 1.6*).[60] Synthetic haemozoin, βH, can be efficiently and reliably synthesised in the laboratory with minimal equipment using a variety of methods.[139,140,160,165] Furthermore, detecting the formation of βH is easily achieved through a number of different techniques as discussed in *Section 1.8.6*.

The first target-based HTS project for βH inhibitors was carried out by Kurosawa *et al.*[159] on over 100,000 compounds from several compound libraries and reported in 2000. They employed an assay which incorporates radiolabelled haematin into the βH crystal, where

the ability of known quinolines to inhibit this process was correlated with their ability to inhibit parasite growth. Initiation of crystal formation was via acetonitrile extracts from *P. falciparum* trophozoite lysates, the major components of which were lipids. At the time, this was the most appropriate assay for the robotic handling involved in HTS. Compounds were obtained from the Roche, SPECS and other diverse compound libraries as well as the NIH Clinical Collection (NCC), where they were isolated from microbial broths. Many of the compounds were initially screened as cocktails of ten samples and then, if a hit was identified, screened individually. A total of 45 non-quinoline compounds were identified as hits (IC_{50} of <50 µM) with IC_{50}s well below the CQ control of 80 µM, and were primarily triarylcarbinols, piperazines, benzophenones, imides, hydrazides, indoles and isoxazoles. However, only four of these compounds exhibited activity <5µM in both the CQ-sensitive NF54 and CQ-resistant K1 strains of *P. falciparum*, resulting in a very low overall hit rate.

An improvement in the hit rate of target-based screening for haemozoin inhibiting antimalarials was reported in 2009 by Rush *et al.*[180] They took advantage of the colorimetric pyridine-based ferrihemochrome assay developed by Ncokazi and Egan,[103] requiring no parasite-derived reagents or radioactive materials. Rather, 9.7 M acetate was used to initiate βH formation. This method was adapted into a robust HTS assay in a 384-well plate format for robotic handling. A total of 16,000 diverse compounds were screened, of which ~600 were hits with IC_{50}s of <220 µM, well below the CQ control of 374 µM in this assay. Subsequently, 17 compounds showed activity against the CQ-sensitive 3D7 and CQ-resistant Dd2 strains of *P. falciparum*, although only seven hits were active below 5 µM.

Large HTS projects for new antimalarials have also recently been carried out as phenotypic screens. Two groups in industry (Novartis[181] and GSK[171]) as well as an academic group (St. Jude Children's Research Hospital[172]) published their screening results for compounds directly active against *P. falciparum* whole cells in 2008 and 2010. These data are publically available from (https://www.ebi.ac.uk/chemblntd/). Novartis screened approximately 800,000 compounds, reporting 5,655 hits with IC_{50}s < 1.25 µM. Priority was given to 200 distinct chemical classes for hit-to-lead investigation. GSK screened approximately 2 million compounds from their libraries and identified ~13,500 compounds, now referred to as the Tres Cantos Antimalarial Set (TCAMS), with IC_{50}s <2 µM. Finally, at the St Jude Children's Research Hospital, ~300,000 compounds were screened against multiple strains of *P.*

falciparum, and 560 validated hits with IC_{50}s <2 μM and >10-fold selectivity against mammalian cells, were reported. Of these, 170 compounds with diverse structures were profiled *in silico* and 80% were predicted to have novel mechanisms and no cross resistance.

1.9.3 The scope for novel βH inhibitors discovered via HTS

The Novartis, GSK and St. Jude's screens were primarily focused on finding novel active compounds intended for development towards clinical antimalarials. However, target-based screening for haemozoin inhibitors offers the additional opportunity to use the data, both active and inactive, to probe mechanistic aspects as well as to construct SAR models. This is especially important regarding non-quinoline chemotypes, since apparent differences between trends in the well-studied quinoline series and that of new series, are likely to elucidate further details of *in vitro* haemozoin inhibition and drug toxicity mechanisms as well as aspects of drug accumulation. For example, Egan *et al.*[182] showed that quinoline analogues that inhibit βH formation but have low parasite activity, can be derivatised into potent antimalarials via the addition of basic side chains for pH trapping. However, it is not known whether this phenomenon is applicable to other chemotypes. In addition, several studies have reported correlations between βH inhibition activity and parasite growth inhibition activity for quinoline and related compounds, yet few series of non-quinoline βH inhibitors have been analysed using this method. Combrinck *et al.*[113] reported an increase in free haem in whole-cell parasites with several quinoline antimalarials, CQ, AQ, QN and MF (*Section 1.7.3*), whereas the extent of free haem accumulation in parasites inoculated with diverse, non-quinoline, βH inhibiting scaffolds has not been investigated.

Another advantage of βH inhibitors is their inherent structural suitability for short-sequence synthesis. This has favourable implications, not only for efficient production of analogues, but also for the potential future of these compounds as antimalarials. Because 90% of malaria-related deaths occur in sub-Saharan Africa, where most regions are burdened with poverty and low-levels of education,[10] designing drugs which can be quickly and inexpensively manufactured is a rational priority. In 2013, 69 million children with malaria did not receive ACTs, which are currently the most effective way of preventing malaria caused by *Plasmodium* parasites.[183] Unfortunately, artemisinin, the key ingredient of ACTs, costs $2.40 per dose, which is prohibitively expensive in most of the malaria-prevalent areas.[184] On an industrial scale, the natural product artemisinin is extracted from the

wormwood plant, *Artemesia annua,* and ACT supply depends heavily on the agricultural production of artemisinin.[185,186] However, problems relating to crop quantities, often as a result of unpredictable climates or natural disasters, have led to shortages and price fluctuations. Despite a decade's worth of research to solve these problems, no impact on the global supply chain has been achieved and little focus has been given to developing a *de novo* total synthesis of artemisinin.[183] At best, an enantioselective five-pot process has been achieved that still demands five purifications as well as toxic, hazardous and environmentally unfriendly reagents. Although this is a cost-effective synthesis, large-scale production of artemisinin via synthesis is still under investigation.[183,187] As such, natural product cultivation and extraction remain essential to meet world-wide antimalarial demands.[188]

The complex chemical structure of artemisinin, which has resulted in long and difficult enantioselective syntheses, has been a liability for investigation into cost-reducing approaches.[187] On the other hand, many βH inhibiting compounds can be synthesised via rapid, simple routes which do not require stereo-selectivity. This is due to the nature of their planar aromatic rings, which appear to be optimal for π-π stacking with haem. Antimalarials that can be synthesised from readily available starting materials using cheap and safe reagents would logically result in improved distribution of drugs to poverty-stricken areas. This would be particularly relevant when combined with a synthetic peroxide antimalarial, such as the ozonide OZ439 drug candidate currently in phase IIb clinical trials with Medicines for Malaria Venture (MMV).[189]

Virtual screening has become an important technique for drug discovery but in malaria, is hindered by the lack of well-validated targets and as a result, few successful examples of this approach exist.[178] An *in silico* technique using Bayesian statistics to create models, involving both active and inactive HTS data, has been utilised in recent literature to enrich hit rates in subsequent screens. Software is now available with a built-in Bayesian classifier, based on Bayes' theorem, which has aided in efficient identification of new actives.[190,191] Since 2004, this method has been applied for modelling kinase inhibitors,[192,193,194] *Escherichia coli* dihydrofolate reductase inhibitors,[195] G protein-coupled receptor (GPCR) ligands,[196] histidine H4 antagonists, phosphodiesterase (PDE) inhibitors, metabotropic glutamate receptor antagonists,[197] oestrogen receptor inhibitors as well as

metalloproteinase, nitric oxide synthase and other non-kinase enzyme inhibitors.[194,198] These are key targets for diseases such as cancer and Alzheimer's.[199] In addition, Ekins *et al*.[200,201] have employed public whole-cell *Mycobacterium tuberculosis* (Mtb) HTS data to demonstrate a 10-fold enrichment on typical hit rates when compounds are prioritized using Bayesian models. Although they were able to apply Bayesian models to phenotypic data, the other literature reports applying this statistical method for drug discovery, appear to include data for specific drug targets. Indeed, Bayesian models have not been applied to antimalarial HTS data, most likely as a result of the lack of sufficient target-defined activity data.

With the recent developments of improved assays for identifying βH inhibitors, haemozoin target-based HTS of large diverse libraries offers a unique and compelling opportunity to acquire urgently needed data for further insight into the problem of malaria.

1.10 Aims and Objectives

1.10.1 Aims

Given the desperate need for alternative antimalarial drug scaffolds, particularly for probing mechanisms aimed at improving activity and overcoming resistance, together with the suitability and convenience of haemozoin formation as a drug target, the aims of the project were to identify novel βH inhibiting compounds via HTS and to select appropriate chemotypes for investigation. The molecular scaffolds were chosen to be amenable to efficient synthesis of a small library of analogues in order to probe mechanistic aspects of activity and to uncover any possible SARs for each scaffold.

1.10.2 Objectives

Comprehensive lists of the specific goals for each component of this project are shown in the introduction to each respective chapter, however a summary of objectives is provided here.

To accomplish the aims highlighted in *Section 1.10.1*, the specific objectives of this project were:

1) To perform HTS on the Vanderbilt University Institute of Chemical Biology compound collection. This required carrying out the detergent-mediated assay on approximately 40,000 compounds, after which, hits were screened in a secondary assay to identify those which inhibit *P. falciparum* growth.
2) To analyse and investigate novel βH inhibiting scaffolds, particularly where some of the derivatives exhibited promising parasite inhibition potency, in order to select chemotypes for further exploration.
3) To identify the minimum pharmacophore for βH and parasite growth inhibition within each selected chemotype, via synthesis or purchase of deconstructed analogues.
4) To synthesise a small library of derivatives for each selected chemotype, in order to determine the βH and parasite growth inhibition activities.

5) Where appropriate, to evaluate the effect of attaching basic side chains believed to cause pH trapping and then to determine CARs.
6) To investigate structure-activity relationships (SARs) to determine the likelihood of future potential for the chosen scaffolds, either as mechanistic probes or as antimalarial drugs.
7) To create and evaluate a predictive model based on Bayesian probability using the combined structure-activity data, including all available HTS data as well as evidence from the synthetic libraries, to aid *in silico* HTS screening in the future.

CHAPTER TWO

HIGH THROUGHPUT SCREENING FOR NOVEL β-HAEMATIN INHIBITORS

2.1 Introduction

In an effort to discover novel βH inhibiting scaffolds active against *P. falciparum*, HTS was conducted on 43,520 compounds, a subset of the Vanderbilt University Institute of Chemical Biology (VU) chemical library. Validation of the βH inhibition assay for use in high throughput experiments was essential to the success of this project and hence, a number of quality-control analyses and a pilot screen were carried out prior to HTS.

2.1.1 Drift, Z' and edge effects in plate-based assays

Assays can be powerful tools for drug discovery, owing to their ability to mimic the intricacy of biological systems on a suitably fundamental level. They also have the potential to be scaled up for HTS. Unfortunately, the chemical complexity gained from this technique can be the unwanted cause of noise and variability, leading to errors that affect assay performance. The factors that can cause uncertainties include errors in addition of reagents to the plate, differential reaction rates due to uneven response across the plate, plate-to-plate variability, run-to-run variability, and edge effects.[202] Since compounds are usually tested only once in primary screens, accuracy and sensitivity is vital for identifying hits.[175]

In light of this, Carter *et al.*[160] validated the βH assay in a 96-well plate for performance and plate uniformity with checkerboard positive and negative controls. Using the NP-40 detergent, a drift response of <12%, significantly below the accepted threshold of 20%, was calculated.[203] The drift response is calculated via *Eq. 2.1.*

$$Drift = \frac{(mean\ abs.\ of\ column\ 1) - (mean\ abs.\ of\ column\ 2)}{overall\ mean} \qquad \textbf{Eq. 2.1}$$

This favourable drift response of <12% indicated that there was uniformity across the plate. Furthermore, linear regression of the negative and positive control resulted in slopes <2.5 × 10^{-3}, confirming negligible edge effects on the plate (*Figure 2.1*). In addition, Sandlin *et al.*[161] reported a favourable Z' value for the NP-40 mediated βH inhibition assay of 0.82, well above the 0.5 threshold value. This Z' is based on *Eq. 2.2* and indicates the degree of UV-vis absorbance error within repetitions of the controls.

$$Z' = 1 - \frac{3(\sigma_p + \sigma_n)}{\mu_p - \mu_n} \qquad \textbf{Eq. 2.2}$$

Where σ_p and σ_n refer to the standard deviation in the absorbance of the positive and negative controls respectively and μ_p and μ_n are the mean absorbance values for the positive and negative controls.

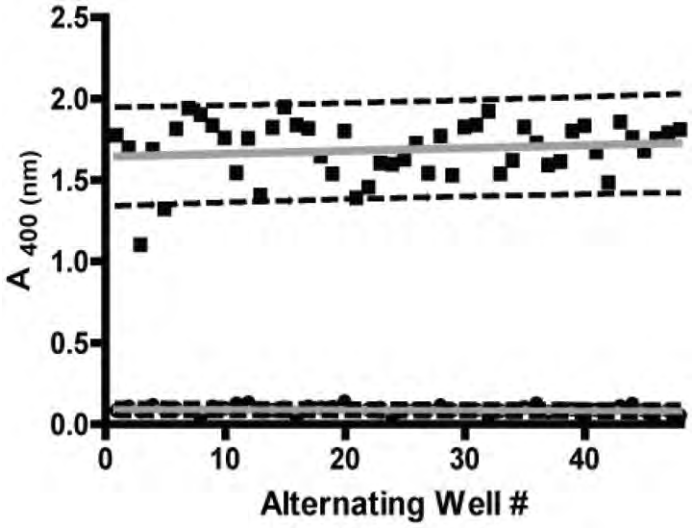


Figure 2.1 Absorbance data for the negative and positive controls in a checkerboard arrangement. Linear regression gave negligible slopes of <2.5 × 10^{-3} showing uniformity in the data across the plate. Copyright © 2014 Bentham Science, Carter *et al.*[160]

2.1.2 Pilot screen for βH inhibitors and antimalarials

In 2010, a pilot screen using the NP-40 assay was carried out at Vanderbilt University.[161] A total of 38,400 compounds at a concentration of 19.3 μM were tested for their ability to inhibit βH formation. This resulted in the discovery of 161 novel hits, none of which were previously identified as βH inhibitors. The hits were defined as those compounds which had a measured UV-vis absorbance at 405 nm above 3 standard deviations of the negative DMSO control on that plate. These hits were then screened against *P. falciparum* in a phenotypic assay utilising SYBR-Green I fluorescence dye to bind to parasite DNA, based on modification of methods published in the literature.[84] Of the 161 βH inhibitors identified, 48 also inhibited ≥90% of parasite growth at 23 μM, with eight showing nanomolar IC_{50}s. The success of this pilot screen prompted the HTS of the remaining 144,330 compounds in the library using a similar approach.

2.1.3 Specific goals

The specific goals for Chapter 2 were to:

1) Carry out HTS using the NP-40 based assay for βH inhibition on 43,520 compounds, a subset of the VU HTS library.
2) Perform data analysis using pre-defined cut-offs to identify hits.
3) Cherry-pick βH inhibitors from compound library source-plates and test them in a dose-response assay to measure activities and detect false positives.
4) Culture CQ sensitive D6 and multi-drug resistant C235 strains of *Plasmodium falciparum* in whole red blood cells for the phenotypic whole cell assay and identify compounds which show *in vitro* activity.
5) Analyse hit scaffolds for further lead-probe optimisation.

2.2 Results and discussion

2.2.1 Assay validation

A standard plate containing positive (AQ at IC_{100}) and negative control wells was analysed to assess the quality and reproducibility of the NP-40 assay (*Figure 2.2*). The mean positive control UV-vis absorbance value was 3.332 ± 0.043 and that of the negative control was 0.567 ± 0.096. Linear regression of the data confirmed negligible drift with slopes of $5x10^{-5}$ and $7x10^{-5}$ for the negative and positive data sets respectively and a drift response across the 24 microplate columns of 2.3%, well below the 20% threshold. These data also resulted in an affirmatory Z' value of 0.85, in good agreement with the 0.82 value determined by Sandlin *et al.*[161]

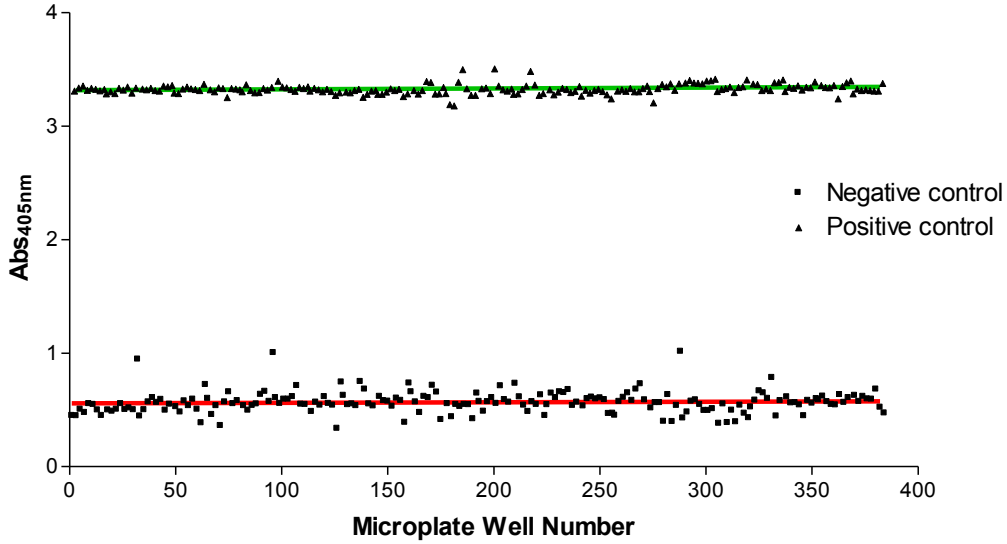


Figure 2.2 Positive and negative control wells on a 384-well microplate, arranged in an alternating checkerboard, revealed a favourable Z' value of 0.85 and low drift response of 2.3%.

Inhibitory concentrations, determined via dose response curves for both AQ and CQ were calculated to be 23.1 ± 2.6 µM and 66.3 ± 3.5 µM respectively (*Figure 2.3*). These IC_{50} values were within an acceptable range relative to previous literature which reported 21.0 and 53.0 µM for AQ and CQ respectively.[161]

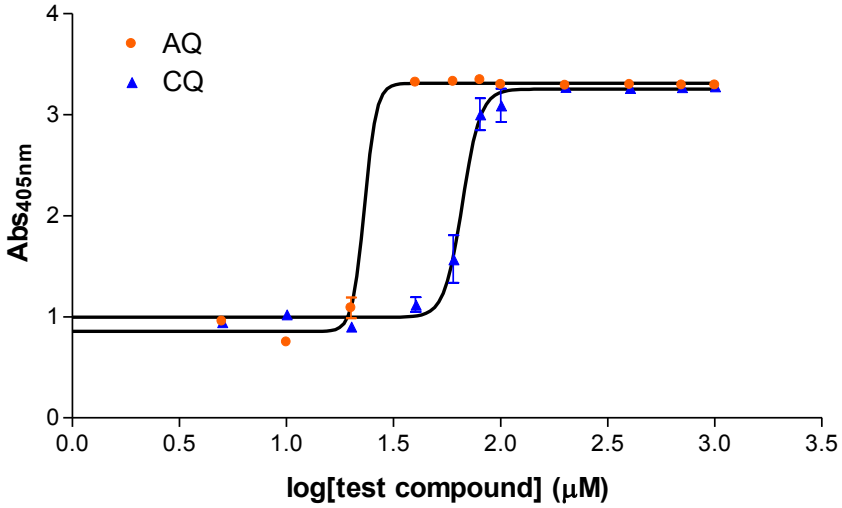


Figure 2.3 Standard antimalarials AQ and CQ inhibited βH formation in a dose-response manner.

2.2.2 Compound screening for βH inhibitors

The NP-40 assay for βH inhibition was performed on 43,520 compounds as part of a larger screen of 144,330 commercially available drug-like compounds from the Vanderbilt compound library.[204] Test compounds were screened at a rate of 320 per microplate and 20-30 microplates per week. Hits were considered to be those which, relative to the controls, inhibited βH formation above 50% at 19.3 μM, approximately the IC_{50} of AQ in this assay. The workflow for the HTS is shown in *Figure 2.4* where a plot of the absorbance data for a typical 384-well microplate highlights the 50% inhibition cut-off and identification of a hit (red circle). A total of 136 microplates were screened and the Z' score was calculated for each plate based on the first and last two columns containing controls. The average Z' across all the plates was calculated to be 0.51 ± 0.08, showing acceptable accuracy relative to the negative and positive controls. For this subset of the library, 758 compounds were classified as hits, resulting in a hit rate of 1.7%. The chemical scaffolds from this subset are shown in *Figure 2.5*. The most common chemotype found (35% of hits) belonged to the quinoline family, which includes quinolones, quinazolines and quinoxalines. The second most frequently occurring group (16% of hits) were the fused five-membered to six-membered rings containing nitrogen, oxygen and sulfur, namely the benzimidazoles, benzoxazoles, benzothiazoles and indoles, followed by the benzamide chemotype (10%).

These were of particular interest since there was little known of the βH inhibiting properties of these classes of compounds.

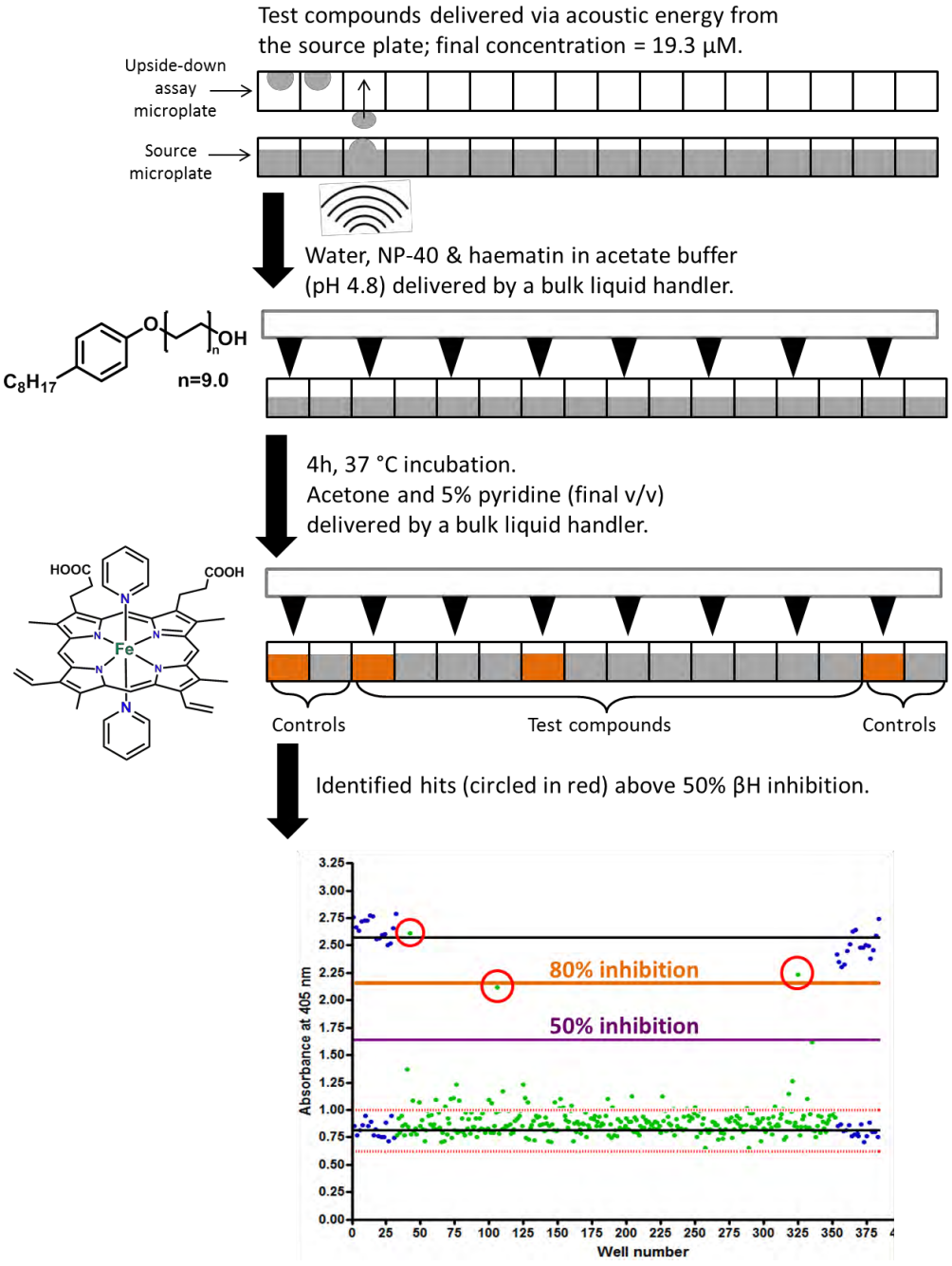
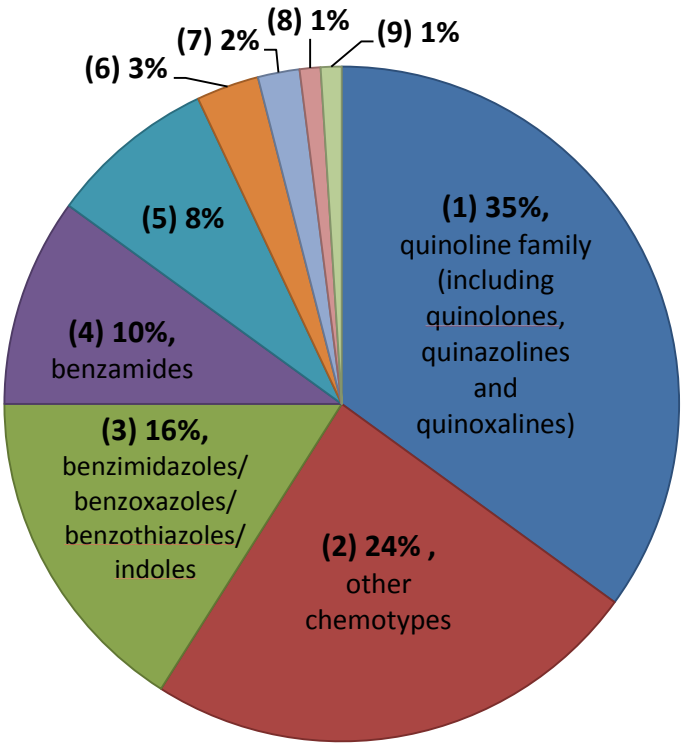


Figure 2.4 The HTS workflow for identification of compounds which inhibit βH formation in the NP-40 detergent (structure shown) assay. Hits were determined by measuring the absorbance of the haem-pyridyl complex shown in the figure, where inhibition values above 50% or 80% at 19.3 µM were chosen as the criteria for hit identification. The higher cut-off was used when analysing the entire library of 144,330 compounds.



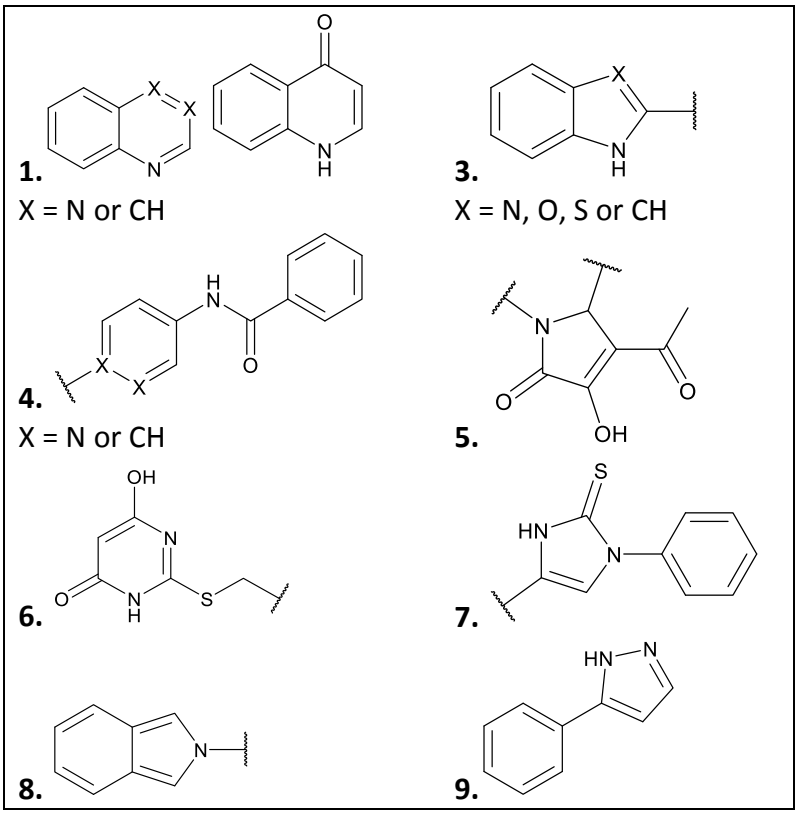
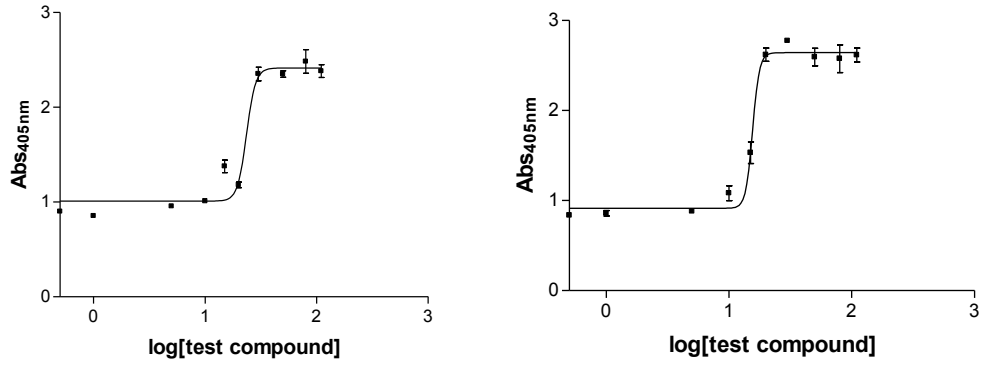


Figure 2.5 The most commonly occurring hit scaffolds with ≥50% βH inhibition.

Upon completion of the entire HTS project, where all 144,330 compounds were tested, the criterion for hit identification was raised to ≥80% inhibition. This cut-off was chosen so that only potent inhibitors were advanced to the next stage of the screening in order to better manage the numbers of compounds tested for their dose-response with respect to βH formation and eventually against the parasite. In order to ensure that comprehensive results are reported and because the entire database of compounds was analysed to determine appropriate scaffolds for synthesis, the following discussion refers to the ≥80% cut-off for the whole library.

Of the 144,330 compounds screened, 729 inhibited ≥80% of βH formation at 19.3 μM, resulting in a hit rate of 0.51%. In this preliminary screen, each compound was tested only once and as a result, false positives were expected to occur. It is important in any HTS project to confirm the validity of the active compounds, since errors are statistically likely to occur when dealing with such large numbers of compounds.[194] In light of this, these hits were selected and their IC_{50} values determined in a dose-response assay. This process allowed the false positives to be found and the potency of the true actives to be measured. A total of 530 compounds were confirmed hits (0.37% hit rate) with IC_{50} values ≤27 μM, indicating a false positive rate of 0.14%. Although this was ten-fold higher than the 0.016% false positive rate determined in the pilot screen by Sandlin *et al.*,[161] the overall hit rate of 0.37% was similar to the hit rate of 0.40% reported previously. Furthermore, 416 were more potent than AQ, with IC_{50} values <23 μM and all 530 compounds had IC_{50} values below that of CQ (66.3 μM). The IC_{50} values of the hits ranged from 0.17 to 26.97 μM with most compounds displaying IC_{50}s between 15 and 18 μM. *Figure 2.6* shows typical IC_{50} curves from the dose-response assay as well as the activity distribution.

a)



b)

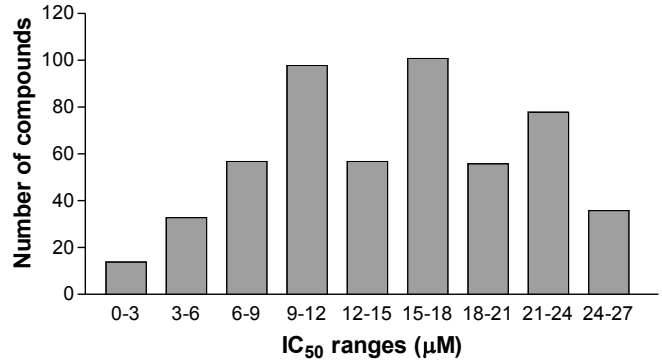


Figure 2.6 a) Dose-response curves for randomly chosen hit compounds from different plates, performed on different days and b) the distribution of βH inhibition activities for the 530 hit compounds.

Analysis of SARs for the hit compounds did not reveal any one-variable correlations between the βH IC_{50} and logP, molecular weight (MW), number of aromatic rings (#AR), number of H bond acceptors (#HBA), number of H bond donors (#HBD), number of rings (#R), number of rotatable bonds (#RB) or fractional polar surface area (FPSA). There was generally a wide range represented for each property within the βH hits (*Table 2.1*), however, the majority of the hits preferred a greater logP, MW, #AR, #HBA, #HBD, #R and FPSA, while also favouring fewer RB. This observation suggests that compounds may be less likely to interact with haem to inhibit haemozoin formation when they have large numbers of rotatable bonds. Having fewer rotatable bonds is believed to help with π-π interactions, which has been reported as one of the primary interactions between haem and quinoline containing antimalarials. This was shown from the crystal structure of the complexes of QD or QN with iron(III) protoporphyrin IX (Fe(III)PPIX), where the closest distance between the quinoline

ring and the the porphyrin pyrroles was calculated to be 3.3 Å, a characteristic indication of π-π interactions.[158]

Table 2.1 Ranges of physical properties for the βH inhibitors identified. The column on the right indicates that, for example, 75% of the compounds had a logP greater than 3.

Property	Range	% of βH inhibitors greater or smaller than a value
logP	-0.29 – 6.97	75% ≥ 3
Molecular weight	159.19 – 607.61	88% ≥ 300
Number of aromatic rings	0 – 5	95% ≥ 2
Number of H bond acceptors	0 – 10	92% ≥ 3
Number of H bond donors	0 – 4	60% ≥ 2
Number of rings	1 – 7	90% ≥ 3
Number of rotatable bonds	0 – 12	66% ≤ 6
Fractional polar surface area	0.03 – 0.56	68% ≥ 0.2

The hits were also analysed using Scaffold Hunter,[205] a program for interactive exploration of chemical space and visual analysis of biological activity data. The scaffold tree (*Figure 2.7*) showed 78 unique single rings which formed the most basic scaffolds of the 530 hits, emphasising the diversity of βH inhibitors. The six parent rings which comprised the largest number of child scaffolds, from most to least active on average, were imidazole, pyridyl, 2,3-dihydropyridin-4(1H)-one, pyrrole, phenyl and thiazole as shown enlarged in *Figure 2.7*. These represent a range of π-deficient and π-excessive rings. Interestingly, they were also reported to be amongst the most common aromatic scaffolds present in bioactive molecules.[206]

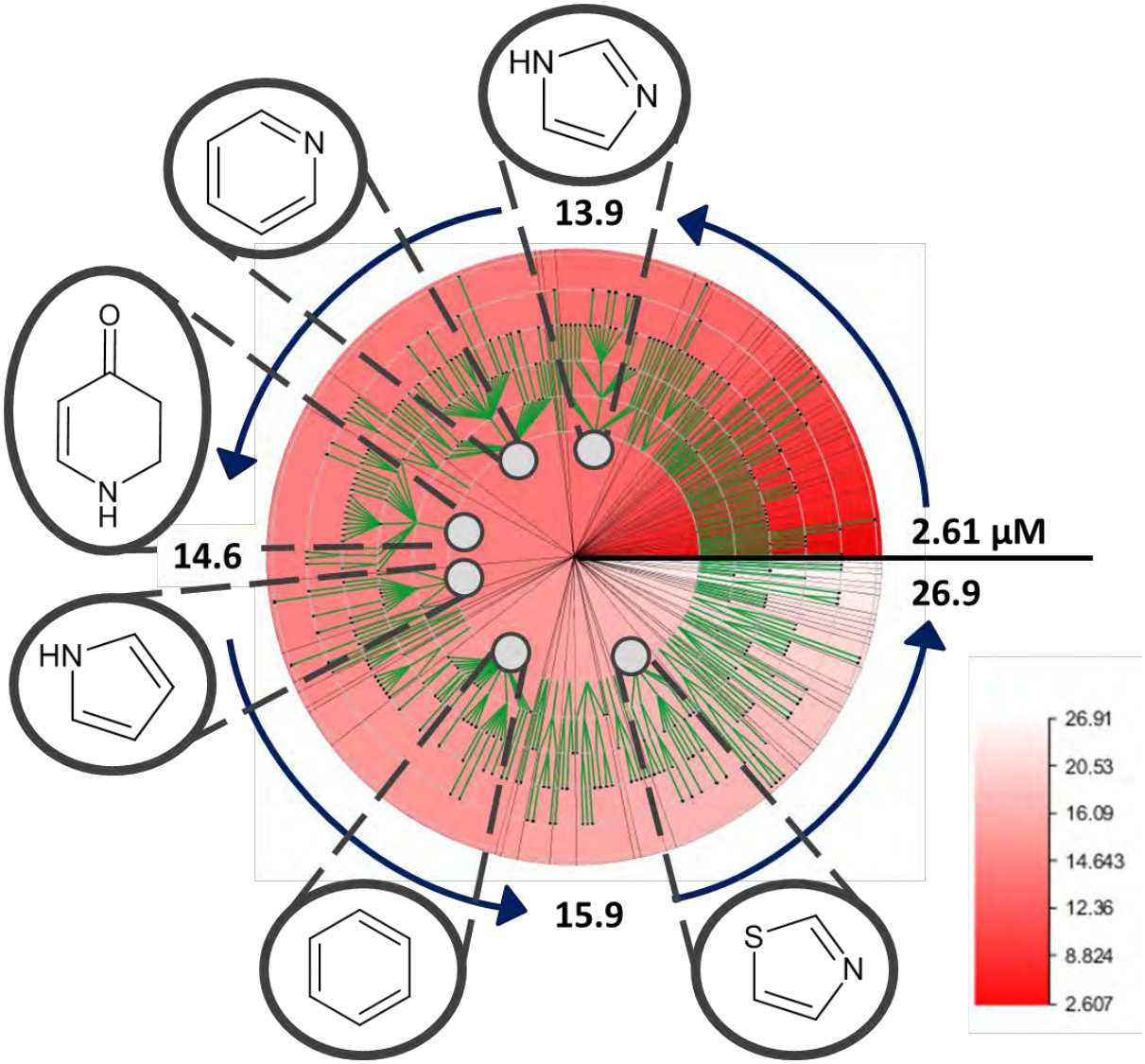


Figure 2.7 Diagram from Scaffold Hunter showing the diverse scaffold tree for βH hits arranged by increasing average IC_{50} value anticlockwise from 2.61 to 26.9 µM. The root of the tree starts as a single aromatic ring. The six scaffolds which led to the largest number of child scaffolds are highlighted.

The top ten most active parent scaffolds, which involved other N, O or S-containing five or six membered heteroaromatic rings, were linearly expanded to single structures on the scaffold tree. These structures are included in *Table 2.2*, which highlights the variety of chemotypes within the top 20 most potent βH inhibitors identified in the screening process. Interestingly, almost all the structures contained aromatic or heteroaromatic rings, with minimal saturated or freely rotatable bonds. Exceptions were **VU0020967** and **VU0013068,** with a cyclohexane ring and hexane moiety respectively.

Table 2.2 The 20 most potent βH inhibitors.

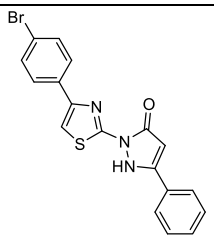
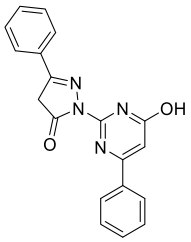
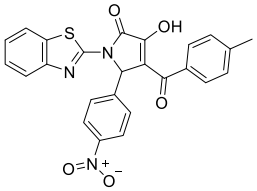
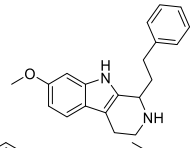
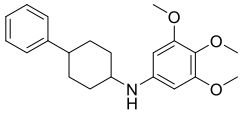
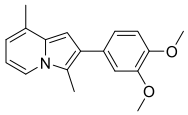
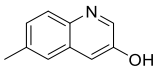
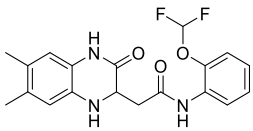
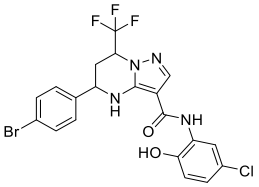
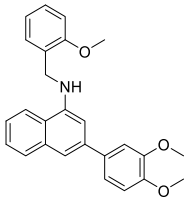
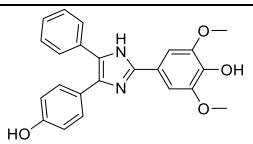
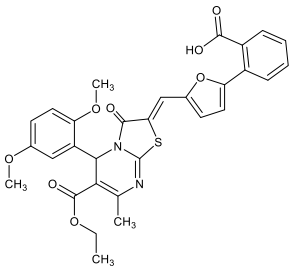
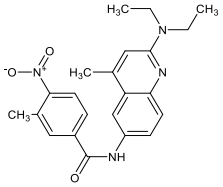
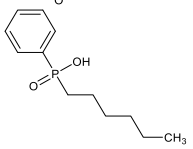
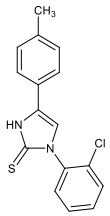
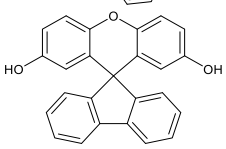
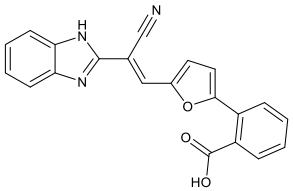
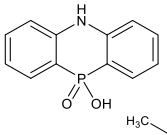
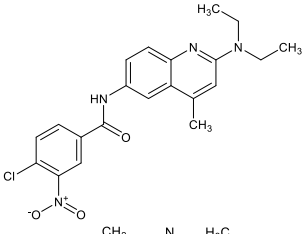
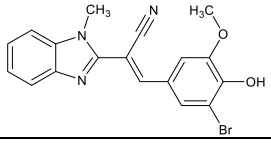
Code	Structure	logP	βH IC_{50} (μM)	D6 IC_{50} (μM)
VU0014379		4.52	0.168	Inactive
VU0014981		4.03	0.4	Inactive
VU0063871		5.19	0.4	Inactive
VU0015078		3.95	1.1	0.81
VU0020967		4.19	1.5	Inactive
VU0042031		3.65	1.7	3.515
VU0000264		2.34	2.0	0.59
VU0099210		3.54	2.2	Inactive
VU0123869		4.57	2.4	Inactive
VU0358176		5.46	2.4	1.72

Table 2.2 (Continued)

Code	Structure	logP	βH IC_{50} (μM)	D6 IC_{50} (μM)
VU0094619		5.33	2.4	0.70
VU0001216		4.71	2.607	Inactive
VU0073706		5.61	2.911	4.313
VU0013068		2.81	2.948	Inactive
VU0112751		4.82	3.049	Inactive
VU0000817		5.51	3.147	Inactive
VU0357956		3.98	3.151	Inactive
VU0009045		1.82	3.181	Inactive
VU0073726		5.70	3.215	Inactive
VU0358099		4.22	3.353	Inactive

2.2.3 Secondary screen of βH inhibitors for *in vitro* parasite activity

The SYBR Green I fluorescence-based assay was used to evaluate the efficacy of the βH inhibitors against *P. falciparum* (0.3 % parasitemia, 2% haematocrit).[84] This was achieved by first pre-screening the 530 compounds against the CQ-sensitive D6 strain, where hits were defined as those which inhibited ≥90% of parasite growth at a single concentration of 23 μM. The percentage inhibition was calculated relative to the positive control which contained the IC_{100} (kill concentration) of CQ and the negative control (DMSO only). Here, 171 compounds were identified as active against D6 parasite growth, giving a high hit rate of 32.3% out of the 530 βH hits. This is in contrast to the relatively low hit rate of 3% published in 2009 by Rush *et al.*[180] in their cell-free haem crystallization screen (CFHCS), which utilised an alternate method for βH formation. Their test compounds were incubated at 60 ˚C for 2 h in the presence of haemin and 9.7 M sodium acetate, with no detergent present. It appears that this lack of lipid or lipid mimic produced a larger number of βH hits. Of the 16,000 diverse compounds screened using their "acetate method", 644 (3.96%) were reported active with >50% inhibition. This hit rate is more than twice that of the hit rate (1.7% at >50% cut-off) found using the NP-40 based assay. Therefore the acetate method identifies a much larger percentage of βH inhibitors which either cannot access the DV or do not inhibit βH in the parasite's DV. It has been proposed that the latter is more likely owing to the less realistic mechanism of βH formation with an acetate mediator since high concentrations of acetate are known to solubilise free haem in a rate-limiting manner, which would not occur in the cell.[152,161,207] In the case of the NP-40 assay however, the favourable percentage of βH inhibitors that are also able to prevent parasite growth indicates the relative reliability of the NP-40 based assay and its ability to more accurately mimic the parasite's lipophilic environment through detergent micelles. The result also agrees with the pilot screen, which determined a hit rate of 30%.[161] The parasite active hits were analysed for any QSAR correlations between D6 IC_{50} and physical properties including MW, surface area and volume, logP, PSA, #HBD, #HBA, H_2O.Hyd and solubility. No statistically significant correlations were found when attempting linear regression with combinations of up to four variables. This result is not surprising given the diversity of scaffolds. Previously reported parasite activity correlations have only been identified within derivatives of a specific chemotype.[101,107,108,208,209] This is also the case for 3D QSAR.[210,211]

The relatively high single-point concentration of 23 µM chosen for the pre-screen (>1000 times that of the CQ IC_{50}) allowed for the maximum number of potentially interesting antiplasmodials to be detected. However, prioritisation of the scaffolds for further analysis towards antimalarial drug development was very dependent on the actual activities of the hits. To this end, the 171 compounds that inhibited parasite growth by ≥90% at 23 µM were selected and a dose response assay performed in duplicate. This resulted in 73 compounds showing activities below 5 µM of which 25 had nanomolar activities (< 1µM). These 25 were also tested against the multidrug resistant C235 strain of the parasite and their activities determined. In order to evaluate the potential for resistance predisposition, the resistance index (RI) of each compound was calculated (C235 IC_{50} / D6 IC_{50}). By comparing the activities of the resistant versus sensitive strains, the level of cross resistance can be assessed.[212] In this case only four compounds had RIs above 3 (16%), leaving 21 compounds which showed very low levels of cross resistance, including 15 compounds with RIs <2. This revealed that most of the novel chemotypes identified are roughly as effective against resistant strains as they are against sensitive strains, a property which gives them good potential for further optimization. *Table 2.3* displays the most potent 25 compounds exhibiting IC_{50}s ranging from 0.11 µM to 0.99 µM which show no apparent correlation with their βH activities. Of these eleven were derivatives of quinolines (including quinazolines and quinoliniums) and nine contained amide moieties.

Table 2.3 The 25 most potent D6 Parasite inhibitors.

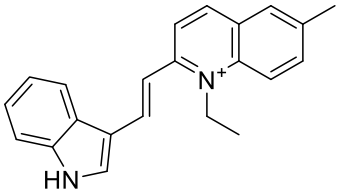
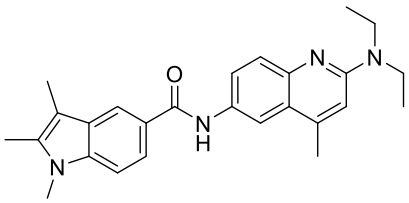
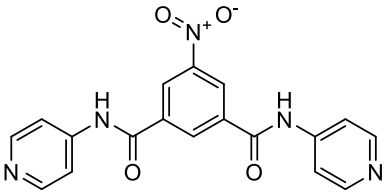
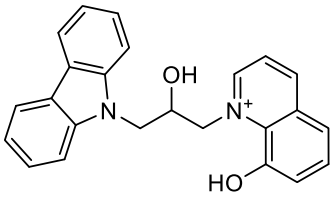
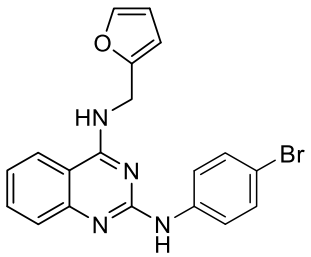
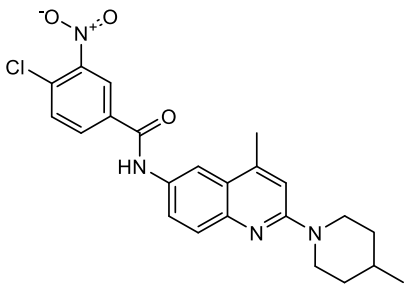
Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)	C235 IC_{50} (μM)	RI
VU0098755		12.6	0.11	0.13	1.2
VU0073687		6.3	0.18	0.55	2.9
VU0001281		5.9	0.19	0.17	0.9
VU0065708		16.2	0.20	0.18	0.9
VU0096505		8.8	0.24	0.22	0.9
VU0107282		17.0	0.29	0.54	1.9

Table 2.3 (Continued)

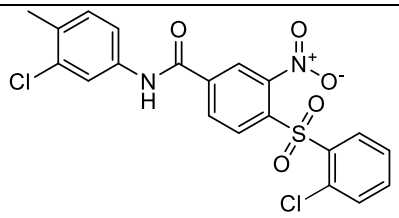
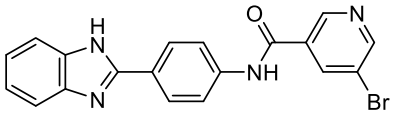
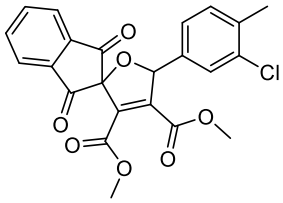
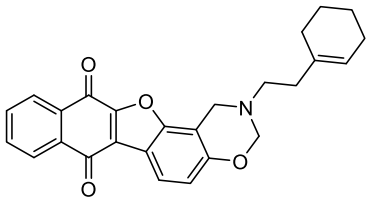
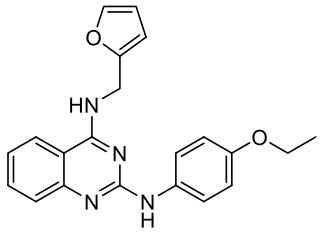
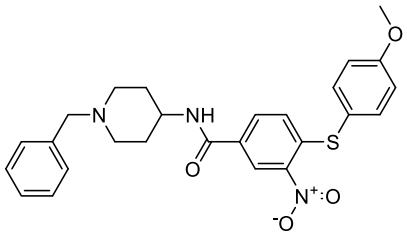
Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)	C235 IC_{50} (μM)	RI
VU0114785		13.4	0.35	4.82	13.9
VU0002101		14.3	0.35	0.41	1.2
VU0028177		13.3	0.35	0.46	1.3
VU0063971		8.9	0.38	0.83	2.2
VU0129813		26.78	0.396	0.294	0.7
VU0114734		22.08	0.461	1.138	2.5

Table 2.3 (Continued)

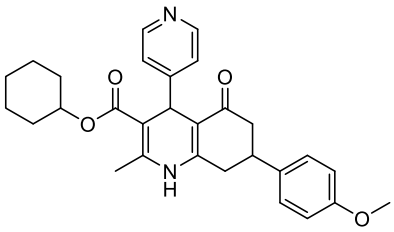
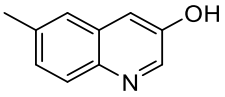
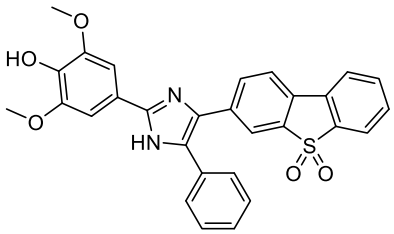
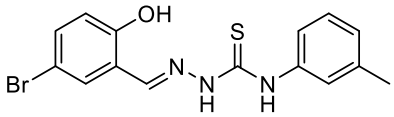
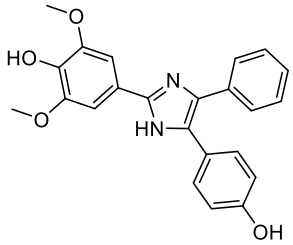
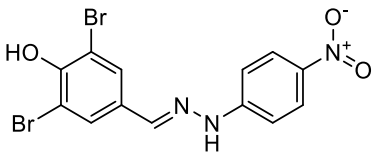
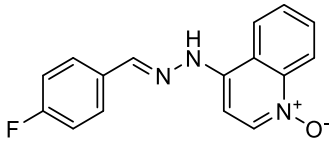
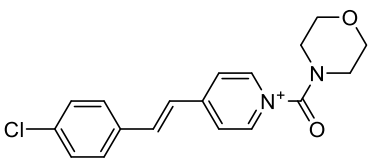
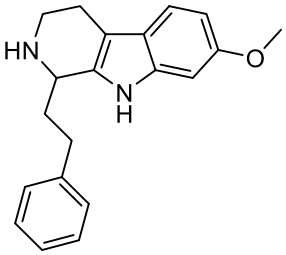
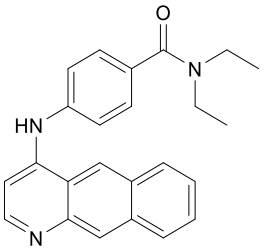
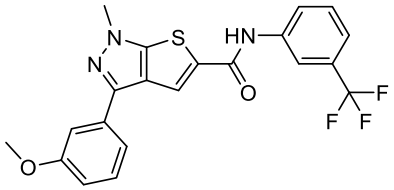
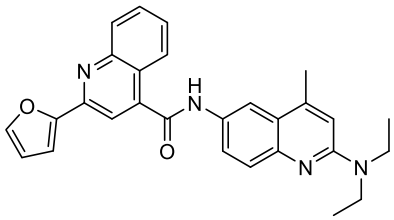
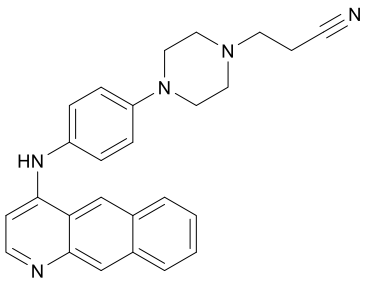
Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)	C235 IC_{50} (μM)	RI
VU0122653		24.54	0.5503	0.8998	1.6
VU0000264		2.043	0.59	0.248	0.4
VU0065892		6.194	0.611	0.693	1.1
VU0010690		25.03	0.673	1.52	2.3
VU0094619		2.394	0.70	0.2842	0.4
VU0012464		7.288	0.708	0.923	1.3
VU0008057		9.567	0.73	1.497	2.1
VU0358149		8.339	0.7782	19.1	24.5

Table 2.3 (Continued)

Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)	C235 IC_{50} (μM)	RI
VU0015078		1.121	0.81	5.744	7.1
VU0119324		21.47	0.8453	0.1864	0.2
VU0078364		5.299	0.922	1.971	2.1
VU0107278		11.77	0.951	3.896	4.1
VU0064165		15.36	0.9926	1.008	1.0
	CQ	53.0	0.016	0.226	14.1
	AQ	21.0	0.011	ND	ND

2.2.4 Scaffold analysis

Discovering novel βH inhibiting chemotypes, particularly ones which maintain activity *in vitro*, was the primary motive for this HTS. Following the collection of all the activity data, the 530 βH hits were further analysed to assess the scaffolds in more detail, specifically in relation to the percentage of each class which are also biologically active. This process revealed that 246 of the hits belonged to a class containing at least seven derivatives (*Figure 2.8*). These could further be traced back to the simpler structures, quinoline or quinazoline, benzamide and benzazole with the exception of scaffolds M and N. Interestingly, the parent scaffold a level up from these simpler structures was identified as an aminophenyl as shown in the scaffold tree diagram in *Figure 2.8*. The compounds from each class that were also active against the D6 strain of *P. falciparum* were considered in order to determine which chemotype holds the most promise for further hit to lead optimisation. Here, the triarylimidazoles (scaffold N) performed the best, with six of seven compounds (86%) maintaining parasite activity while the diarylpyrrolones (scaffold M) exhibited no compounds active *in vitro*. Approximately 50% of the quinoline and benzamide-based hits were also parasite actives, with the exception of scaffold C with a significantly lower value of 22%. However, these percentages must be treated with caution since the results are biased by the initial compounds available in the HTS library. The remaining 54% of the βH hits were not assigned a specific scaffold class, an affirmation of the diversity of the identified active compounds.

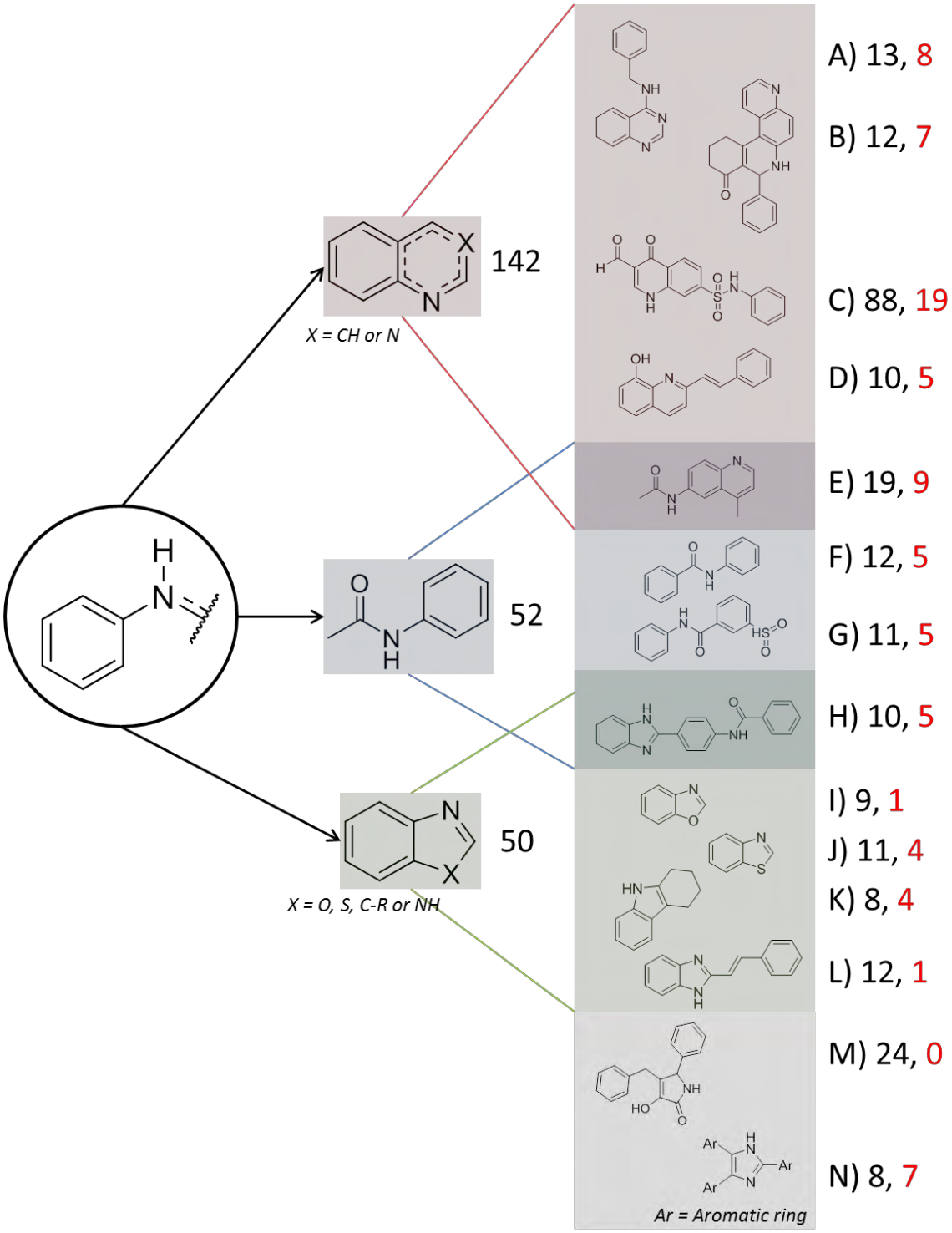


Figure 2.8 Fourteen βH hit scaffolds (A to N containing at least seven derivatives each), twelve of which were linked back to simple parent scaffolds. Scaffolds E and H are each derived from two parent scaffolds. The values indicate the number of βH hits (≥80% inhibition) belonging to the scaffold followed by the number of compounds (red) which are also active against the D6 *P. falciparum*.

2.2.4.1 Haemozoin target validation via cell fractionation

In addition to analysing the percentage of compounds within each scaffold that prevent parasite growth, it was also important to validate the therapeutic target within *P. falciparum*. This is due to the fact that a compound need not necessarily inhibit haemozoin within the whole-cell, even though it is active in the NP-40 assay. One reason for this may be that the compound is unable to access or accumulate in the DV to inhibit haemozoin, while still being to act on an alternate parasite target, allowing for favourable *in vitro* activity. Alternatively, the compound may be able to access the DV but interacts with haem and the other cellular components differently *in vitro* to how it interacts with buffered haematin in the NP-40 mediated assay. If this were found to be the case, the reliability of the NP-40 assay would be called into question. Fortunately this was not the case, as cell fractionation studies carried out at Vanderbilt University in collaboration with the University of Cape Town's (UCT) pharmacology department revealed all twelve compounds tested were haemozoin inhibitors, in contrast to a negative control.[204] The cell fractionation procedure was based on previously published methods developed at the UCT, where the negative control employed was the antifolate, sulfadoxine/ pyrimethamine.[113] However, in this study all compounds were analysed relative to a different negative control, α-[(cyclohexylamino) methyl]-9H-Carbazole-9-ethanol, which was inactive in the NP-40 assay but active in the parasite (D10 IC_{50} of 0.99 μM).[204] The twelve compounds were selected in order to represent the majority of the different scaffolds, including **VU0001281** (*Table 2.4*), a pyridyl analogue of the phenyl benzamide scaffold (F). The activity values for each compound found in the D6 parasite assay were used to give the required testing concentrations (0, 0.5, 1, 2 and 3 times the IC_{50} value). Haemozoin inhibition was measured via the quantification of D6 *P. falciparum* Hb, free haem and haemozoin remaining after incubation with the test compound. An active haemozoin inhibitor was one which resulted in a decrease in haemozoin with an increase in free haem. In addition, some compounds were shown to cause significantly larger levels of free haem at their respective IC_{50}s. The negative controls did not display any increase in free haem relative to baseline levels, despite being able to cause parasite death, indicating an alternate mechanism of drug action.

The cell fractionation study by Combrinck *et al.*[113] indicated that CQ caused approximately a 10% increase in free haem (5% at 0 nM to 15% at 2.5 times the IC_{50}), as discussed in

Section 1.7.3. Tested compounds from the HTS based on the quinoline scaffolds (B, C, D and E) showed a relatively low increase in free haem (8 to 14%) despite having moderate-high D6 activity. This was especially evident for the compound representing scaffold E which only produced a 14% increase in free haem while also having a very low IC_{50} value. In contrast, compounds from scaffolds G and I induce high levels of free haem but had relatively low activity. Scaffolds F (phenyl benzamide), F* (pyridyl benzamide) and N (triarylimidazole) contained compounds which caused the largest increase in free haem at their respective IC_{50}s, despite that fact that they exhibited a wide range of activities from 0.19 – 9.06 µM. These compounds were particularly interesting from a mechanistic perspective, since haemozoin inhibition is almost certainly the primary target. All the compounds induced a significant increase in free haem relative to that of 0.4% for the negative control (*Figure 2.9*).

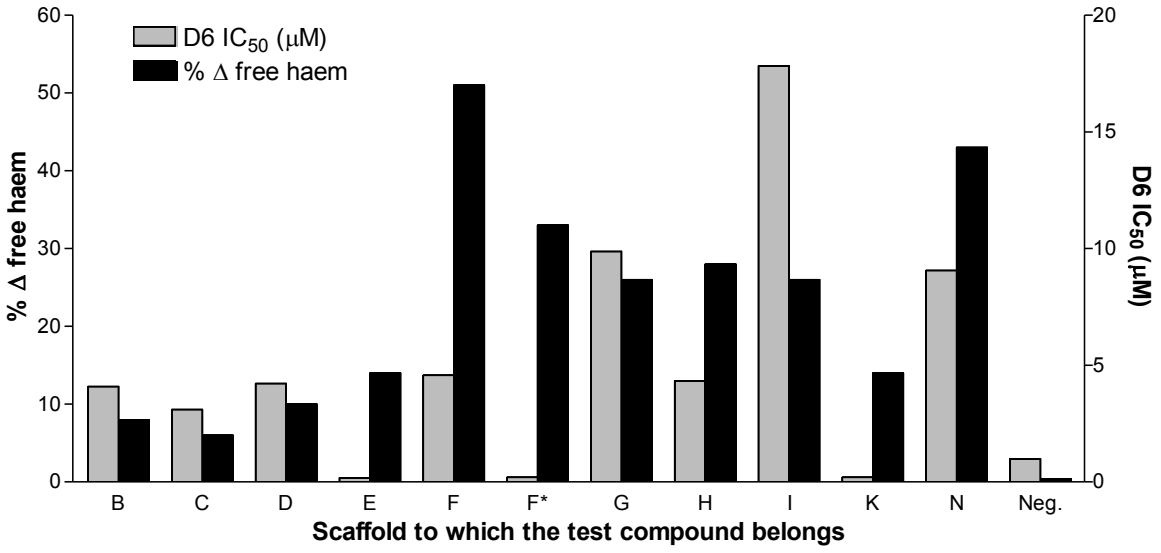


Figure 2.9 The increase in free haem at the IC_{50} found for each compound belonging to the specified scaffold. F* represents the pyridyl derivative of scaffold F (phenyl benzamide). Scaffolds F, F* and N showed the largest increase in free haem at their IC_{50}s.

2.2.4.2 Phenyl and pyridyl benzamides (Scaffold 1)

The phenylbenzamides have been reported for showing activity against the 3D7 strain of *P. falciparum* using the parasite lactate dehydrogenase as a measure of parasite growth. These

data were as a result of the GSK *P. falciparum* screen published in 2012.[178] The Novartis group went on to optimise the solubility and selectivity of a benzamide which led to enhanced *P. falciparum* activity. (*Figure 2.10*).[213] Another study reported antimalarial activity for a series of benzamide-containing stilbene derivatives.[214] However, it is not known whether these compounds are haemozoin inhibitors.

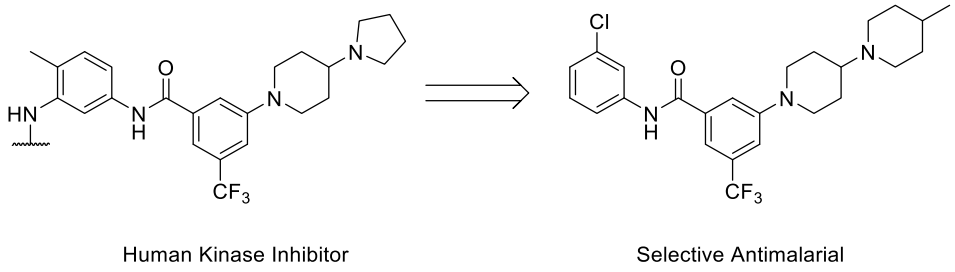


Figure 2.10 The benzamide compound discovered via the GSK antimalarial screen and optimised by Wu *et al.*[213] from the Novartis group.

A list of the βH inhibiting benzamides were analysed by mining the 530 hits for the subset which contained scaffold F and F*. This was carried out using the Molecular Substructure Miner (MoSS) software, freely available online at http://www.borgelt.net/moss.html.[215] The twelve benzamides which were active against the D6 strain of the parasite and do not contain a quinoline ring are displayed in *Table 2.4*. Six of the benzamides also contained a benzimidazole and one contained a benzoxazole moiety. Compound **VU0001281** was particularly interesting given that it showed the best *in vitro* activity (and the third most potent out of the whole screen) and because of its large increase in free haem (about three times greater than the quinoline compounds). However, there were some structural alerts for this compound in terms of its potential drug profile. Firstly, the aromatic nitro group has been known to result in toxic metabolites via reduction to the carinogenic nitroso toxicophore. Sequential reduction forms further toxicophores, aromatic hydroxylamines (ArNHOH) and anilines (ArNH2).[216,217] Compounds containing nitro groups can also lead to methaemoglobinemia, a blood disorder characterised by too much ferric [Fe3+] versus ferrous [Fe2+] Hb.[218,219] Secondly, the 4-pyridyl moiety in **VU0001281** is a potential cytochrome P450 (CYP) enzyme inhibitor which may result in oxidation to the pyridine-oxide by one or more CYP enzymes.[216]

Table 2.4 Non-quinoline containing benzamides from the HTS (scaffold F and F*) which showed ≥90% parasite activity at 23 µM. Compounds are listed in order of decreasing D6 activity and % free haem increases are shown in red where data are available.

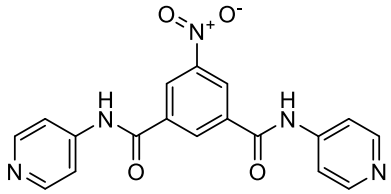
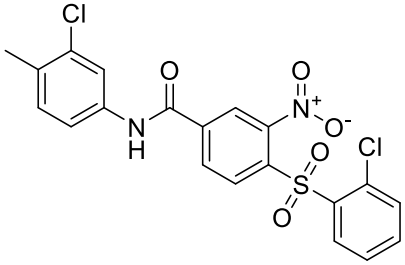
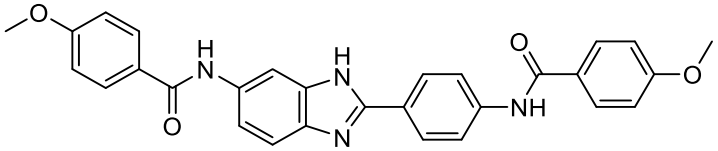
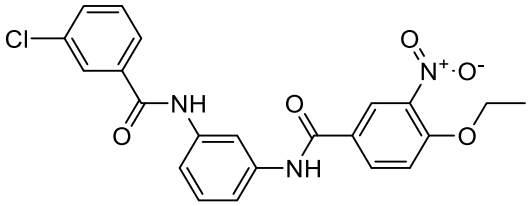
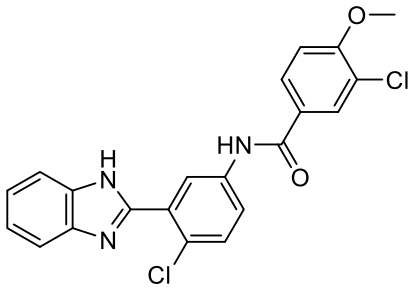
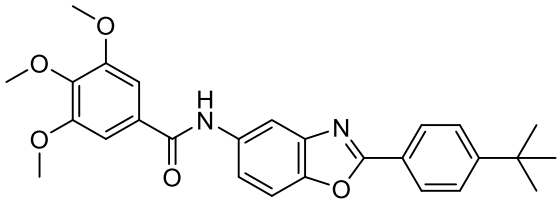
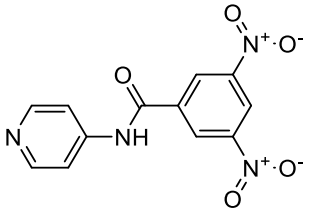
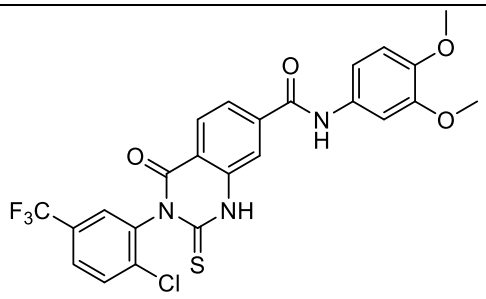
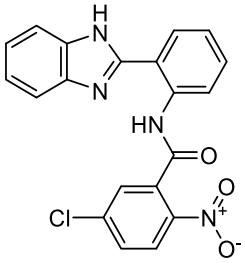
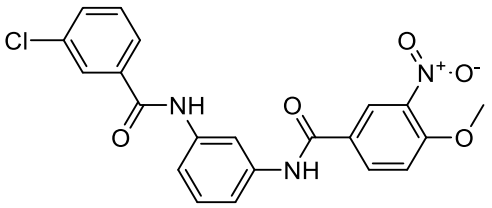
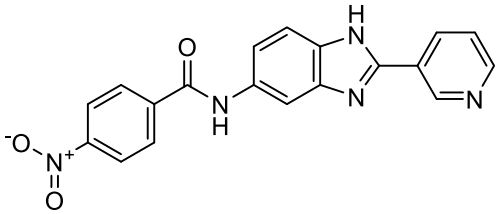
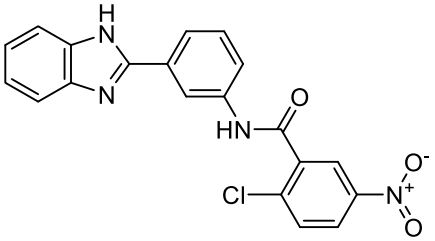
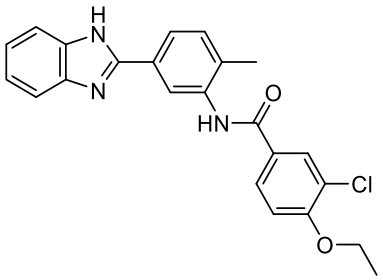
Code	Structure	βH IC_{50} (µM)	D6 IC_{50} (µM)
VU0001281		5.86	0.194 (33% Δ free haem)
VU0114785		13.43	0.3462
VU0117011		10.42	1.686
VU0054768		22.57	3.471
VU0122425		21.51	4.317
VU0054902		17.94	4.565 (51% Δ free haem)
VU0358764		9.038	4.758

Table 2.4 (Continued)

Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)
VU0105653		20.02	5.308
VU0122447		17.26	7.856
VU0054351		17.04	9.096
VU0093551		10.95	11.08
VU0060787		17.09	14.17
VU0056416		21.2	16.13

The potential toxicity linked to compound **VU0001281** prompted some *in silico* modelling using *MetaSite*[220,221] and *VolSurf+* software.[222] Structural alerts for potential metabolic instability were confirmed by *MetaSite*, which identified the reduction of the nitro group, N-oxidation of the pyridyl moiety and amide hydroxylation as sites of possible metabolism in the liver. However, *VolSurf+* predicted the benzamide derivatives from the HTS to have suitable metabolic stability (within the 95% confidence boundary) when mapped relative to the chemical space of known drugs (*Figure 2.11*). Furthermore, very successful drugs such as Lipitor, Lidoderm and paracetamol contain benzamide moieties and it was hypothesised that replacing the nitro and pyridyl moieties with other functionalities could overcome potential metabolic weaknesses. Other models within *VolSurf+* predicted appropriate CACO2 permeation, protein binding, solubility and volume of distribution within 95% confidence boundary of known drug space.

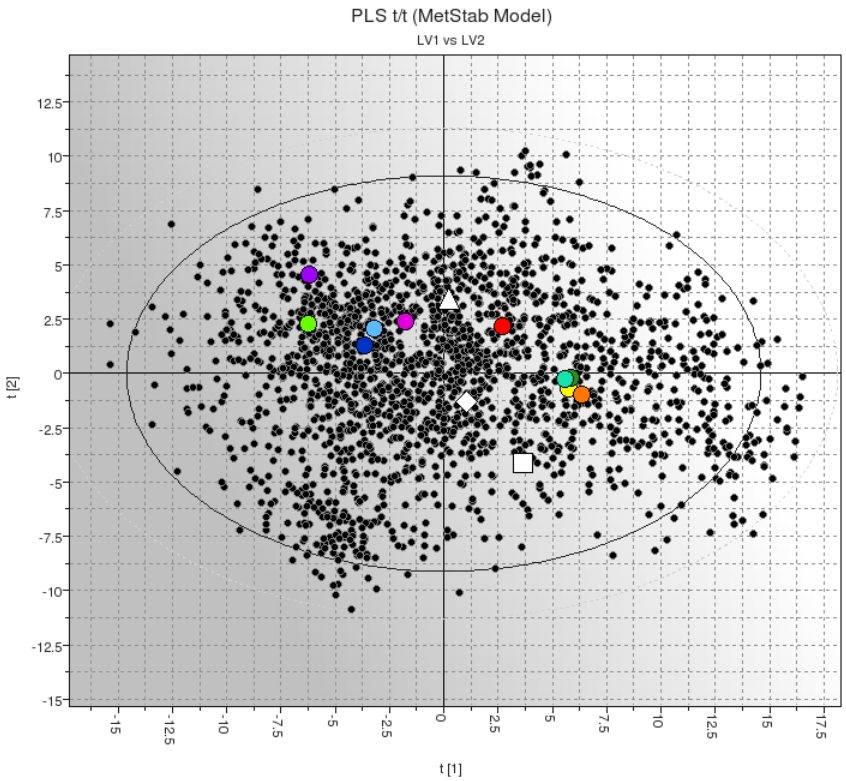
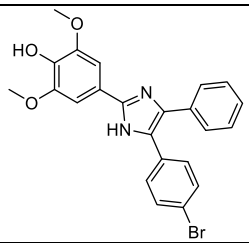
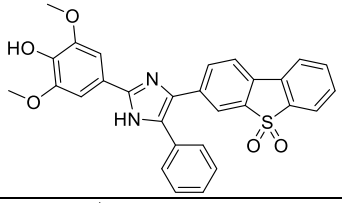
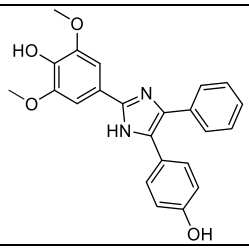
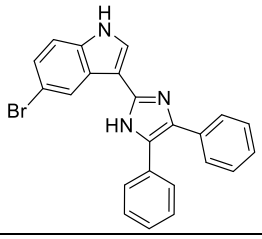
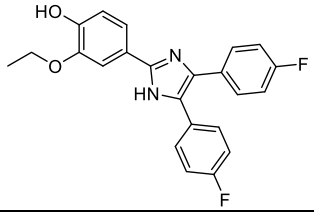
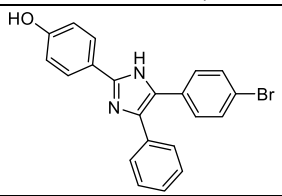
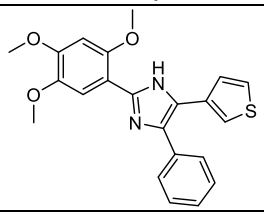


Figure 2.11 Metabolic stability space of known drugs forming the training set of model compounds (black circles), a range of pyridyl benzamides (coloured circles) and known antimalarials, CQ (white diamond), AQ (white triangle) and QD (white square). All the benzamides and known drugs were within the outer circle, representing a 95% confidence for the metabolic stability space.

2.2.4.3 Triarylimidazoles (Scaffold 2)

Scaffold N (triarylimidazoles) displayed a very high rate for parasite activity (six of seven βH inhibitors were active against *in vitro*). In addition to this, compound **VU0099289** induced a large increase in free haem in the cell fractionation assay of 43% relative to the negative control. The identified compounds showed activities from 0.61 to 11.71 μM, with small structural changes having significant influence on the activity as evident when comparing **VU0132562** (0.3 μM) **VU0094619** (0.7 μM), **VU0129734** (3.4 μM) and **VU0099289** (9.0 μM). Imidazoles have a wide application owing to their range of biological activities, such as anti-inflammatory,[223] anti-allergic,[224] anti-analgesic,[225] anti-cancer[226] and antibiotic.[227] Additionally, a series of 2,4,5-triarylimidazoles was shown to have antibacterial properties.[228] Rathore *et al*.[229] screened 110,000 drug-like, diverse heterocyclic compounds and identified 305 inhibitors of *P. falciparum* haem detoxification protein (HDP), one of them being the triarylimidazole, 4,4'-(4-phenyl-1H-imidazole-2,5-diyl)bis-phenol. However no follow-up reports of antimalarial activity could be found in the literature. Recently, several triarylimidazoles were discovered to have 3D7 parasite activity as a result of the phenotypic screens carried out by GSK and Novartis.[171,181] Due to the scope of imidazole-based compounds as biologically active agents, this scaffold was identified as a potentially interesting starting point for a series of derivatives to probe haemozoin inhibition. Moreover, an investigation into βH inhibition SARs for this class was attractive owing to the novelty of the chemotype as an antimalarial. The active triarylimidazoles are shown in *Table 2.5*.

Table 2.5 The seven βH inhibiting triarylimidazoles which were active against P. falciparum. Compounds are listed in order of decreasing D6 activity and % free haem increases are shown in red where data is available.

Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)
VU0132562		12.6	0.310
VU0065892		6.194	0.611
VU0094619		2.394	0.700
VU0123773		12.72	2.245
VU0129734		18.68	2.396
VU0099289		5.652	9.061 (43% Δ free haem)
VU0099279		10.02	11.71

2.3 Summary and conclusions

A total of 43,520 compounds were screened as a significant contribution to a larger HTS effort, involving a total of 144,330 compounds from the VU chemical compound library. Of the total screened 530 (0.37%) were identified and subsequently confirmed as βH inhibitors with potent IC_{50} values all below that of the CQ control. The structurally diverse hits generally possessed physicochemical properties, such as large #AR, #HBD and #HBA and low #RB, classifying them as planar, aromatic compounds capable of hydrogen bonding and π-π interactions with Fe(III)PPIX. Upon testing the hits from the primary screen against *P. falciparum*, 171 inhibited parasite growth by ≥90% at 23 μM, (32.3% of the 530 βH inhibitors) and 25 compounds exhibited favourable IC_{50} values below 1 μM. Furthermore, cell fractionation studies were able to successfully validate haemozoin formation as the mechanism of action in parasites for all tested compounds, representing twelve of the fourteen novel scaffolds identified.

In this study HTS was successfully applied to the discovery of numerous yet diverse chemotypes, active against malaria parasites, the majority of which have not been reported as having this activity. Moreover, all the hit scaffolds were identified for the first time as haemozoin inhibitors, including those also found in previous phenotypic screens by St. Judes, GSK and Novartis, where the drug targets are unknown. Identifying the mechanism of action for newly validated hits from phenotypic screens remains a challenge. Historically, very low hit rates have been attained when attempting to select inhibitors of specific targets from various other sets of phenotypic hits. However, Pink *et al.*[230] suggest that this is probably due to insufficient validation of the targets rather than to the methods themselves. Not only has this target-based screen allowed for 144,330 compounds to be evaluated for their βH inhibiting antimalarial activity in a more efficient, robust and cost-effective manner than whole-cell screening, but a diverse range of compounds were proven to be haemozoin inhibitors *in vitro*. This, along with the unprecedented activity retention between the two utilized assays, shows the value and reliability of the initial NP-40 assay for accurately mimicking the haemozoin formation process within the parasite DV. While other large-scale screening efforts, such as the GSK screen have been very successful in finding thousands of antimalarial molecules, including one already in clinical trials, the failure rate for new antimalarials in phase 2 trials is 50%.[230] Therefore it is crucial to utilise the potential

of the existing antimalarial hits for maximum benefit in order to probe mechanistic aspects of activity.[178] This can be achieved most successfully by having access to multiple hits which act on the same target in order to understand SARs. Improvements in public disclosure and sharing of data will also aid in bridging the gaps in the knowledge of potential starting points for antimalarials. Two hit scaffolds from this HTS, the benzamides and the triarylimidazoles, were particularly interesting since upon inoculation with representative compounds, much larger quantities of free haem were measured in the parasite, compared with the known quinoline antimalarials. Furthermore, their rates of βH inhibition hit occurrence were relatively high and parasite activities varied significantly between derivatives within each scaffold. These novel chemotypes were chosen for further investigation, requiring the purchase or synthesis of analogues to establish the minimum active pharmacophores and SARs in each case.

CHAPTER THREE

SYNTHESIS

3.1 Introduction

In order to determine the potential of the benzamide and triarylimidazole scaffolds discovered via HTS (discussed in *Section 2.2.4*) for further hit-to-lead development, a more detailed understanding of the SARs of each pharmacophore was required. The promising initial parasite activities for at least one compound from each chemotype led to the prospect of improved analogues that could be efficiently synthesised from cheap starting materials. Investigating novel βH and antimalarial scaffolds requires a library of derivatives with different physical properties and functional groups. While these derivatives are not necessarily intended to become antimalarial drugs, they are required as probes for understanding the mechanism of action. Choosing chemotypes that do not require demanding, expensive and multi-step syntheses allows for more derivatives to be synthesised in a given time frame, and is a rational way of effectively probing haem-drug interactions and finding SARs. This is especially important for quantitative structure-activity analysis, since having more synthesised compounds for activity input data may lead to an improved statistical significance and more reliable relationships.

3.1.1 Scaffold 1: Benzamides

3.1.1.1 Amide bond formation

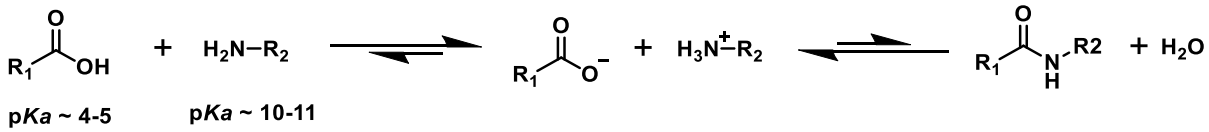
Benzamides have been synthesised and extensively studied owing to their wide range of applications, which include pharmaceuticals and medicinal chemistry,[231,232] polymer materials[233] and supramolecular chemistry.[234] As a result, many methods have been developed for their synthesis, most of which most focus on amide-bond formation. The amide functionality is present in proteins, which are involved in almost all biological processes and therefore ubiquitous in life.[235] Moreover, the carboxamide functionality appears in 25% of all known drugs analysed from the Comprehensive Medicinal Chemistry

database.[236] This may be because they are neutral, stable and also have hydrogen-bond donating and accepting properties. The most prevalent method of synthesis involves condensation between an acid or its activated derivative and an amine, in which a whole host of methods are available.[235,237]

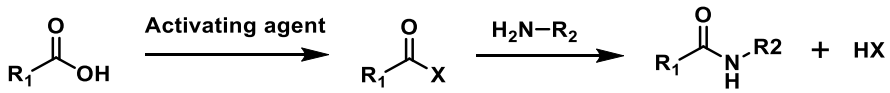
Other examples of amide bond synthesis include the hydrolysis of aromatic nitriles,[238] aminocarbonylation between an aromatic halide, an amine and carbon monoxide with a palladium catalyst[239,240] and the acid-induced rearrangement of oximes (Beckmann rearrangement).[241-243]

Formally, amide bond formation between an acid and an amine is a condensation reaction, thereby liberating H_2O. However, for this to take place, adverse thermodynamic factors need to be overcome, owing to the stable salt which forms upon mixing an acid with an amine (*Scheme 3.1a*). As a result, direct condensation of the salt can only take place at high temperatures exceeding 160 °C, which are generally incompatible with other functionalities.[244] In order to overcome this, the traditional and easiest method, involving conversion of the hydroxyl into a good leaving group, can be employed. Here, the acid is activated after introducing the leaving group to the acyl carbon of the acid, allowing for amine attack by virtue of a lower activation energy and hence under milder conditions (*Scheme 3.1b*). A variety of coupling reagents exist for this purpose, enabling the generation of acid chlorides, mixed anhydrides or active esters.[245]

(a)



(b)

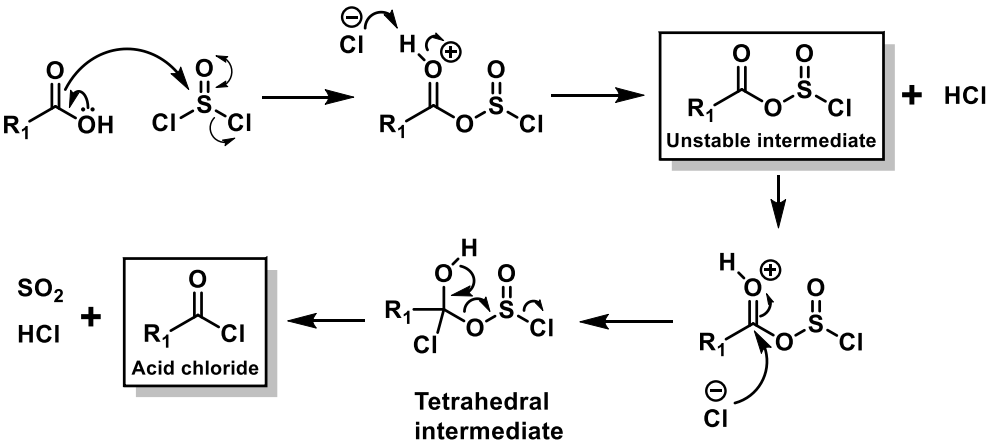


Scheme 3.1. Aspects of amide bond synthesis.

Owing to the plethora of methods and reagents for amide bond formation, it is sometimes necessary to screen a variety of different conditions for reaction optimisation. Several factors often need to be considered, which may include yield and selectivity improvement, or reduction of by-product formation. In other cases, the aim may be to facilitate the purification process or take advantage of more economical reagents.[235] Amide bond formation during peptide synthesis can be more complicated as a result of possible epimerisation, usually through an oxazoline intermediate. In this circumstance, peptide coupling reagents such as dicyclohexylcarbodiimide (DCC) are generally used, in which an activated intermediate O-acylurea is formed *in situ*.[245] Other popular peptide coupling reagents include 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium-3-oxidhexafluoro phosphate (HATU).

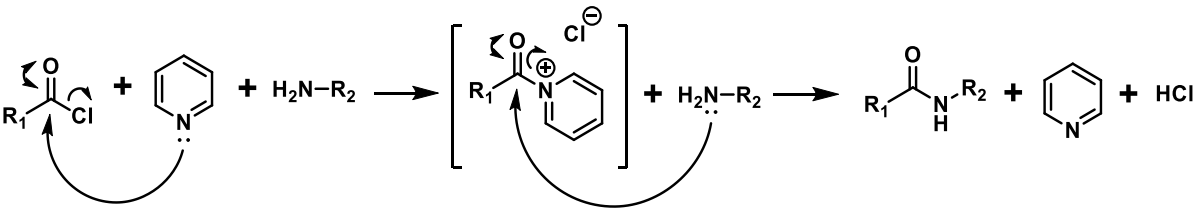
For amide bond formation which does not involving peptides and where there is no danger of hydrolysis, racemisation or cleavage of protecting groups; acyl chlorides or acid chlorides are one of the most popular and easiest to prepare of the activated intermediates. They can be readily made from the acid using a host of reagents such as thionyl chloride ($SOCl_2$),[246,247] oxalyl chloride ($COCl_2$),[248,249] phosphorus trichloride (PCl_3),[250] phosphorus oxychloride ($POCl_3$),[251] or phosphorus pentachloride (PCl_5)[250]. The sulphur, carbon or phosphorus in these reagents is necessary for removing the oxygen from the acid, and therefore acid chlorides cannot be formed from HCl alone.[252] Oxalyl chloride is potentially hazardous due to the stoichiometric production of three gases, HCl, carbon dioxide and carbon monoxide and therefore care needs to be taken when considering this reaction. Generally, the phosphorus-containing reagents are employed in cases where the acid does not react readily with $SOCl_2$, which produces sulphur dioxide and HCl as by-products.[235] The attack of the acid oxygen on the electrophilic $SOCl_2$ sulphur results in an unstable, highly electrophilic chlorosulfite ester intermediate, which is then protonated by expelled HCl. This increases the electrophilicity of the carbonyl carbon resulting in addition of the choride ion to form a tetrahedral intermediate, which collapses irreversibly to the acyl chloride with release of sulphur dioxide (SO_2) and HCl as shown in *Scheme 3.2.*[252] The production of HCl may, in some cases be problematic, especially in acid sensitive reactions such as those involving Boc-protected amines. In this instance, cyanuric chloride (2,4,6-trichloro-1,3,5-triazine),

which maintains basic pH conditions, can be used to form the acid chloride in the presence of triethylamine (Et_3N).[253]



Scheme 3.2. Mechanism for the formation of the acid chloride using thionyl chloride.

Once the acid chloride has been formed, a simple coupling reaction with the appropriate amine is carried out to form the amide. Once again, this reaction liberates HCl, and a base is necessary to avoid protonation of the amine to form an unreactive ammonium chloride salt. Furthermore, other basic groups on the acid or amine substrates may become protonated under these conditions. For this reason, reactions are generally carried out in the presence of non-nucleophilic tertiary amines, such as Et_3N or Hünig's base (*i*-Pr_2NEt), in inert dry solvents.[254] In order to decrease reaction times and improve solubility in some cases, pyridine can be used as a solvent. Pyridine acts as an acyl-transfer catalyst in the aminolysis by undergoing a nucleophilic acyl substitution (S_NAc) reaction with the acid chloride faster than with the primary amine to generate an intermediate acyl-pyridinium salt.[235] This is due to the pyridine nitrogen's relatively soft and strongly Lewis basic character. The ionic species formed is more electrophilic than the acid chloride, and attack by the amine is driven by the positive charge on the pyridine nitrogen (*Scheme 3.3*). *N*,*N*-dimethylaminopyridine (DMAP) can also be used for this purpose.[255]

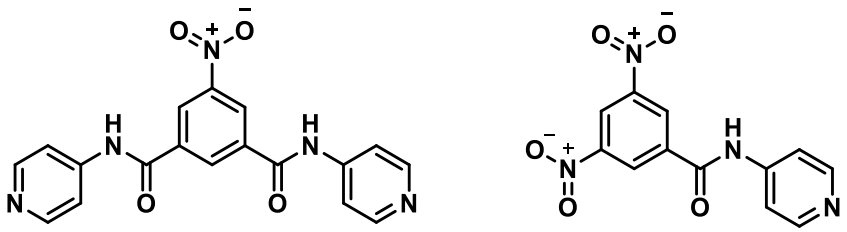


Scheme 3.3. Mechanism of acyl-transfer with pyridine in the aminolysis.

3.1.1.2 Synthetic targets for the benzamide scaffold

The benzamide that showed the highest activity in the HTS against the D6 strain of the *P. falciparum* parasite was a di-pyridylbenzamide (**VU0001281**) with an IC_{50} of 0.194 µM (*Figure 3.1a*). Cell fractionation studies showed that this βH inhibitor also induced a large increase in free haem in the parasite, supporting haemozoin as the therapeutic target (*Table 2.4*). Related to the dibenzamide compound was the mono-pyridylbenzamide (**VU0358764**) containing a nitro group in place of the second *N*-pyridyl amide (*Figure 3.1b*). While this compound showed similarly potent activity against βH formation, it was significantly less active towards the parasite in the D6 strain relative to **VU0001281**, with an IC_{50} of 4.76 µM. These results prompted a study into the structure-activity relationships (SARs) of this scaffold.

a) b)



VU0001281
IC_{50} = 0.194 µM

VU0358764
IC_{50} = 4.76 µM

Figure 3.1 The mono- and di-pyridylbenzamide identified in the HTS.

In order to generate a library of mono- and di-benzamide derivatives, several structural aspects were chosen for variation in order to probe the effect of varying the electronic and spatial properties of the substituents on both βH and parasite activity, and to address improving the toxicity profile of the compound. *Figure 3.2* shows the parts of the mono- or di-benzamide scaffolds which were varied either by purchase (indicated by the *) or synthesis in order to probe SARs. Several mono-benzamides were purchased according to information that emerged from the Topliss decision tree, a method designed to aid in the selection of new analogues, based on maximising the chances of improving activity. It also allowed for the maximum discrimination between substituent properties such as resonance vs inductive effects as well as substituent size and hydrophilicity.[256] In the case of the di-

benzamides, all derivatives were symmetrical except for the benzimidazole ring (**), which was attached to the one amide nitrogen, with phenyl on the other.

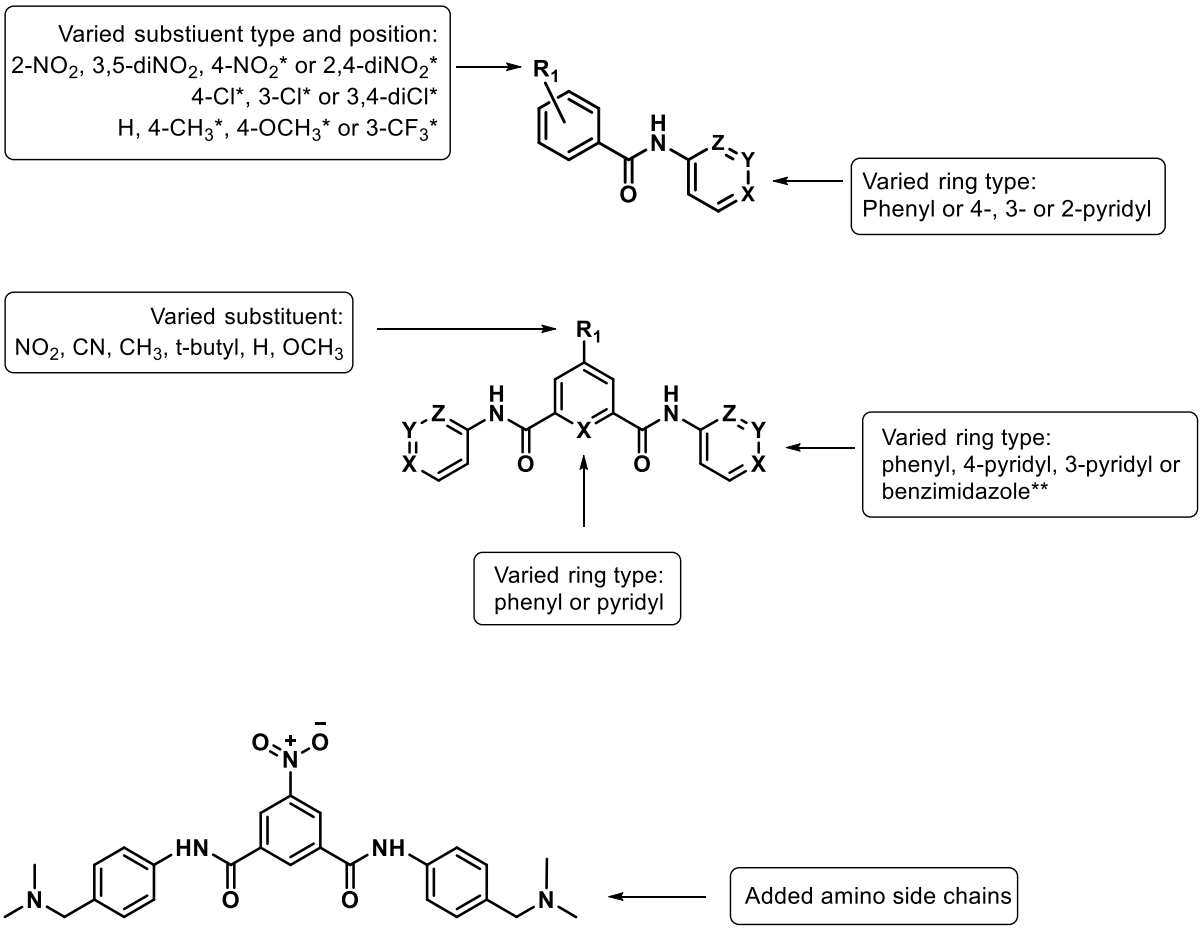
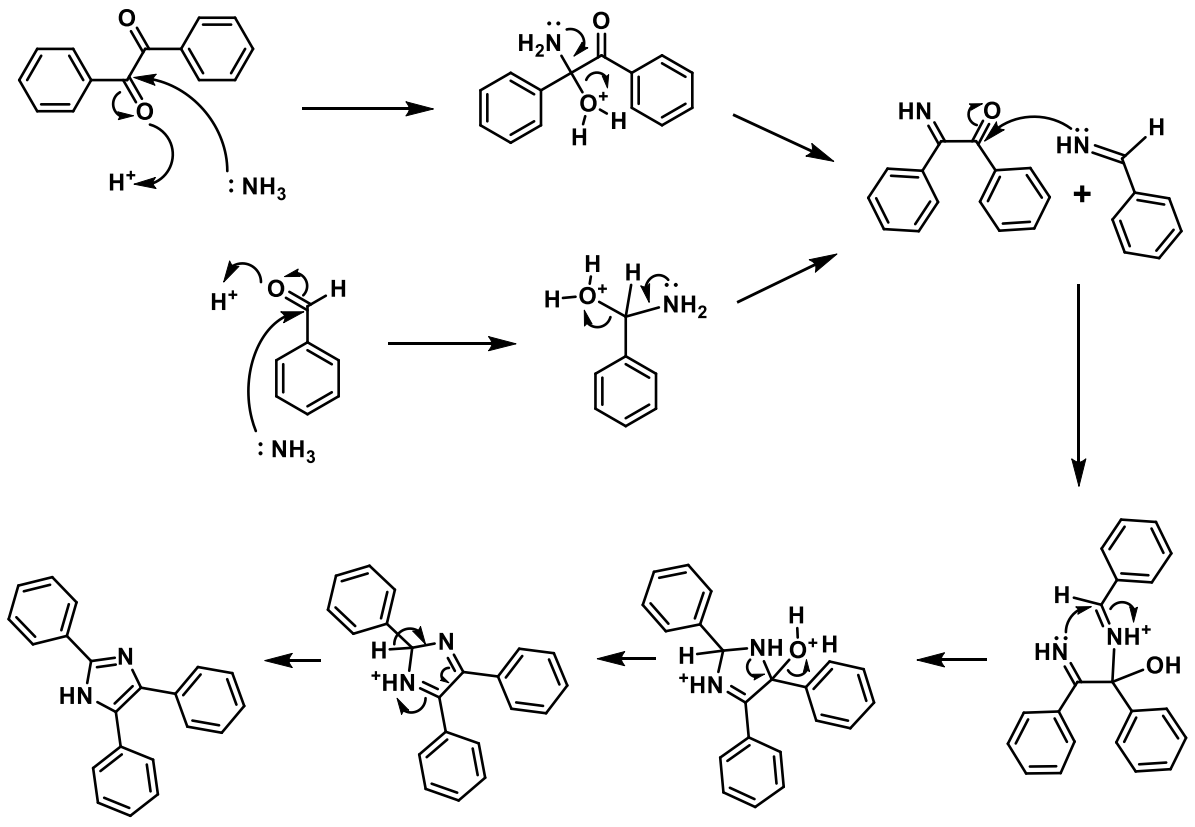


Figure 3.2 Derivatives of the mono- and di-benzamides obtained from commercial sources or via synthesis.

3.1.2 Scaffold 2: Triarylimidazoles

3.1.2.1 Literature methods for the synthesis of triarylimidazoles

Many synthetic routes to 2,4,5-triarylimidazoles have been reported in the literature, most often in conjunction with probing their biological activities,[228,257] fungicidal or herbicidal properties[258,259] or their potential usage in carbon-based materials.[260] The standard procedure for accessing a 2,4,5-triphenylimidazole core involves the condensation reaction of benzaldehyde, benzil, and ammonium acetate. This can be performed in a variety of organic solvents,[259,261] glacial acetic acid[260] or under solvent-free conditions.[257,262] In addition, Chauveau *et al.*[263] have shown that the reaction can be carried out in water under microwave irradiation, achieving yields over 70%. Some studies have also involved the use of catalysts, such as Brønsted acidic ionic liquids, to increase yields and reaction rates or to avoid the use of organic solvents.[228,257,261,262,264] The general mechanism for the formation of triarylimidazoles in this way is shown in *Scheme 3.4.*



Scheme 3.4 Mechanism of the condensation reaction between benzaldehyde, benzil and ammonium acetate.

Azizi *et al.*[259] optimised the catalyst-free reaction conditions by varying the organic solvent and type of amine or ammonium. Ethanol (EtOH) and methanol (MeOH) were found to give the highest yields of 70% and 97% respectively with ammonium acetate (NH_4OAc) at 60 °C for 4 h. Another catalyst-free study used glacial acetic acid, by refluxing the reaction at 118 °C for 4 h to give a *p*-nitro-substituted analogue in 92% yield. The product was then reduced to the amine using iron powder and 2 drops of 38% HCl, which required heating at 80 °C for 6 h.[260] The advantages of these methods were the efficient work-ups and simple procedures for product isolations, requiring no catalyst removal. As a result, these were the methods followed in the synthesis of the selected analogues.

3.1.2.2 Synthetic targets for the triarylimidazole scaffold

Seven βH-inhibiting triarylimidazoles were identified as active against the D6 strain of *P. falciparum* in the HTS. The three compounds with the greatest activity (<1 µM) against the parasite contained the same substitution pattern on the 2-phenylimidazole (ring A), specifically *p*-hydroxy and *m,m*-dimethoxy groups. Interestingly, removal of the dimethoxy groups in **VU0132562** to give **VU0099289** resulted in an increase in βH activity as indicated by a drop in the IC_{50} from 12.6 to 5.7 µM, but with a large decrease in parasite activity from 0.31 to 9.1 µM. The substituents on the 4,5-phenylimidazole (rings B and C) also influenced activity with the *p*-bromo derivative (**VU0132562** – *Figure 3.3a*) showing a 2-fold better activity than the *p*-hydroxy (**VU0094619** – *Figure 3.3c*). These compounds are shown in *Figure 3.3.*

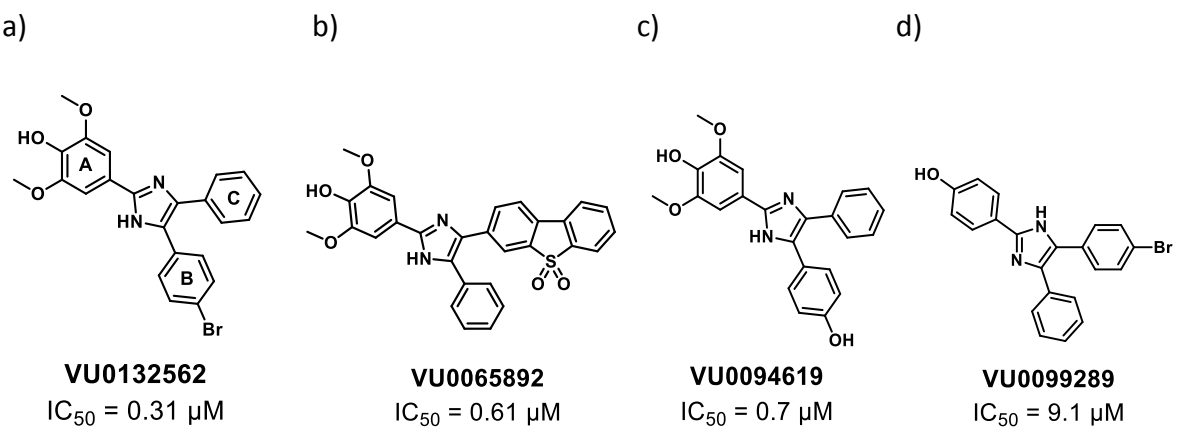


Figure 3.3 Examples of triarylimidazoles found in the HTS showing D6 parasite activities.

These intriguing structure-activity relationships encouraged the synthesis of a small library in order to firstly discover the structural requirements for βH inhibition, and secondly, to probe the influence of the substituents of ring A, specifically to determine whether other functionalities could maintain activity. *Figure 3.4* displays the proposed variations of the triarylimidazole scaffold, including the removal of the B and C phenyl rings to give 2-phenylimidazole, which was commercially available.

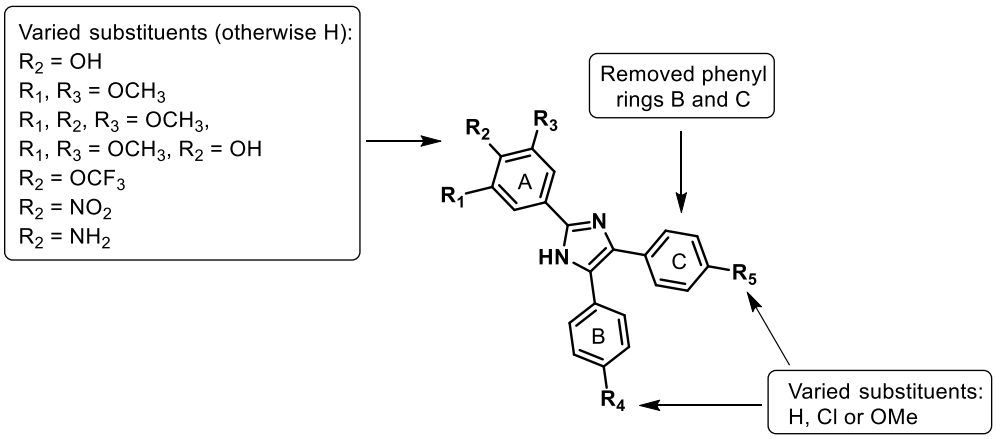


Figure 3.4 Derivatives of the triarylimidazoles obtained from commercial sources or via synthesis.

3.1.3 Specific goals

The specific goals for Chapter 3 were to:

- Synthesise and characterise a series of mono-, di- and tri-benzamide derivatives.
- Synthesise and characterise a series of triarylimidazole derivatives.

3.2 Results and discussion

3.2.1 Scaffold 1: Benzamides

3.2.1.1 Synthesis of the monobenzamides

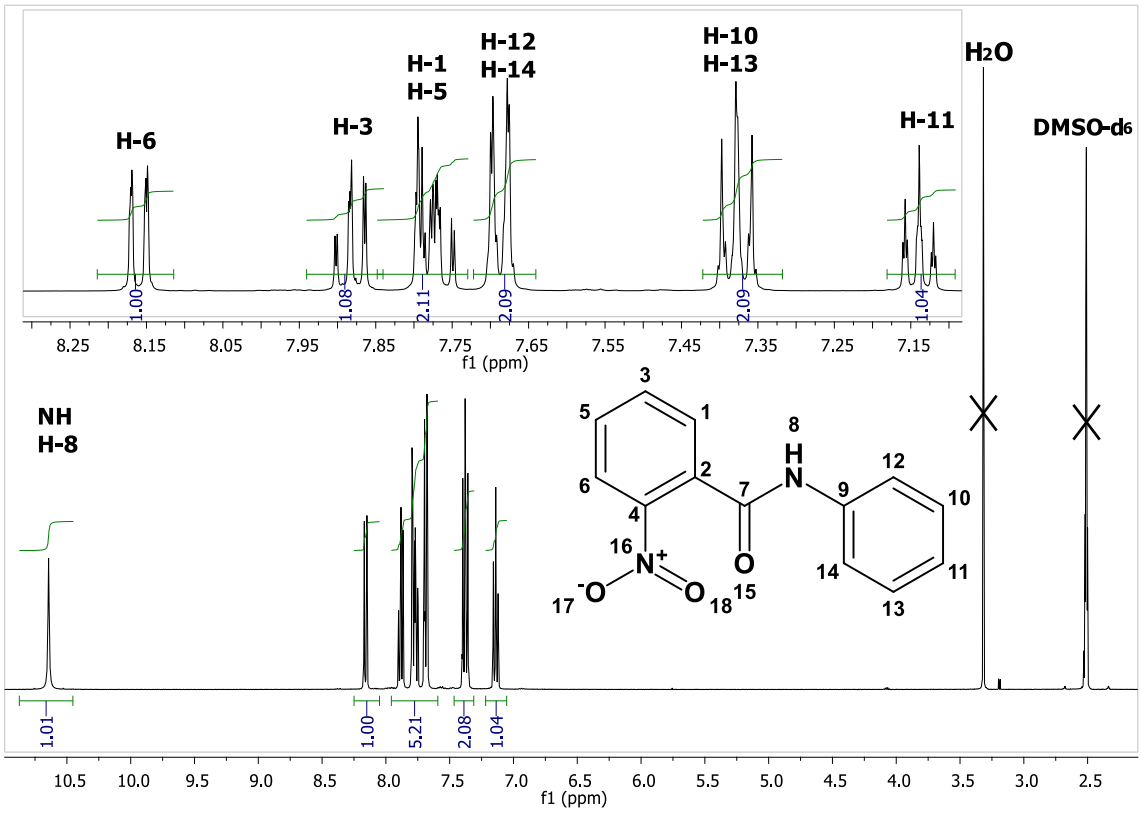
Synthesis was initially directed at generating benzamide derivatives of scaffold 1 in which the synthetic procedure for amide formation was optimised on the simplest target molecule, 2-nitro-*N*-phenylbenzamide (**2a**). The general method in the literature, whereby aniline and 2-nitrobenzoyl chloride (1:1) are stirred in dichloromethane (DCM) with Et_3N (1 eq.) for 2 h at RT, was used as a guideline.[265,266] However, pyridine was chosen as a solvent instead of, or in addition to DCM. This variation allowed for faster, cleaner reactions owing to the improved solubility of the starting materials and the ability of pyridine to act as an acyl-transfer catalyst. Furthermore, pyridine provided the base to consume the expelled HCl and therefore the addition of Et_3N was unnecessary for reactions with aniline as the amine.

A work-up of the pyridine-stirred reaction with ethyl acetate (EtOAc) and 1 M HCl enabled isolation of the product in the organic layer, eliminating the pyridine as a pyridinium salt dissolved into the aqueous layer. However, several washings of the EtOAc layer with 1 M HCl were required, otherwise a fraction of pyridine remained in the organic layer and had to be removed later under reduced pressure.

The reaction product did not require purification by column chromatography but the product could be crystallised directly to give an excellent yield of 99%. The synthesis of **2a** was confirmed by nuclear magnetic resonance (NMR) analysis (*Figure 3.5*) in deuterodimethyl sulfoxide (DMSO-d_6). All NMR analyses of the benzamides were carried out in this solvent, since it allowed for the NH peaks to be detected, an important indicator of amide bond formation, and was able to dissolve all the derivatives, many of which were too polar for dissolution in deuterochloroform ($CDCl_3$). Peak assignments were confirmed with correlation spectroscopy (COSY) and heteronuclear single-quantum correlation (HSQC) spectroscopy (*Figure 3.6*). The aromatic protons were observed in two distinct coupling systems with those on the nitro-substituted ring more downfield owing to the resonance-based electron-withdrawing character and deshielding effect of a nitro group. The ortho-proton adjacent to the nitro (H-6) group was assigned as the most deshielded on the basis of a nitro substituent being more electron-withdrawing than a carboxamide. Assignment of H-3 as the triplet at 7.87 ppm and H-5 as the triplet at 7.80 ppm was based on the lack of

coupling in the COSY spectrum of H-3 to H-6, the latter resonating at 8.15 ppm, while H-5, adjacent to H-6, did show a cross-peak to indicate coupling. The downfield NH proton peak and C=O carbon peak supported the successful formation of the amide bond. The fully assigned spectra are shown below.

a)



b)

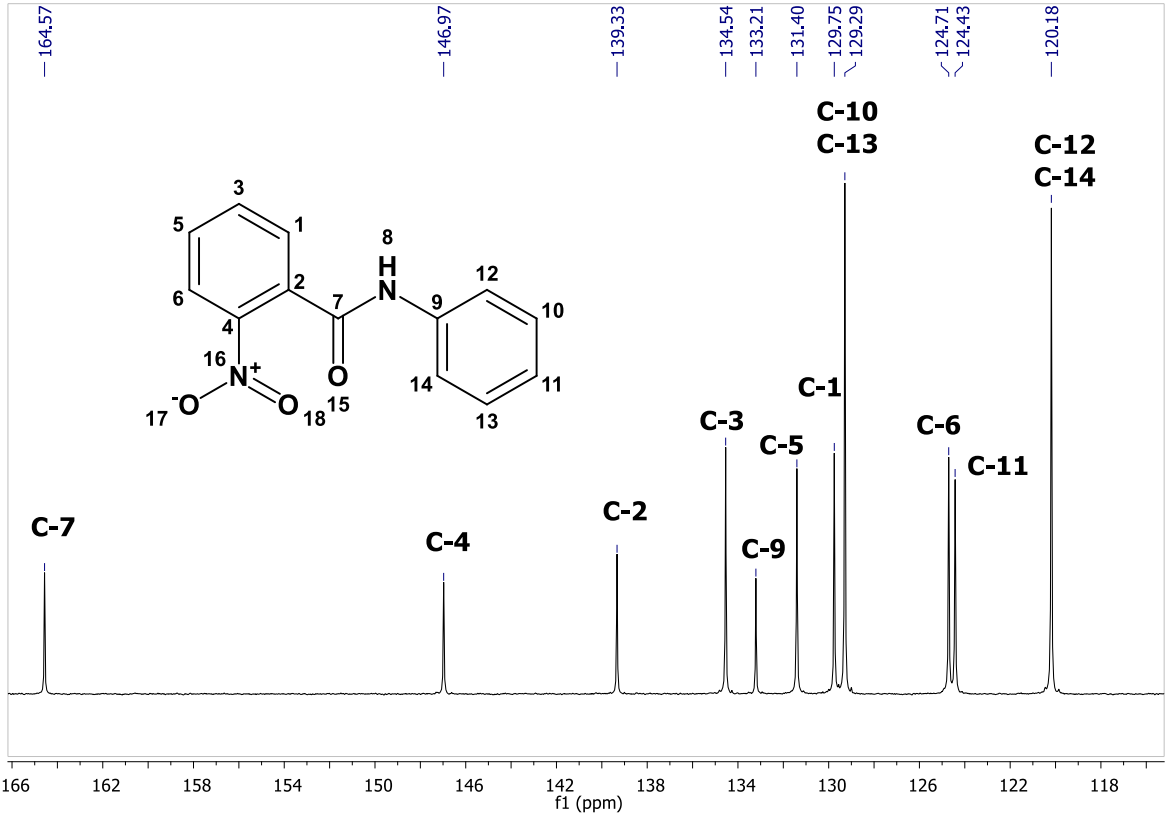
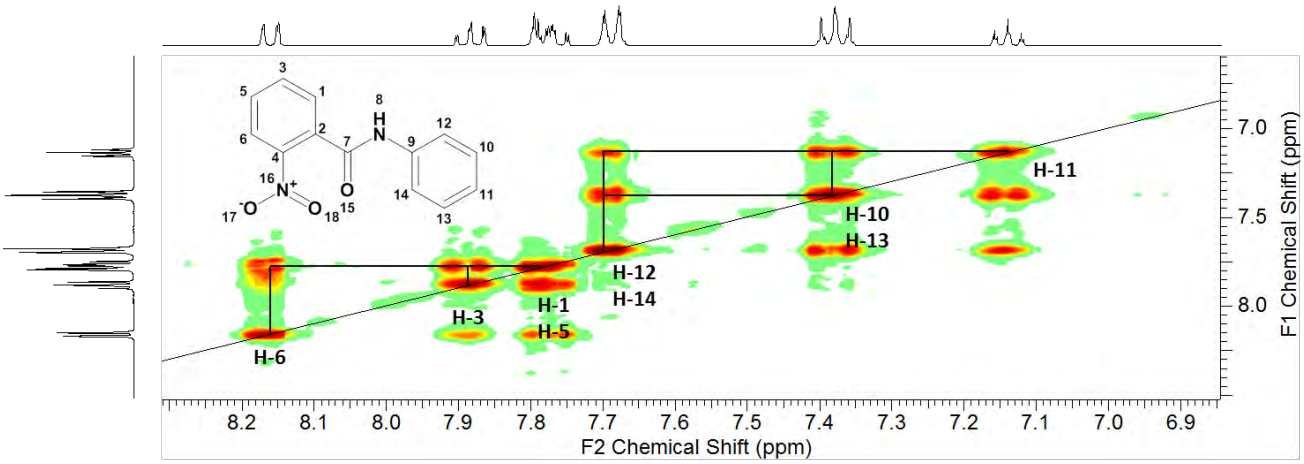


Figure 3.5 The assigned a) ^{1}H and b) ^{13}C NMR spectra for **2a** in d_6-DMSO.

a)



b)

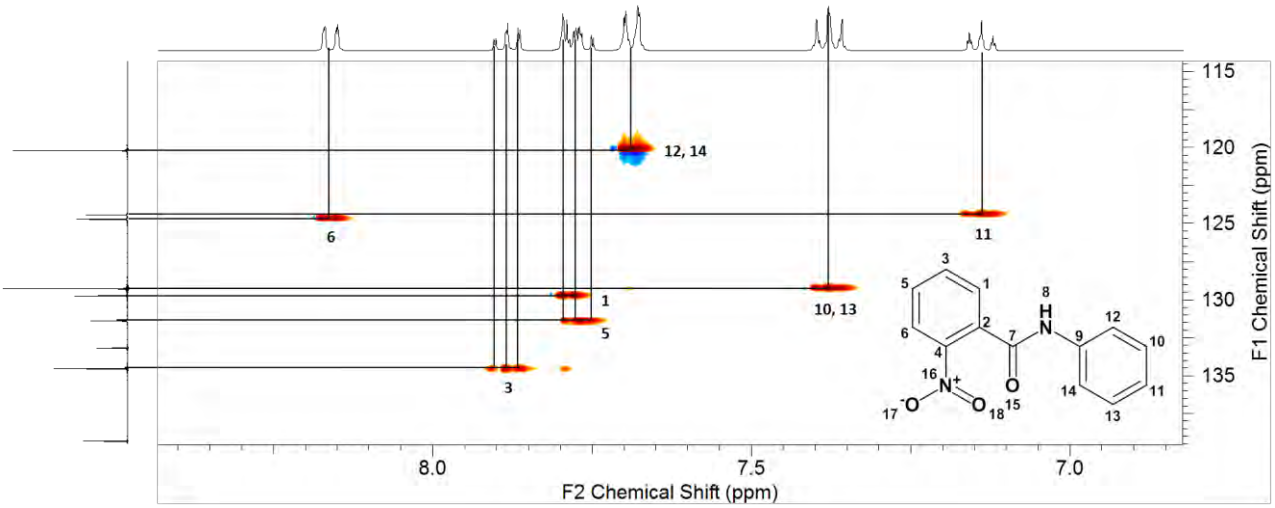
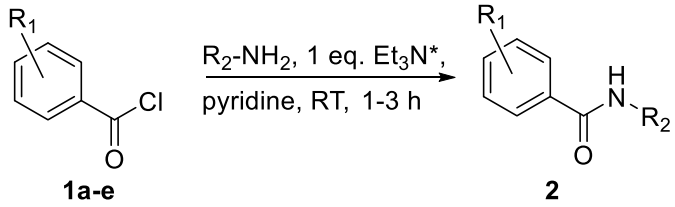


Figure 3.6 The assigned a) COSY and b) HSQC spectra showing the correlation of ^{13}C and ^{1}H peaks for assignment of compound **2a**.

The remaining phenyl and pyridyl monobenzamides (**2b-e**) were synthesised from the relevant commercially available acid chlorides (**1b-e**) as per *Scheme 3.5*. Et_3N (*) was only used in the case of the pyridyl derivatives to ensure that the neutral species of the product was obtained. These derivatives were fully characterised by ^{1}H and ^{13}C NMR spectroscopy, where the correct multiplicity and number of resonances supported structural identity. Furthermore, all compounds were sent for high resolution mass spectrometry (HRMS) in which observed $[M-H]^+$ ions corresponded to calculated theoretical masses within 5 ppm. Purity was assessed by melting point and high pressure liquid chromatography (HPLC) analysis. All the compounds had melting point ranges of ≤3 °C and HPLC purities of >95.0%.



2a: R_1 = 2-NO_2; R_2 = phenyl
2b: R_1 = 3,5-diNO_2; R_2 = phenyl
2c: R_1 = 3,5-diNO_2; R_2 = 4-pyridyl
2d: R_1 = 3,5-diNO_2; R_2 = 3-pyridyl
2e: R_1 = 3,5-diNO_2; R_2 = 2-pyridyl

Scheme 3.5 Synthesised derivatives of the monobenzamides.

In some of the targets, particularly those containing a pyridyl ring, removal of pyridine using 1 M HCl in the work-up was inappropriate. This was due to the similarity in pK_a of the molecule's conjugate acid to that of the pyridinium ion. A higher value for the conjugate acid pK_a indicates a stronger base since the concentration of protonated base (BH^+) in the equilibrium equation increases (*Eq. 3.1*).

$$\mathbf{BH^+ + H_2O} \underset{}{\overset{\boldsymbol{K_a}}{\rightleftharpoons}} \mathbf{B + H_3O^+} \qquad \textbf{Eq. 3.1}$$

$$\boldsymbol{K_a} = \frac{[\mathbf{B}][\mathbf{H_3O^+}]}{[\mathbf{BH^+}]}$$

$$\mathbf{p}\boldsymbol{K_a} = \log\frac{[\mathbf{BH^+}]}{[\mathbf{B}][\mathbf{H_3O^+}]}$$

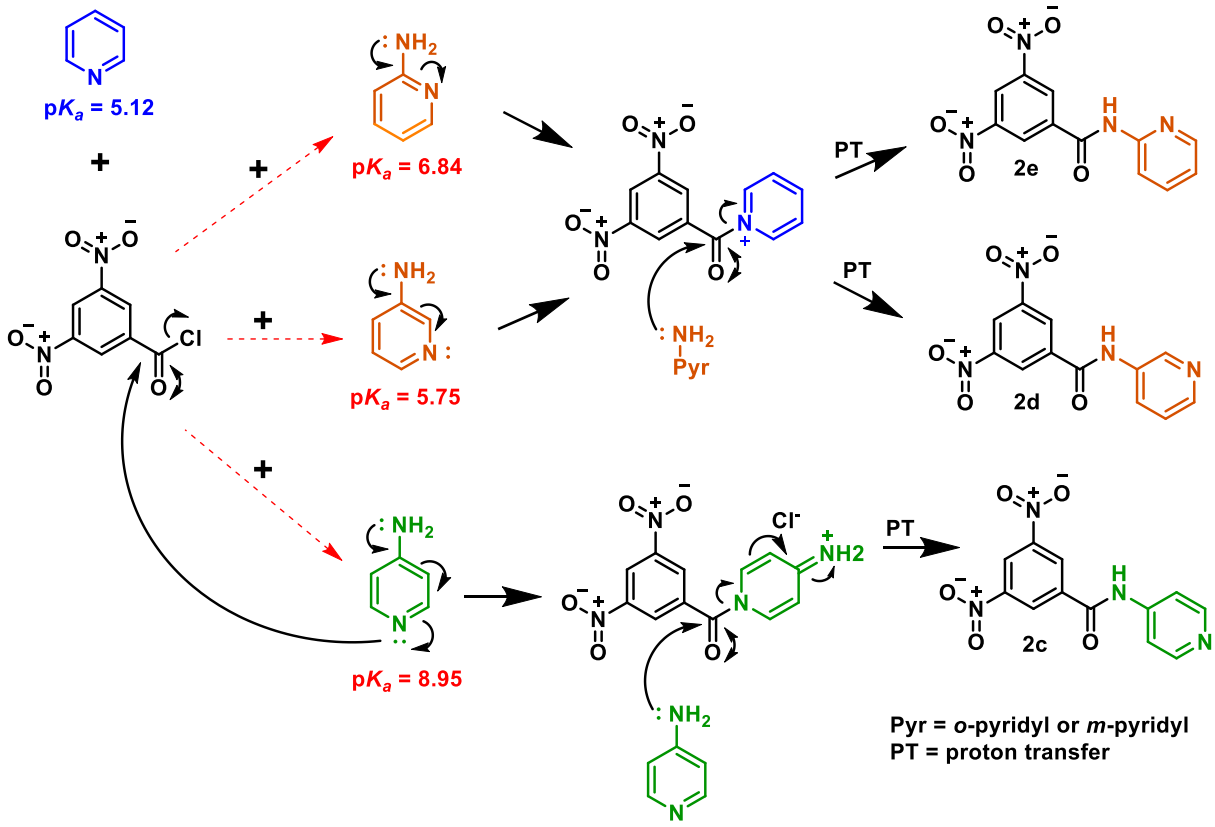
In this case, azeotropic removal of the pyridine with toluene on the rotary evaporator at ≥60 °C could be employed to remove the majority of the pyridine. Toluene and pyridine form a positive azeotrope whereby the boiling point of the mixture is lower than that of either constituent. For example, at atmospheric pressure, toluene and pyridine boil at 110.63 °C and 115.3 °C respectively, while the mixture boils at 110.0 °C with a composition of 78% and 22% respectively.[267,268] When the pressure is reduced, the boiling point decreases allowing toluene and pyridine to be removed together at a lower temperature than pyridine alone.

This method was tested for 3,5-dinitro-*N*-phenylbenzamide (**2b**), which achieved complete removal of the pyridine under reduced pressure with toluene without using an acidic work-

up. This was important for the synthesis of the pyridyl derivatives, in view of the similarity of the two pK_a's (5.12 for pyridine; 5.62 for 3,5-dinitro-*N*-(pyridin-4-yl)benzamide (**2c**)).[269]

In the cases where 4-aminopyridine was the starting material, the reaction proceeded via the benzoylpyridin-4(1H)-iminium chloride salt (*Scheme 3.6*). This is because the nitrogen of 4-aminopyridine (predicted conjugate acid pK_a of 8.95) is almost 10,000 times more basic than that of pyridine (predicted conjugate acid pK_a of 5.12) and much more nucleophilic as a result of resonance from the amine into the pyridyl ring. Therefore it is preferentially this nitrogen that reacts with the acyl chloride carbonyl to form an acyl-pyridinium ion intermediate which is then irreversibly converted to the product by transfer to the amino group of an unacylated 4-aminopyridine molecule. This process is slower since the pyridyl amine is less nucleophilic than the pyridyl nitrogen, but is driven by the charge on the iminium ion.

The reaction to form **2d**, the 3-pyridyl derivative, involved 3-aminopyridine as the reactant, which has a conjugate acid pK_a of 5.75, comparable to that of pyridine. The amino group nitrogen lone pair is not able to resonate onto the pyridyl nitrogen, increasing the nucleophilicity of the amino group nitrogen but decreasing that of the ring nitrogen, relative to those in 4-aminopyridine. In this case, it is most likely that pyridine, being in a large excess, acts as the transfer agent. This is likely also the case for the synthesis of **2e** with 2-aminopyridyl (conjugate pK_a of 6.84), where although resonance communication can occur, steric factors decrease the nucleophilicity of the ring nitrogen relative to that of 4-aminopyridine. There did not appear to be a significant difference in reaction times between reactions forming 3-, 4- and 2-pyridyl compounds when pyridine was used as a solvent. *Scheme 3.6* summarises these ideas and shows the mechanism for the reaction with 4-aminopyridine, which possesses the largest conjugate acid pK_a and the most nucleophilic ring nitrogen. This was the only aminopyridine starting material that was likely to act as the acyl transfer catalyst, rather than pyridine.



Scheme 3.6 The mechanism for the formation of pyridylbenzamides **2c-e**.

3.2.1.2 Synthesis of the dibenzamides

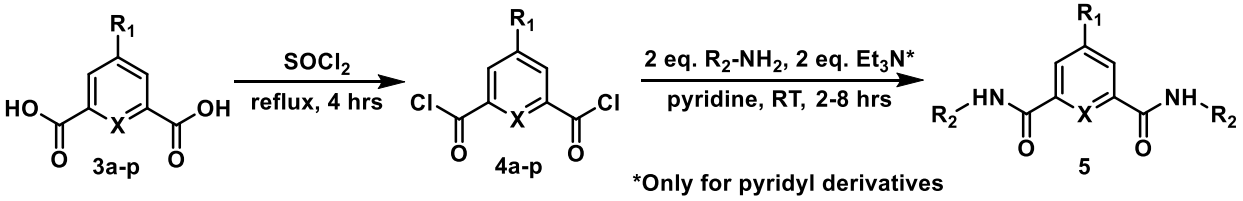
As with the mono-benzamides, the di and tri-benzamides were also synthesised via N-acylation using the acid chloride. However, in the case of the dibenzamides, the relevant isophthalic dichloride had to be prepared fresh using $SOCl_2$, as they were unstable or not commercially available. Once prepared, the acid chlorides were not isolated but kept under vacuum or under N_2 and then dissolved in pyridine within the same reaction flask.

3.2.1.2.1 Symmetrical dibenzamides

The pyridyl derivatives **5a**, **5b, 5d, 5f, 5h, 5i, 5k, 5m** and **5o** (*Scheme 3.7*) were synthesised from the appropriate freshly prepared isophthaloyl dichlorides and the appropriate amine as in **2c** but with 2 equivalents (eq.) of Et_3N. In some cases, the reactions were slower than that of the monobenzamides and were left overnight to insure that both amide bonds were formed and that purification by column chromatography was not necessary. In general the 5-nitro (**5a** and **5b**), 5-cyano (**5d**) and 5-methoxy (**5m**) derivatives needed longer to react than the 5-methyl (**5f**), 5-*tert*-butyl (**5h** and **5i**) and the unsubstituted (**5k** and **5o**)

derivatives. The reasons for these differences are not known but may involve factors such as solubility or differences in quantities synthesised. The reaction times did not correlate with electron-withdrawing versus releasing capabilities of the substituents since the derivatives with the electron-withdrawing nitro and cyano substituents did not react faster than the electron releasing methoxy substituent. This is possibly due to the *meta* relationship between the acyl groups and the position-5 substituents of target molecule **5**, which are not in direct electronic contact through resonance. Basic workups with sodium carbonate ensured that the pyridyl product was neutralised and could be extracted in EtOAc. Pyridine was then completely removed with toluene via azeotropic evaporation.

The phenyl derivatives **5c**, **5e**, **5g**, **5j**, **5l**, **5n, 5p** (*Scheme 3.7*) were also synthesised from the appropriate freshly prepared isophthaloyl dichlorides but with aniline instead of the aminopyridine. The reactions were faster and cleaner than those of the pyridyl derivatives, partly owing to the fact that aniline, in contrast to 2-, 3- or 4-aminopyridine, is a liquid which assisted with solubility. Another reason for the shorter reaction times was the better nucleophilicity of the aniline amine relative to the pyridyl amines, where the amine electrons are delocalised onto the ring nitrogen. In order to accurately deliver the aniline, a micropipette was used, however, any excess aniline was removed in the workup with 1 M HCl, since the predicted pK_a of the conjugate acid of aniline is 4.64. The pyridine was also removed in this way, allowing for efficient isolation of the products in good to excellent yields.

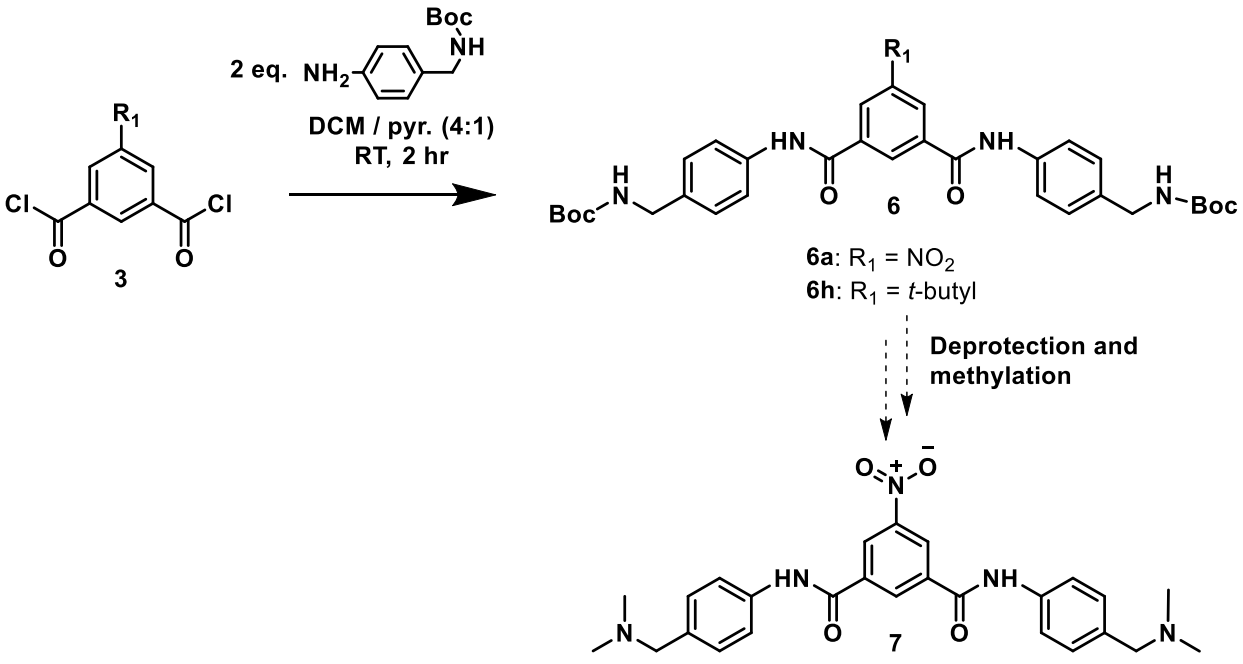


5a: X = CH; R_1 = NO_2; R_2 = 4-pyridyl
5b: X = CH; R_1 = NO_2; R_2 = *3*-pyridyl
5c: X = CH; R_1 = NO_2; R_2 = phenyl
5d: X = CH; R_1 = CN; R_2 = *4*-pyridyl
5e: X = CH; R_1 = CN; R_2 = phenyl
5f: X = CH; R_1 = CH_3; R_2 = *4*-pyridyl
5g: X = CH; R_1 = CH_3; R_2 = phenyl
5h: X = CH; R_1 = *t*-butyl; R_2 = *4*-pyridyl
5i: X = CH; R_1 = *t*-butyl; R_2 = *3*-pyridyl
5j: X = CH; R_1 = *t*-butyl; R_2 = phenyl
5k: X = CH; R_1 = H; R_2 = *4*-pyridyl
5l: X = CH; R_1 = H; R_2 = phenyl
5m: X = CH; R_1 = OMe; R_2 = *4*-pyridyl
5n: X = CH; R_1 = OMe; R_2 = phenyl
5o: X = N; R_1 = H; R_2 = *4*-pyridyl
5p: X = N; R_1 = H; R_2 = phenyl

Scheme 3.7 Synthesised derivatives of symmetrical phenyl and pyridyl dibenzamides.

Many studies have suggested that increased accumulation of compounds in the parasite has a large influence on their activity *in vitro*.[98,58,99,101,108] For the 4-amino quinoline compounds, this is believed to be as a result of pH trapping in the acidic DV, owing to the basic amine side chains in the known antimalarials CQ, AQ, QN and QD.[182] In this case, derivatives with basic side chains that can be protonated at pH 4.8, may be more active owing to larger VARs. This is discussed in detail from *Sections 4.2.3*.

Synthesis of a derivative of the hit compound from the HTS (**5a**) with basic amino side chains in the form of target molecule **7** (*Scheme 3.8*), was attempted in several different ways. Initially this involved the formation of Boc-protected intermediate **6a** and **6h** from the relevant acid chloride and 4-[(*N*-Boc)aminomethyl]aniline. Although **6a** was intended as the appropriate intermediate to target molecule **7**, trial reactions were performed using **6h**, the *t*-butyl analogue, owing to the availability of larger quantities of the starting material **3h** than **3a**. The protecting group on the amine was necessary since the aliphatic methylamine is more nucleophilic than the aromatic amine, which releases electron density into the ring via resonance.



Scheme 3.8 Formation of target molecule **7** from the Boc-protected intermediate **6**.

Formation of **6h** was supported by TLC analysis and confirmed by ^{1}H NMR analysis in MeOD (*Figure 3.7*). Aromatic peaks corresponding to all three phenyl rings were observed, along with the 4.34 ppm resonance, integrating for the four aminomethyl protons corresponding to H-20 and H-28. In addition, the two magnetically equivalent Boc *t*-butyl functionalities and the one *t*-butyl at R_1 were observed upfield in the spectrum at 1.58 and 1.56 ppm respectively. The amide and carbamate NH protons were not observed in the spectrum owing to hydrogen-deuterium exchange between the solvent and **6h**.

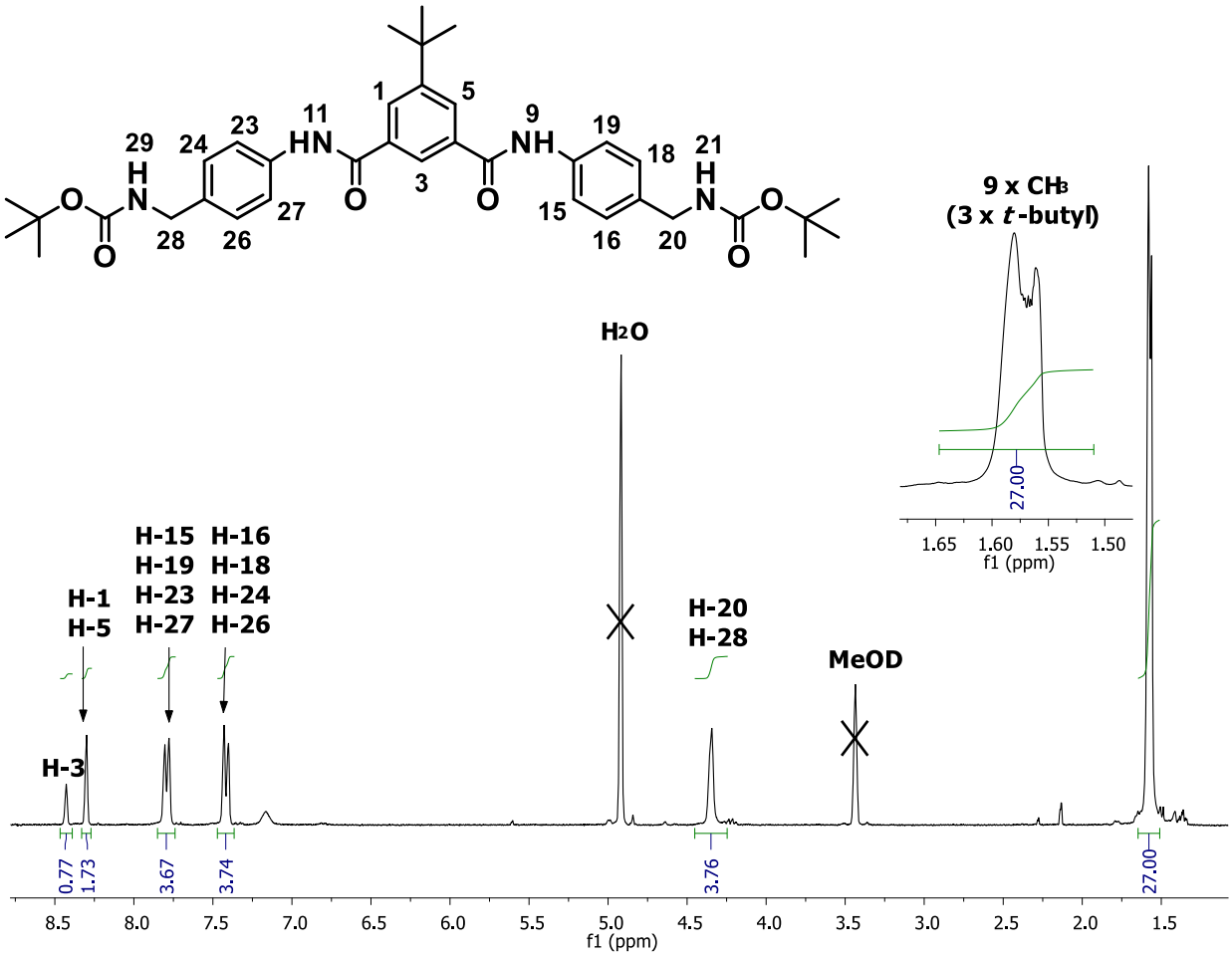
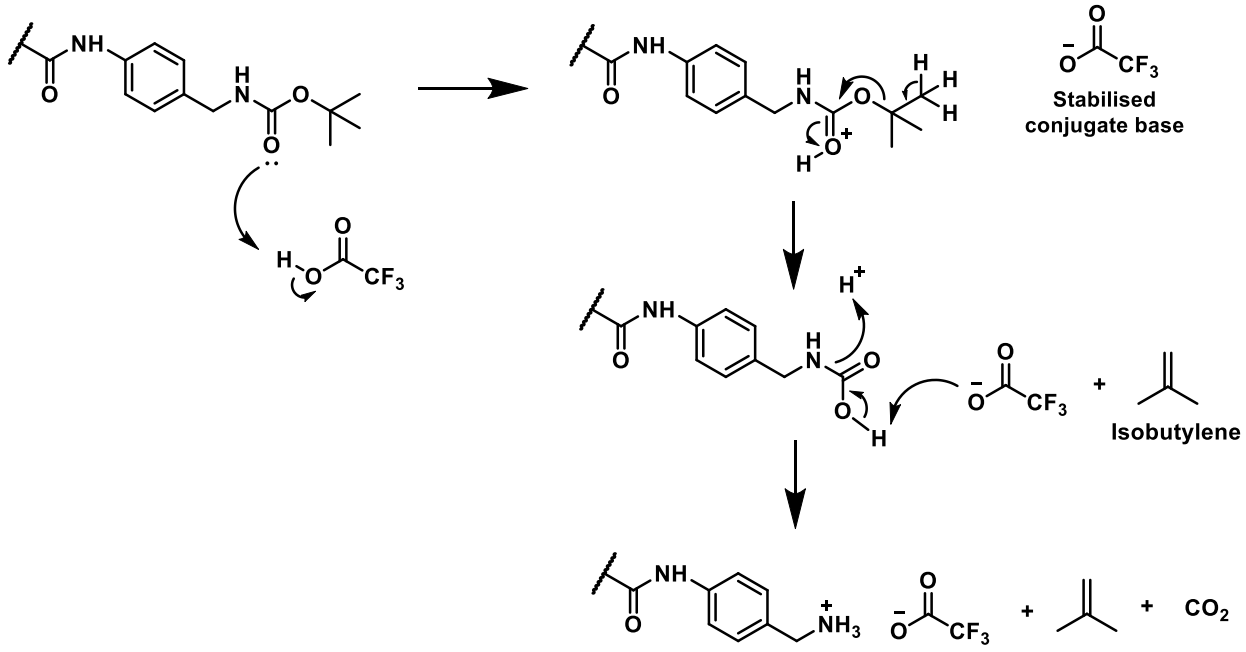


Figure 3.7 The ^{1}H NMR spectrum of intermediate **6h** in MeOD.

In the next step, the Boc group of **6h** was deprotected using trifluoroacetic acid (TFA) in dry DCM. This acid is particularly strong, owing to the electronegativity of the fluorine atoms, which not only destabilise the O-H bond, but also stabilise the anion of the conjugate base. As a result, the deprotection was carried out at 0 °C to ensure that the amide was not

hydrolysed by the TFA. Deprotection proceeded via protonation of the carbamate followed by the entropy-driven loss of isobutylene and then the liberation of carbon dioxide to form the ammonium trifluoroacetate salt (*Scheme 3.9*).



Scheme 3.9 Boc deprotection of **6** with TFA.

The low boiling DCM and excess TFA could be easily removed under reduced pressure after which the free amine was isolated by stirring the product with potassium carbonate (K_2CO_3) in MeOH. This step resulted in an ion exchange to give the free amine **6**, potassium trifluoroacetate and potassium bicarbonate. The differential solubility of these species in MeOH, in which only the free amine remained dissolved, enabled separation by filtration through Celite. The Boc-deprotection was confirmed by the disappearance of the Boc *t*-butyl groups and the upfield shift of the methylamine protons in the ^{1}H NMR spectrum of **6h** after deprotection (*Figure 3.8*). As in the spectrum of **6h** before deprotection (*Figure 3.7*), NH protons were not observed in the solvent, MeOD.

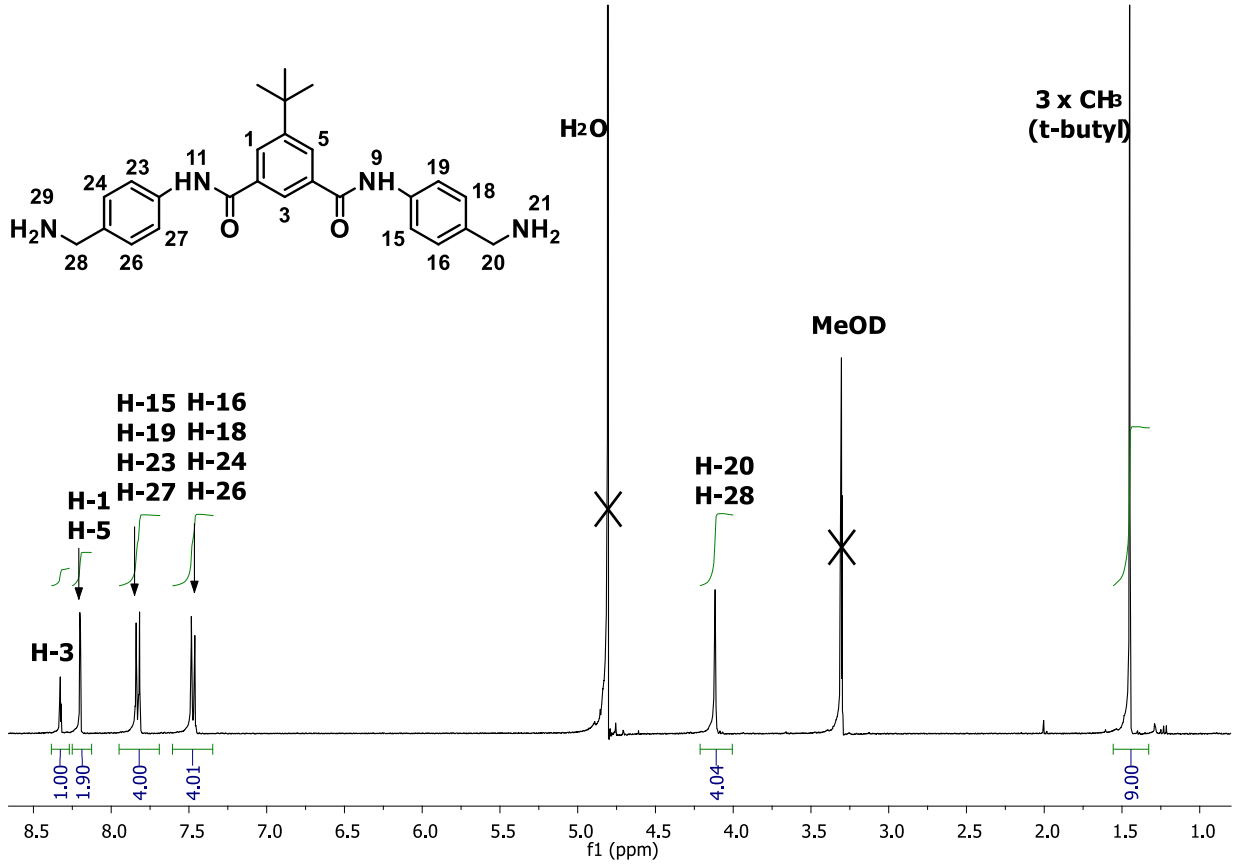
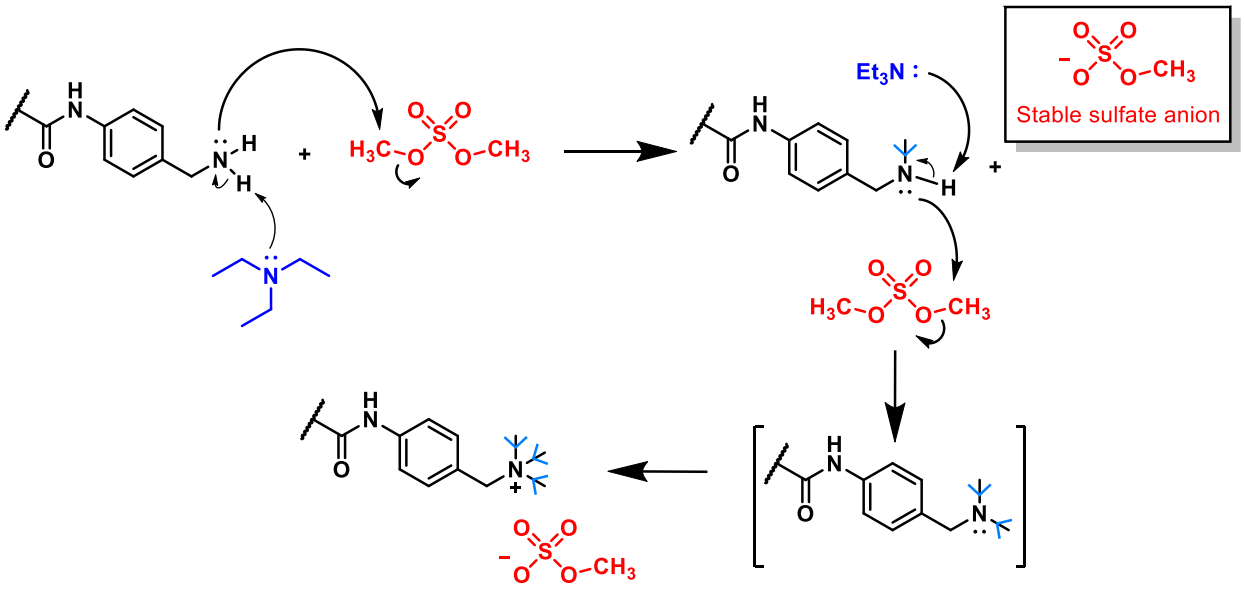


Figure 3.8 The ^{1}H NMR spectrum in MeOD of the **6h** after Boc deprotection.

Following deprotection, methylation was first attempted with dimethylsulfate (DMS), a relatively hard electrophilic methylating agent. Trimethylamine (Et_3N), as an external base, was needed to neutralize the $MeHSO_4$ species from the methylation. However, adding the Et_3N prior to the DMS resulted in a possible intramolecular rearrangement of the deprotected **6h** to a N-(4-aminobenzyl)acetamide within 15 mins, most likely as a result of a base-catalysed rearrangement involving the nucleophilic primary amine. Evidence of this was indicated by TLC analysis, in which less polar UV active spots that stained dark purple with ninhydrin spray were observed before the addition of DMS.

TLC evidence of the reaction with 5 eq. DMS, followed by 7 eq. of Et_3N, gave evidence of quaternisation to the ammonium salt, visualised as a polar UV-active spot on the TLC baseline. This can be explained by the increase in nucleophilicity due to inductive effects of going from secondary (2°) to tertiary (3°) amine status. *Scheme 3.10* outlines the mechanism of formation of 2°, 3° and 4° amines via methylation of the primary amine with

DMS (red). The blue arrows represent the flow of electrons onto the nitrogen atom via an inductive effect.



Scheme 3.10 Methylation to the quarternary amine with dimethylsulfate (DMS).

Owing to the added complication of having two sides of the molecule to methylate, the second strategy that was pursued involved the methylation of 4-(aminomethyl)aniline using DMS, to form 4-((dimethylamino)methyl)aniline, before amide bond coupling. This reaction involved using 2 eq. of DMS followed by 3 eq. of Et_3N. Unfortunately, TLC analysis of this reaction also gave evidence in support of quaternisation, and a mixture of inseparable amines was produced. This can be seen in an example of a ^1H NMR spectrum of the crude reaction product, obtained in very low yield from one of the reactions after extraction into EtOAc (*Figure 3.9*). Here, the 2° and 3° amines were observed in approximately a 1:1.2 ratio, having removed the 4° product into the aqueous layer during the work-up.

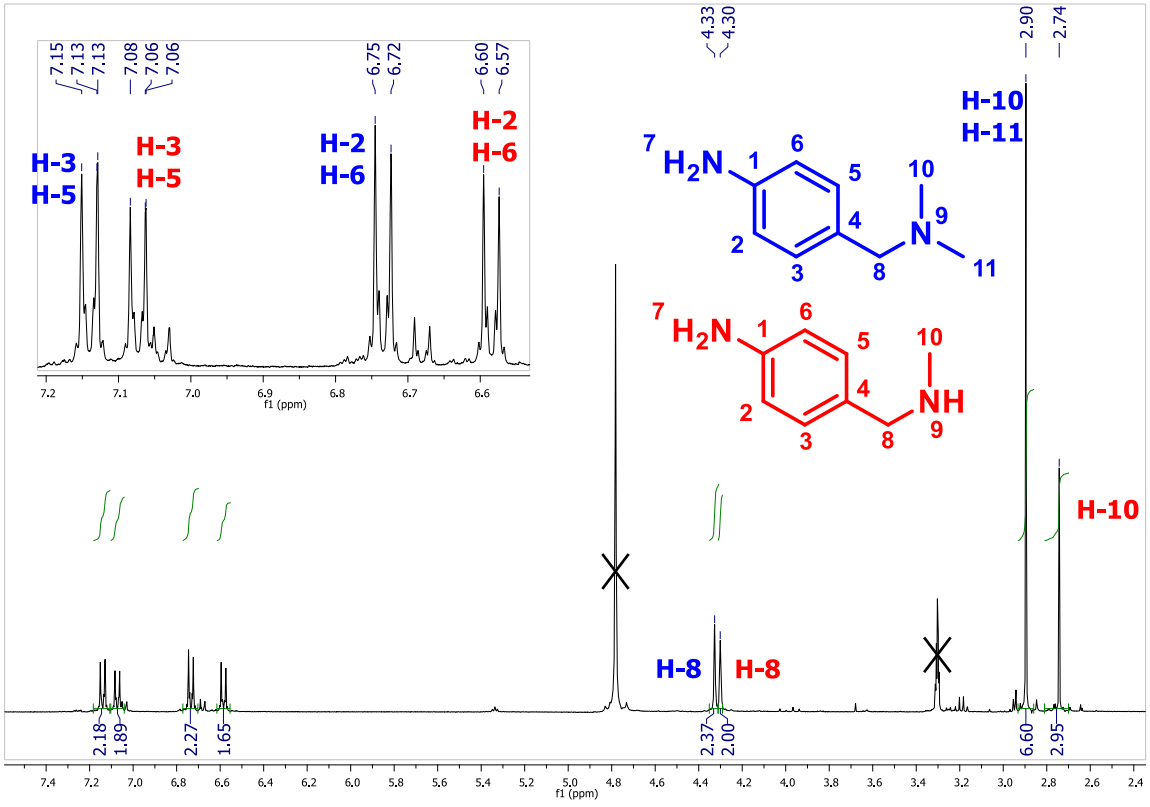
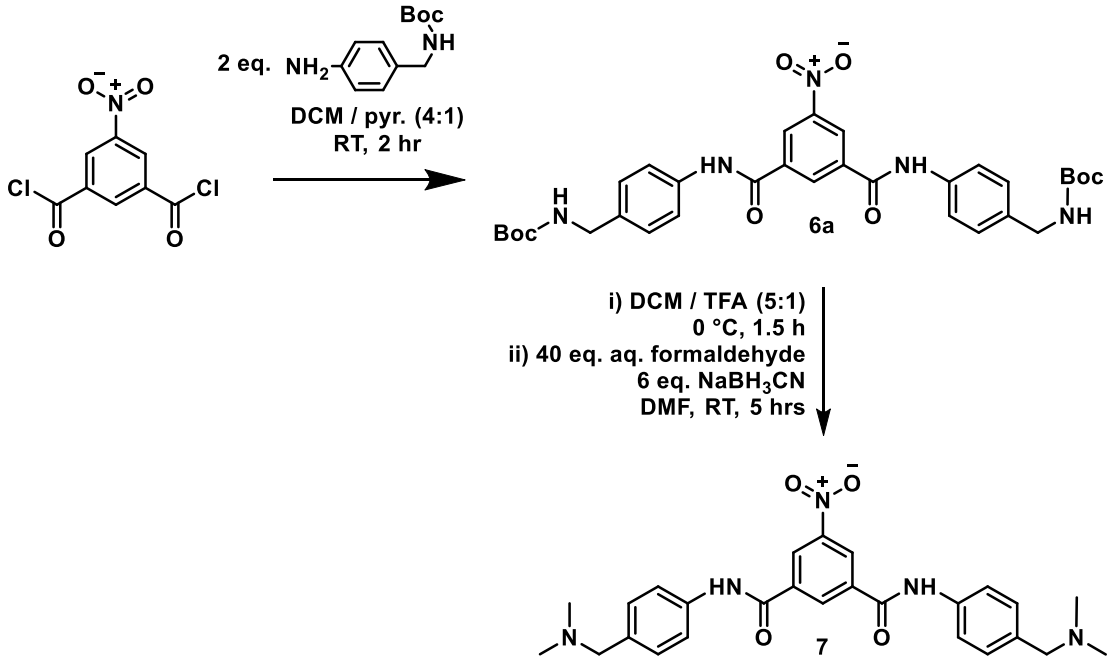


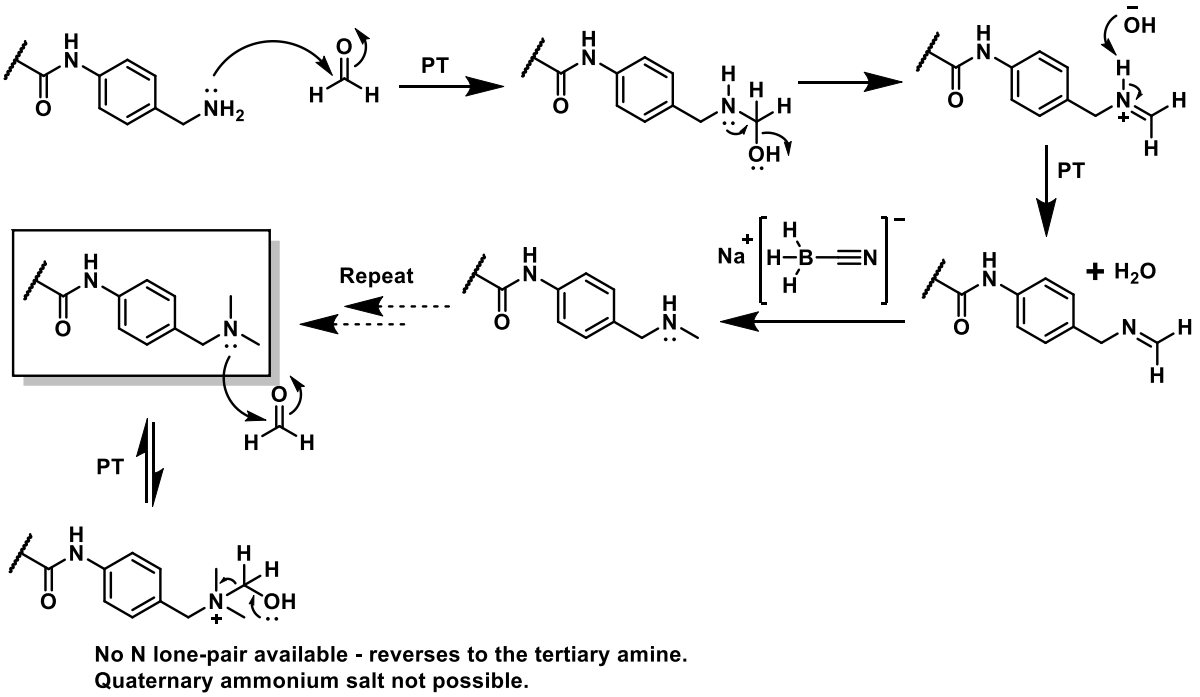
Figure 3.9 The ^{1}H NMR spectrum in MeOD of the methylation reaction with DMS.

Hence, as a result of not being able to control the chemoselectivity of N-alkylation, an alternative, more selective method was adopted, as reductive amination, which involves a different C-N bond-formation mechanism (*Scheme 3.11*).



Scheme 3.11 Formation of **7** via reductive amination from **6a**.

Specifically, the primary amine from Boc deprotection was formylated with formaldehyde to produce the imine, which was reduced with sodium cyanoborohydride ($NaBH_3CN$) in situ.[270,271] The resultant secondary amine then participated in the sequence again resulting in formation of the tertiary amine ultimately, using an excess of formaldehyde (40 eq.) and $NaBH_3CN$ (6 eq.) The one-pot reaction allowed both sides of the dibenzamide to be alkylated and, importantly, avoided the formation of the quaternary ammonium ion, which was not possible via this mechanism. The reason for this is presented in *Scheme 3.12*, which shows the necessity of the nitrogen lone pair for elimination of water. Sodium cyanoborohydride was used since it is milder than sodium borohydride, owing to the electron-withdrawing nature of the cyano group. This helps to stabilise the negative charge on the boron so that the hydride anion is less nucleophilic, resulting in a reducing agent which is selective for imines. Furthermore, the cyano derivative is stable in water, which was used as a co-solvent in the DMF, whereas sodium borohydride reacts with water to form hydrogen gas.[252,272]



Scheme 3.12 The mechanism for the formation of the tertiary amine via reductive amination of the imine with sodium cyanoborohydride.

The reaction was carried out with aqueous formaldehyde in DMF, a solvent known to assist reactions that proceed via a polar mechanism,[252] such as the assisted E1-like elimination of water in this reaction (*Scheme 3.12*). This high-boiling solvent was partly removed in an aqueous work-up with EtOAc and then by evaporation under reduced pressure with heating. The NMR spectrum (*Figure 3.10*) of the purified product shows the assignment of the peaks for the target compound **7**, whose relative simplicity bears testimony to its C-2 symmetry. The expanded aromatic region represents the two doublets for the 1,4-substituted phenyl rings.

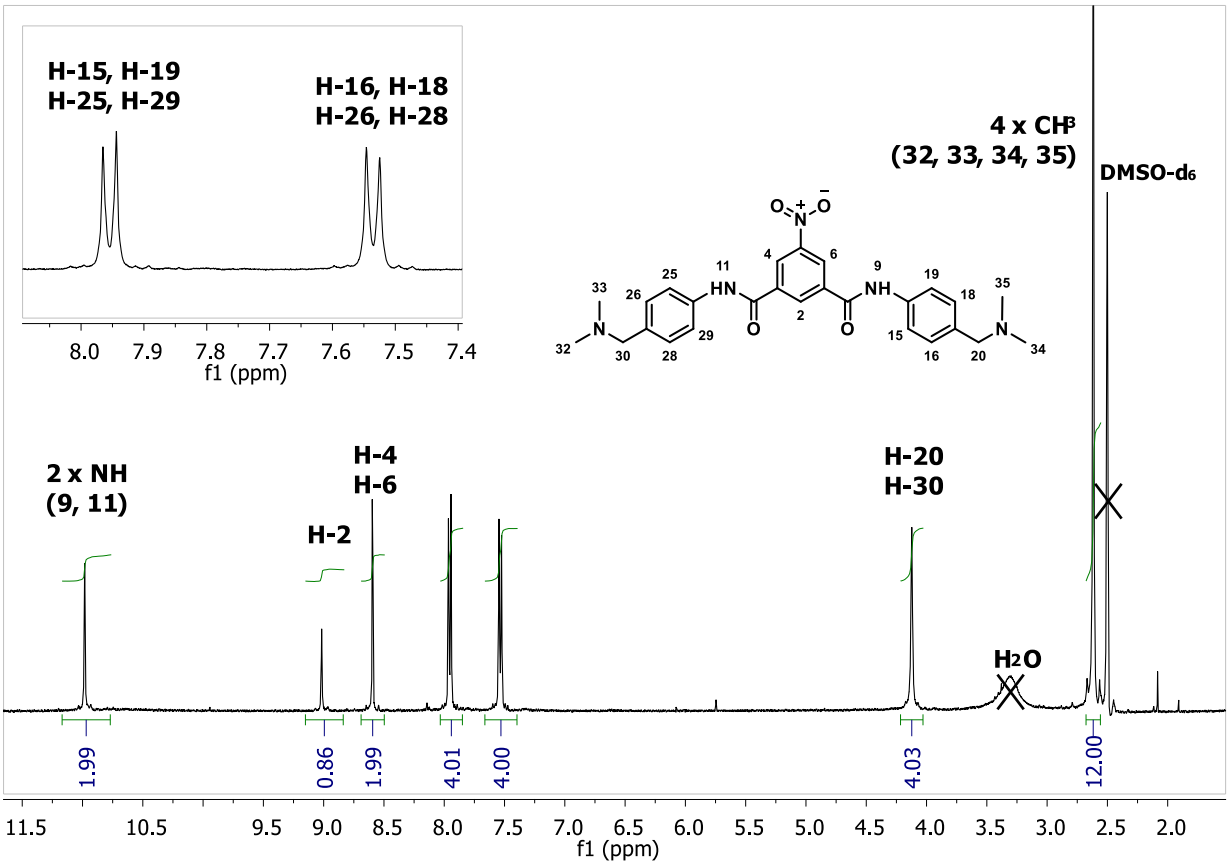
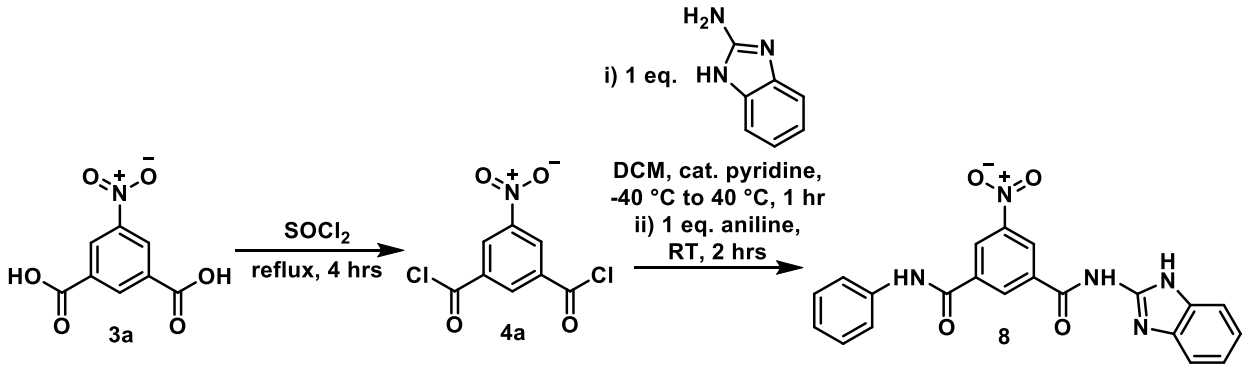


Figure 3.10 The ^{1}H NMR spectrum of **7** in DMSO-d$_6$.

3.2.1.2.2 Unsymmetrical dibenzamide (8)

In order to evaluate the effect of increasing the aromatic molecular surface area available for haem association, a fused ring system, in the form of a benzimidazole, was selected for the R_2 position in place of pyridyl or phenyl. This is discussed in detail in *Section 4.2.4.*

The dibenzamides discussed in *Section 3.2.1.2.1* were all symmetrical with respect to amide N-substituents and could therefore be synthesised easily with 2 eq. of the amine. Compound **8**, as the only unsymmetrical compound, was synthesised owing to the fact that having two benzimidazoles resulted in an insoluble compound, which made isolation and purification essentially impossible. Since this type of molecule was undesirable for further assaying and analysis, a monobenzimidazole as a mixed derivative **8** was synthesised via sequential addition of two different amines instead (*Scheme 3.13*).



Scheme 3.13 Formation of the unsymmetrical benzamide **8**.

The first attempt to achieve this involved reaction of the acid chloride with 1 eq. of aniline in an acetonitrile/liquid nitrogen cooling bath at -40 °C by slow addition. The lower temperature was chosen in order to slow down the rate of the reaction, with the intention of optimizing formation of the mono-phenylbenzamide. However, TLC analysis revealed remaining starting material together with two spots, most likely corresponding to the mono- and di-phenylbenzamide. Realising the inefficiency of this method, an alternative route was found, whereby the 5-nitroisophthaloyl dichloride was reacted with the less nucleophilic 2-aminobenzimidazole first before addition of the more nucleophilic aniline. The resonance of the aminobenzimidazole vs aniline amines into the aromatic ring are shown in *Figure 3.11*. The stabilising ability of the nitrogen in the benzimidazole results in a less nucleophilic amine than that of the aniline.

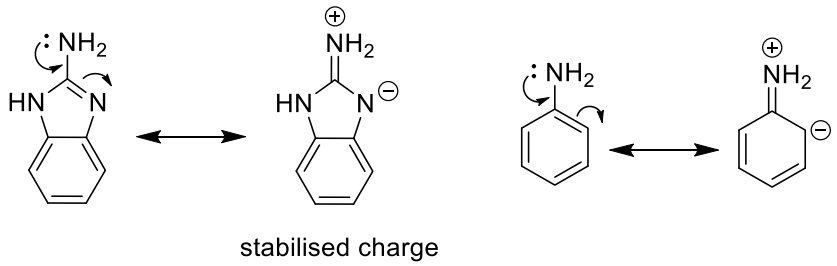
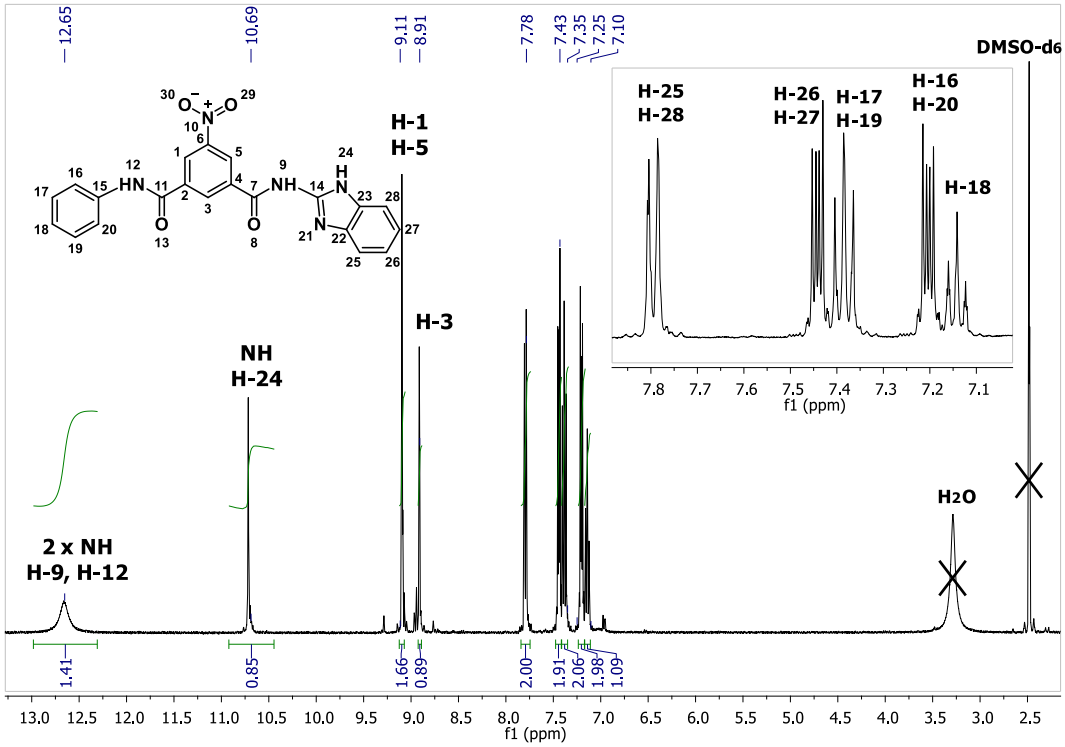


Figure 3.11 Resonance structures for the benzimidazole and aniline nucleophiles.

To this end, one third of 1 eq. of the 2-aminobenzimidazole was slowly added to the 5-nitroisophthaloyl dichloride in DCM without pyridine at -40°C. These conditions enabled the reaction to be maximally controllable since the solubility of the nucleophile in DCM was poor. After warming to RT, two drops of pyridine were added which assisted with the solubility and catalysed the amide bond formation. The remaining portion of the aminobenzimidazole as the limiting reactant was then also added slowly at -40°C before warming the solution and heating at 40°C to enable adequate solubility of the starting materials. Once TLC indicated that the amine had been consumed, aniline in excess was added to form the second amide bond, and a major product spot on the TLC, which did not correspond to that of the diphenylbenzamide product, was observed. After extraction into EtOAc, and hot filtrations to remove the impurities, ^{1}H NMR analysis confirmed the successful synthesis of the unsymmetrical dibenzamide, **8** (*Figure 3.12*). As a result of the insolubility of compound **8** in many organic solvents, further purification by recrystallisation was only possible in hot pyridine, resulting in traces of pyridine in the ^{13}C NMR spectrum after drying under reduced pressure. The peaks were characterised from HSQC NMR analysis (*Figure 3.13*). As for the other dibenzamides, HRMS confirmed structural identity in addition to NMR techniques. However, compound **8** had the lowest HPLC purity of all of the compounds synthesised of 93.2%, with UV-active trace amounts of pyridine, and a melting point range of 3°C. The remaining dibenzamides were all found to be >95% pure with a ≤3 °C melting point range.

a)



b)

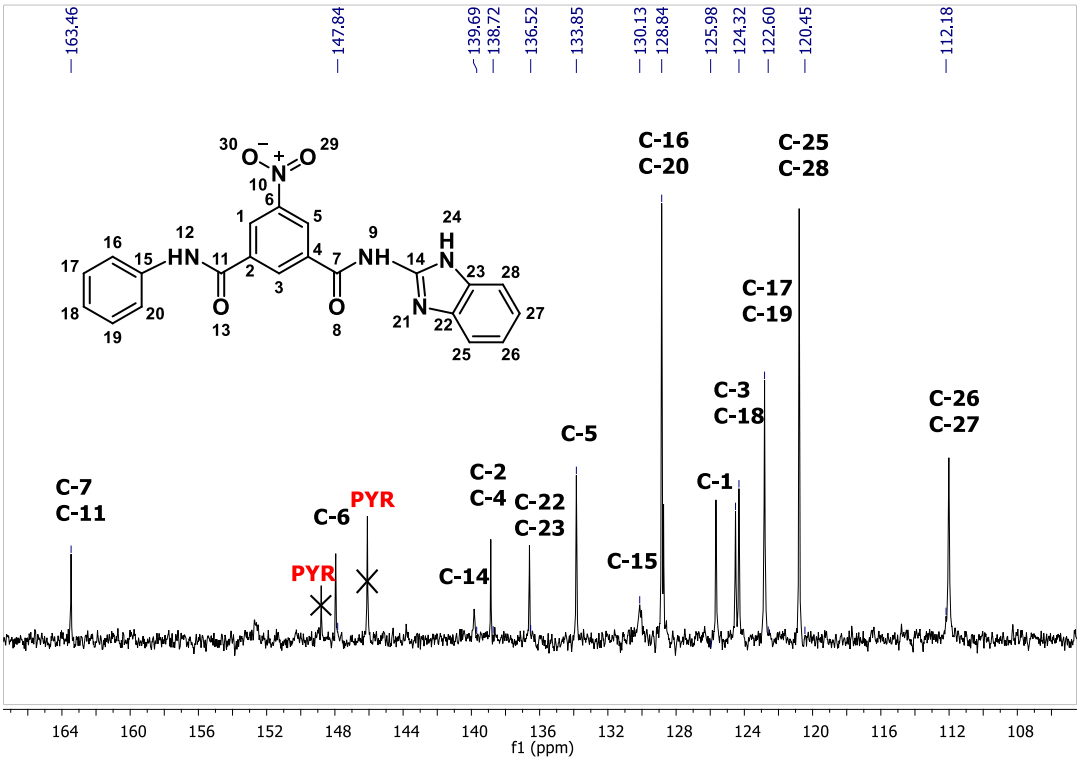


Figure 3.12 The assigned a) ^{1}H NMR (before recrystallisation in pyridine) and b) ^{13}C NMR spectra (after recrystallisation in pyridine) for **8**, confirming the successful synthesis. Recrystallisation resulted in trace amounts of pyridine observed in the spectra (^{1}H ppm 9.7, 8.8, 8.4; ^{13}C ppm 148.8, 146.1, 128.7).

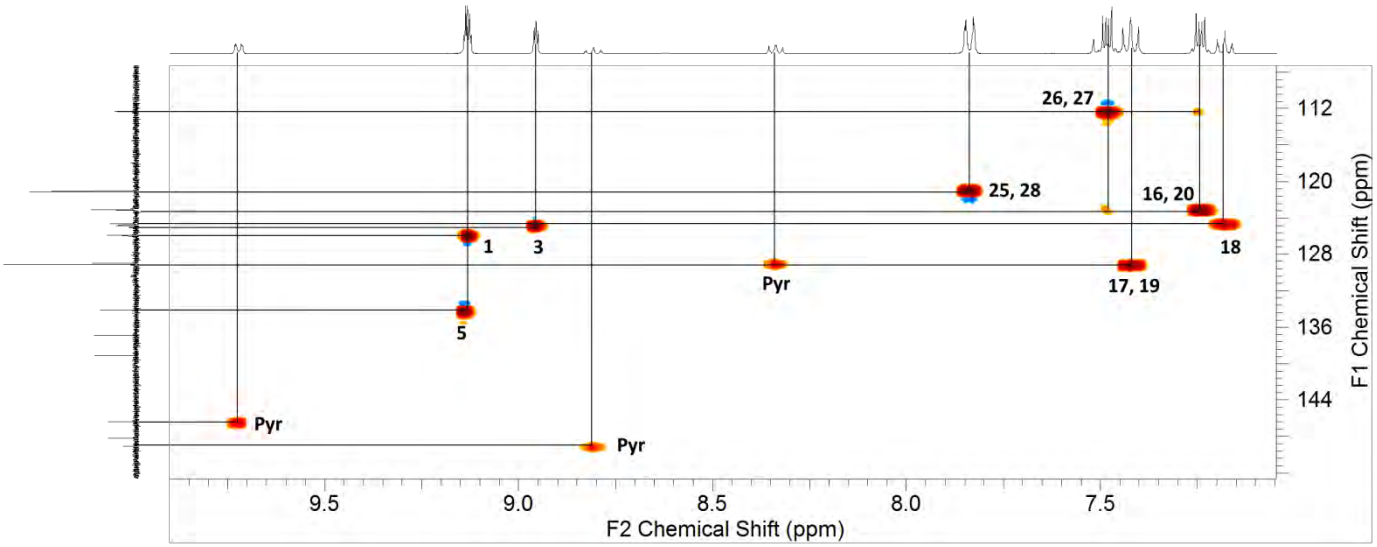
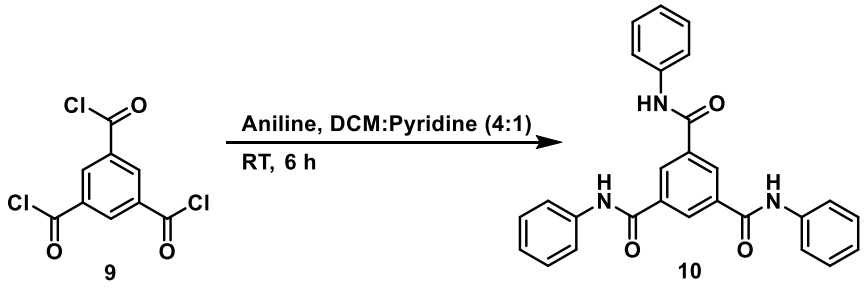


Figure 3.13 The assigned HSQC spectrum showing the correlation of ^{13}C and ^{1}H peaks for assignment of compound **8**.

3.2.1.3 Synthesis of the tribenzamide

Having synthesised a range of mono- and di-benzamides, access to a tri-benzamide was the next rational step in order to enrich the SAR analysis. Hence, symmetrical triphenylbenzamide **10** was synthesised which was easily achieved from the commercially available benzene-1,3,5-tricarbonyl trichloride with aniline and a 4:1 mixture of dry DCM and pyridine in 55% yield (*Scheme 3.14*).



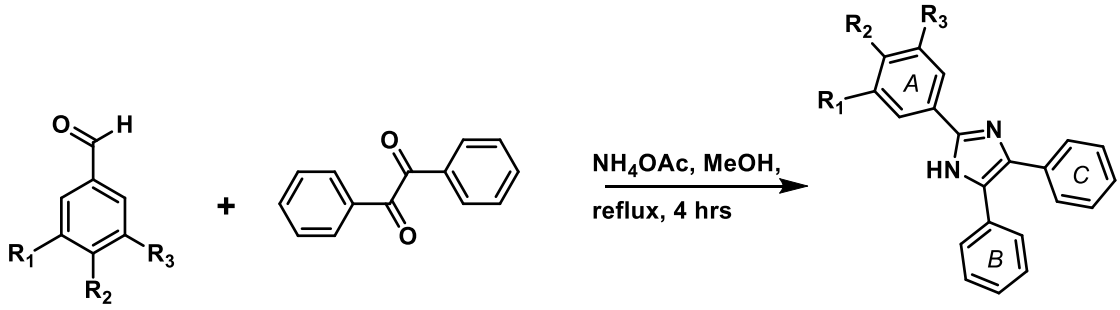
Scheme 3.14 Formation of tribenzamide **10**.

3.2.2 Scaffold 2: Triarylimidazoles

3.2.2.1 Synthesis of triarylimidazoles with unsubstituted B and C rings

The standard method available in the literature was used to prepare triarylimidazoles **15a-d** by refluxing the appropriate aldehyde, benzil and sodium acetate in MeOH, according to *Scheme 3.15*. Benzil, being unsubstituted in the phenyl rings, resulted in products containing unsubstituted B and C rings, while substitution could be readily introduced into the A-ring by appropriate choice of aldehyde.

High temperatures were preferred owing the release of water in the reaction mechanism, which is entropically favoured. The Gibbs free energy of reaction, which is related to the enthalpy (ΔH), temperature (T) and entropy (ΔS) by ΔG = ΔH - TΔS, is more negative with higher temperatures since ΔS is positive for entropically favoured reactions. The decrease in ΔG typically results in a higher shift in the reaction equilibrium to the product side.[273]



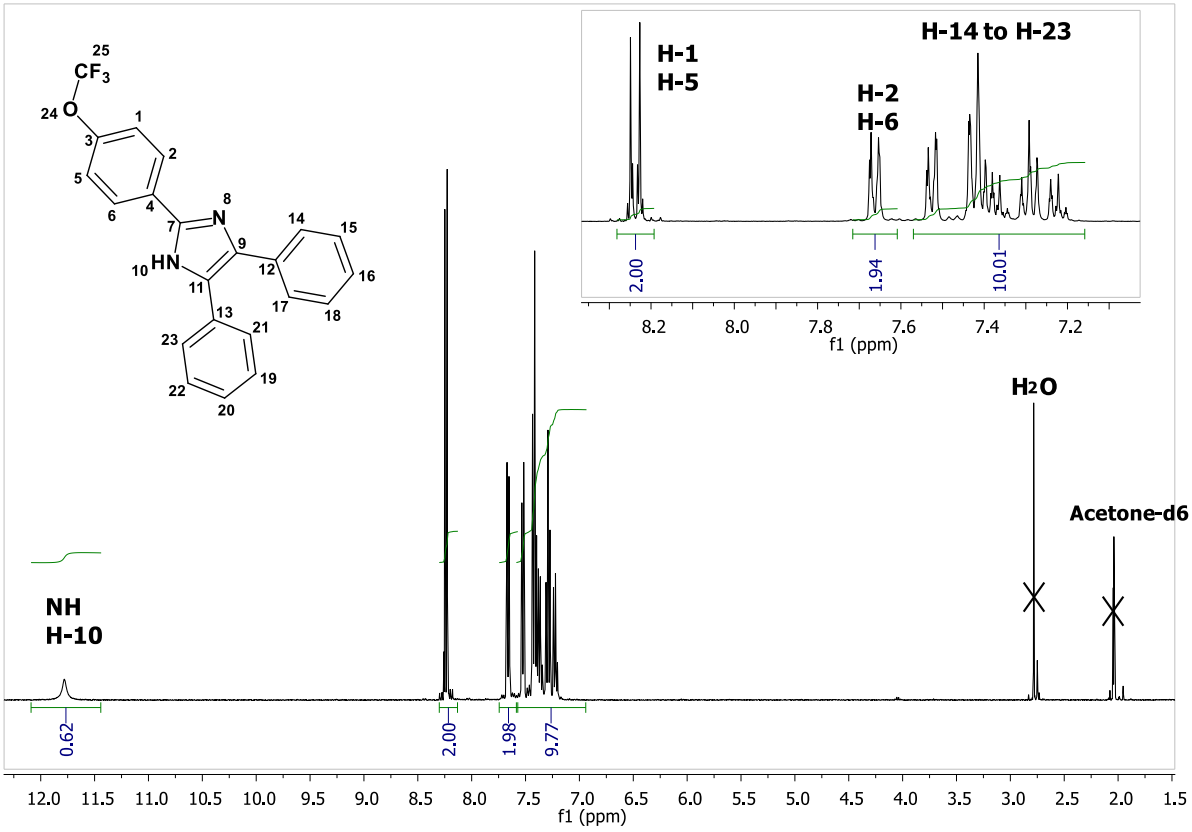
13a-d 14 15

15a: R_1 = OMe; R_2 = H; R_3 = OMe
15b: R_1 = OMe; R_2 = OMe; R_3 = OMe
15c: R_1 = OMe; R_2 = OH; R_3 = OMe
15d: R_1 = H; R_2 = OCF_3; R_3 = H

Scheme 3.15 Preparation of the triarylimidazole derivatives **15a-d**.

The optimized conditions presented by Azizi *et al*.[259] were followed, which gave good to excellent yields for **15a-d**. NMR data or melting points correlated with that of the literature for **15a**[274], **15b**[264] and **15c**[275] although **15d** was a new, previously unreported compound, with a similar reaction time to that of **15a-c**. There was no apparent trend observed between the electron donating ability of the aldehyde substituent and the reaction time, despite the expectation from a mechanistic standpoint that an electron-withdrawing substituent would increase the reaction rate, due to the greater electrophilicity of the

aldehyde carbonyl carbon. *Figure 3.14* shows confirmation of the synthesis of **15d**, via ^{1}H and ^{13}C NMR analysis in deuterated acetone (acetone-d_6) and DMSO-d_6, respectively. The downfield NH peak in the ^{1}H spectrum (~11.8 ppm) confirmed the formation of the imidazole ring. The ^{13}C peak corresponding to the trifluoromethoxy carbon was further downfield at ~129 ppm compared to a methoxy carbon (normally at 50-60 ppm)[276] owing to the strong electron-withdrawing ability of the three adjacent fluorine atoms, and was observed as a relaxed quartet (J = 144.9 Hz) due to a large germinal C-F coupling (I_F = 0.5). This quartet is indicated by the red asterisks in *Figure 3.14*.



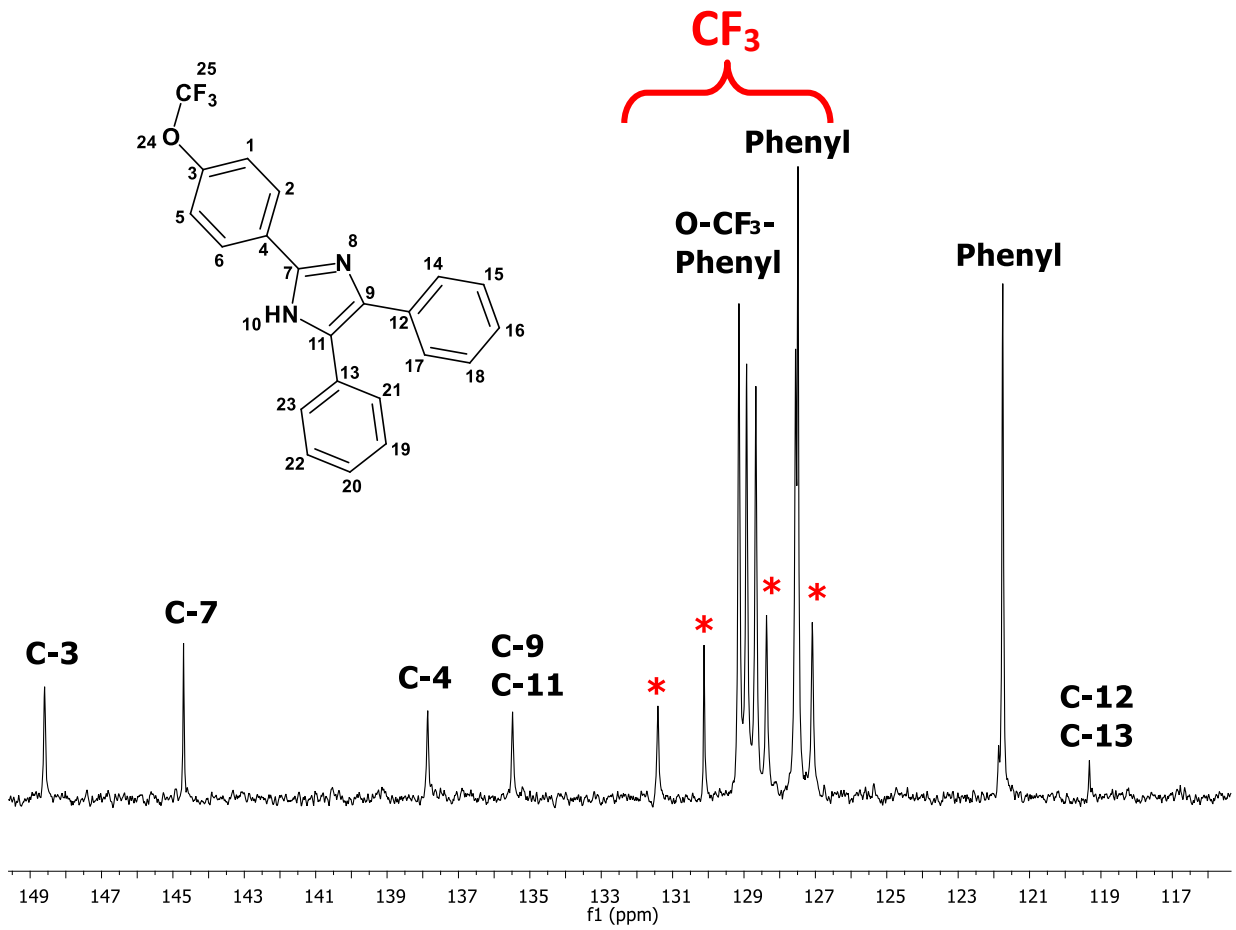
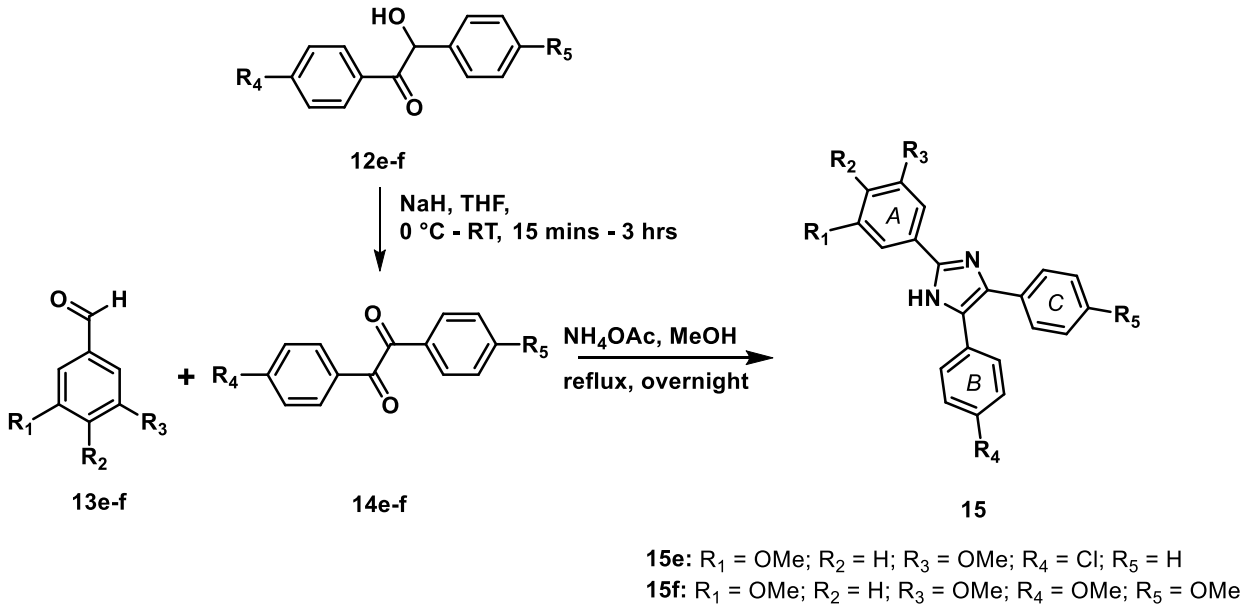


Figure 3.14 ^{1}H and ^{13}C NMR spectra of **15d** in DMSO-d_6.

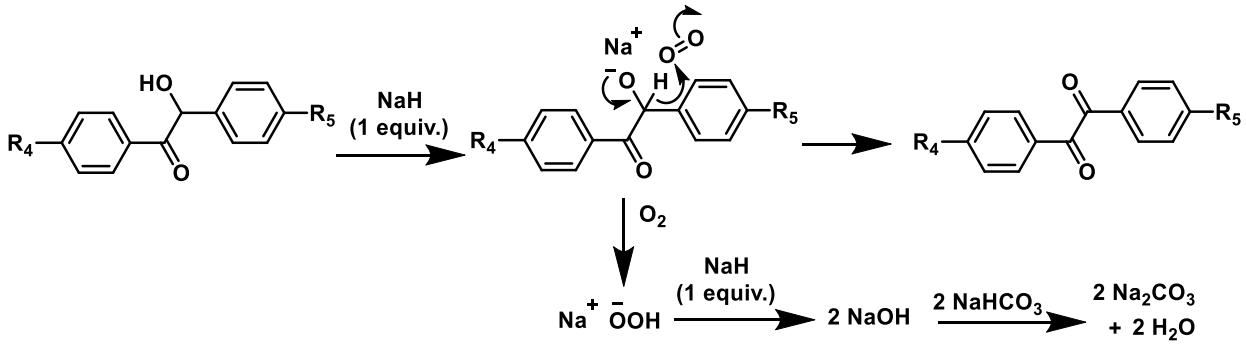
3.2.2.2 Synthesis of triarylimidazoles with substituted B and C rings

Target compounds **15e** and **15f** containing substituents in the B and C rings required 4-chlorobenzil and 4,4'-dimethoxybenzil as the starting materials respectively (*Scheme 3.16*). Since these compounds were not commercially available at the time of synthesis, they were synthesised from the corresponding benzoins, **12e** and **12f** via oxidation of the hydroxyl to the carbonyl with sodium hydride (NaH). Here, oxidation proceeds via the anion in a process driven by conjugating the two carbonyl groups into resonance with both aromatic rings.



Scheme 3.16 Preparation of the triarylimidazole derivatives **15e** and **15f**.

The proposed mechanism by Joo *et al*.[277] first involves the deprotonation of the hydroxyl group. The resultant alkoxide then undergoes an interesting hydride elimination with transfer to molecular oxygen, which is reduced to hydroperoxide. This then reacts with a second eq. of NaH to produce 2 eq. of sodium hydroxide (NaOH). The driving force for this transfer is undoubtedly that of achieving full conjugation between both phenyl rings and the two carbonyl groups. A work-up with sodium bicarbonate ($NaHCO_3$) helps to wash out the NaOH, allowing for negligible by-product formation (*Scheme 3.17*). The role of oxygen was investigated by carrying out the reaction under oxygen-free conditions, which resulted in a 33% decrease in product yield, confirming the involvement of oxygen in the mechanism.



Scheme 3.17 The proposed mechanism for the oxidation of benzoins to benzils using NaH.[277]

The two oxidation reactions of **12e** and **12f** progressed at significantly different rates. The formation of **14e** (R_4 = Cl, R_5 = H), was completed after only 15 mins of stirring from 0° C to RT. The rapid rate of oxidation for this chloro derivative agreed with the previously reported time of 6 mins at RT. This was in contrast to **14f** (R_4 = R_5 = OMe), the 4,4'-dimethoxy derivative, which required stirring at RT for 3 h in order to completely consume the benzoin starting material. A likely reason for this difference in rate is the electronic effect of the chloro (Cl) vs methoxy (OMe) substitiuent. The reported reaction time for benzoin (R_4 = R_5 = H) is 1.5 h, significantly longer than the chloro derivatives, which show times ranging from 6 to 15 mins. Since the order of electron donating ability is OMe > H > Cl and the corresponding reaction times were 3 h > 1.5 h > 15 mins, the electron density within the phenyl ring is likely to have an important effect on the degree of stabilisation of the developing carbonyl group. Here, presumably mesomeric electron release (OMe) or inductive withdrawal (Cl) destabilises or stabilises respectively, the developing carbanionic character of the benzylic position in the transition state (*Figure 3.15*). Although the oxidation of 4,4'-dimethoxybenzoin with NaH has not been reported, that for the 3,3'-dimethoxy derivative *meta* to the C=O and C-OH, has been published. Interestingly, this reaction has a reported time of 9 mins, remarkably shorter than both the *p*-dimethoxy and the unsubstituted benzoin. This is in agreement with the proposed hypothesis since the *m*-methoxy cannot conjugatively release electron density via resonance towards the carbonyl, although it does have an inductive withdrawing effect which may assist with *α*-hydride elimination. Alternatively, the negative charge adjacent to the hydroxyl carbon in the *p*-methoxy derivative may decrease the rate of *α*-hydride elimination, since negative charge must flow from the sodium alkoxide oxygen to the adjacent carbon in order for this

elimination to take place. The negative charge on the phenyl ring cannot assist with this elimination since it is part of the π-orbital framework and the *α*-hydride leaves with electron density through the sigma orbital framework. *Figure 3.15* summarises these ideas for a *p*-methoxy group.

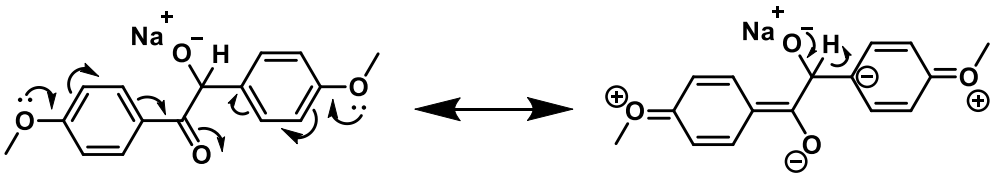


Figure 3.15 Resonance structures for 4,4'-dimethoxybenzoin.

Formation of the required benzils **14e** and **14f** was confirmed by ^{1}H NMR analysis, whose data corresponded to that of the literature,[278] and the spectra of which are shown in *Figures 3.16* and *3.17*, respectively. In both cases resonances in the aromatic region shifted downfield upon oxidation owing to the electron-withdrawing effect of the oxidised carbonyl. For **12e**, coupling (*J* = 6.2 Hz) was observed between the hydroxyl and benzylic protons. This is due to the hydrogen bond accepting capability of the solvent used, acetone-d_6, which decreases the rate of hydroxyl proton exchange and results in doublets for each of these two resonances. However, this coupling was not observed for **12f**, in $CDCl_3$, since intermolecular exchange of the hydroxyl protons, leading to broadening of the signals, occurs in this solvent.[279] The resonances corresponding to the hydroxyl and benzylic protons disappeared upon oxidation to **14e** and **14f**. The oxidation to **14f** was particularly interesting as the molecule becomes symmetrised upon forming two carbonyls. This is very evident in the ^{1}H NMR spectra (*Figure 3.17*) which shows the two methoxy peaks, initially in different chemical environments converging to one peak once they are chemically and magnetically equivalent.

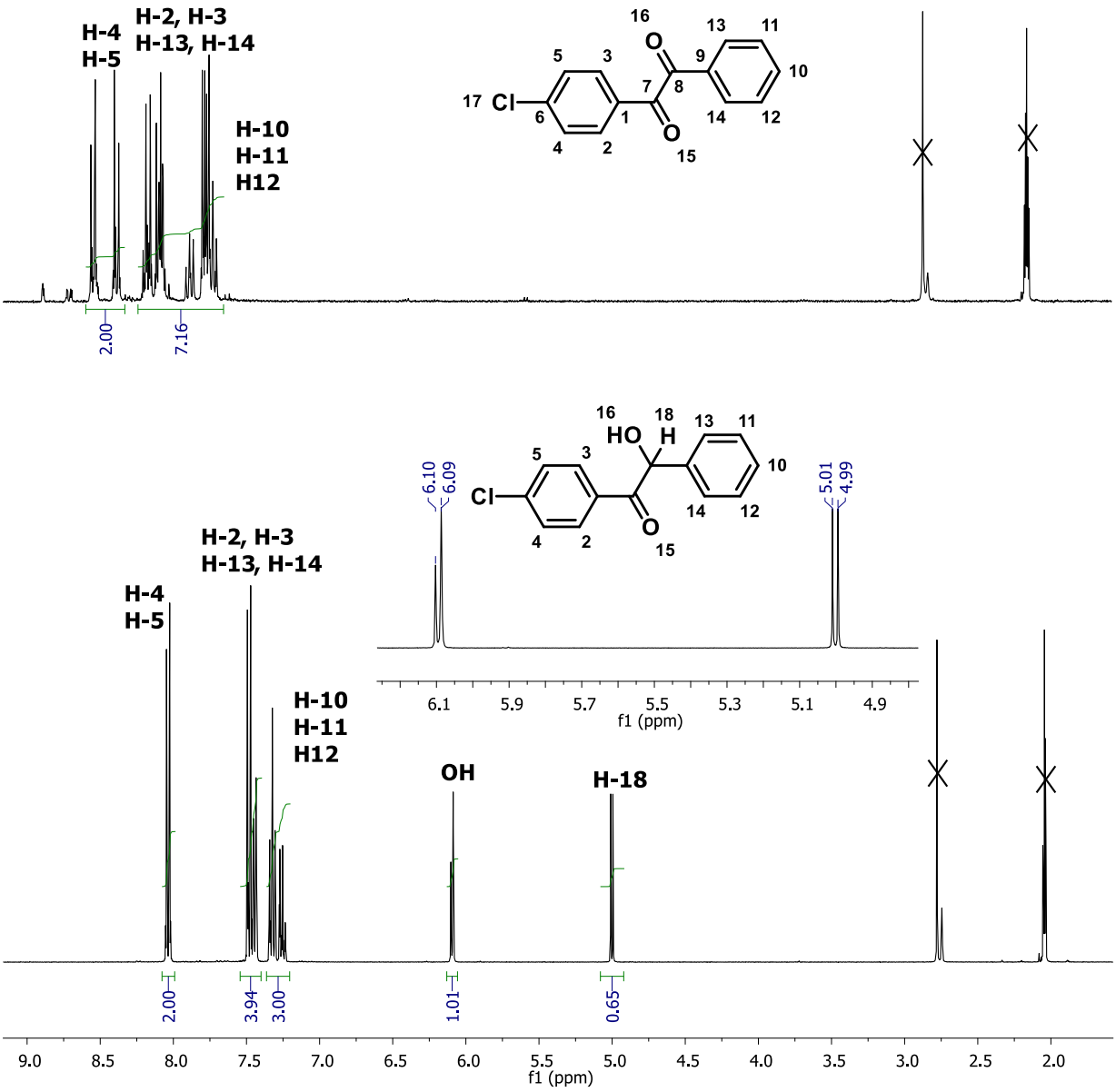


Figure 3.16 The ^{1}H NMR spectra of 4-chlorobenzoin (**12e**; bottom) and 4-chlorobenzil (**14e**, top) in acetone-d_6.

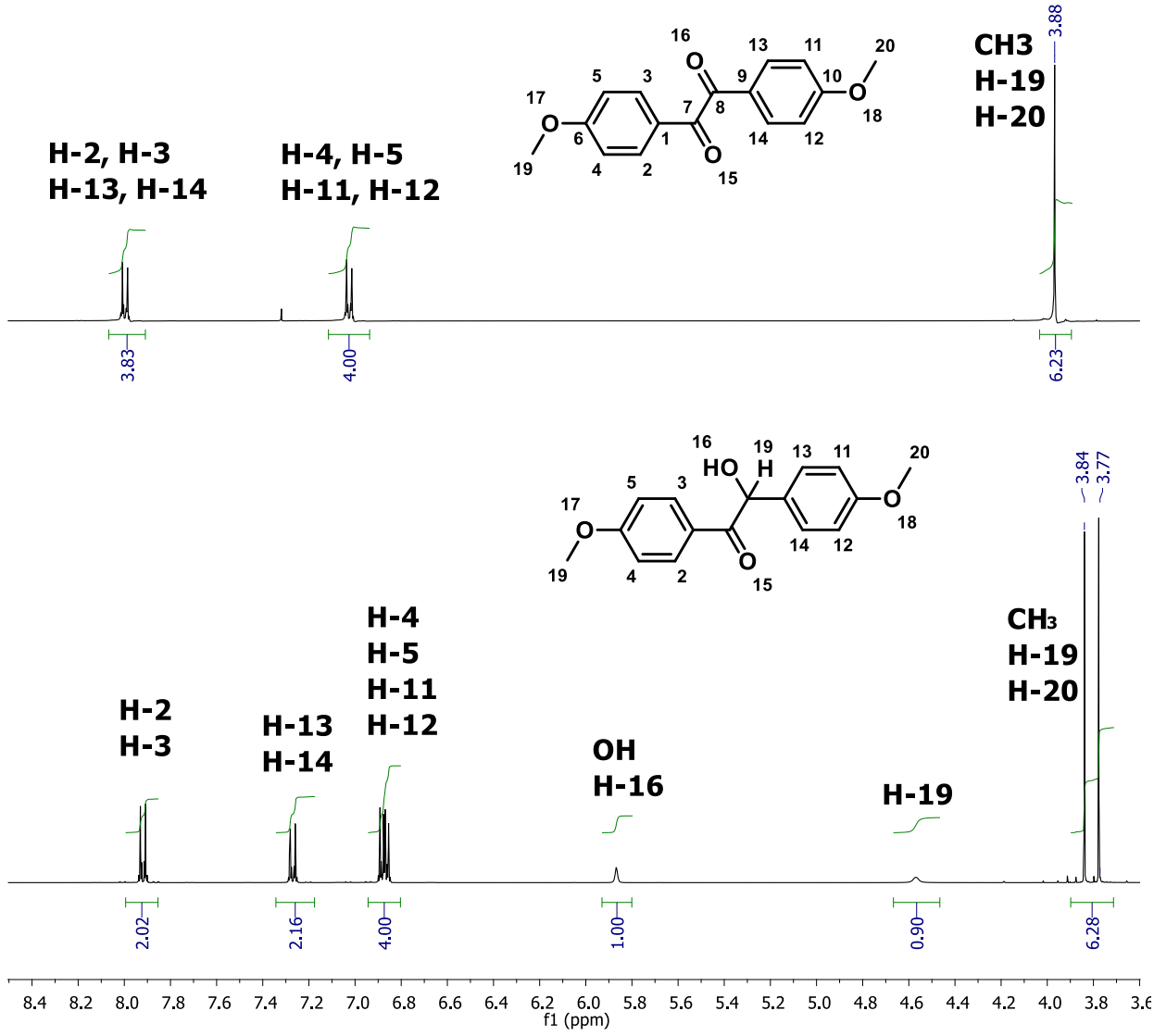
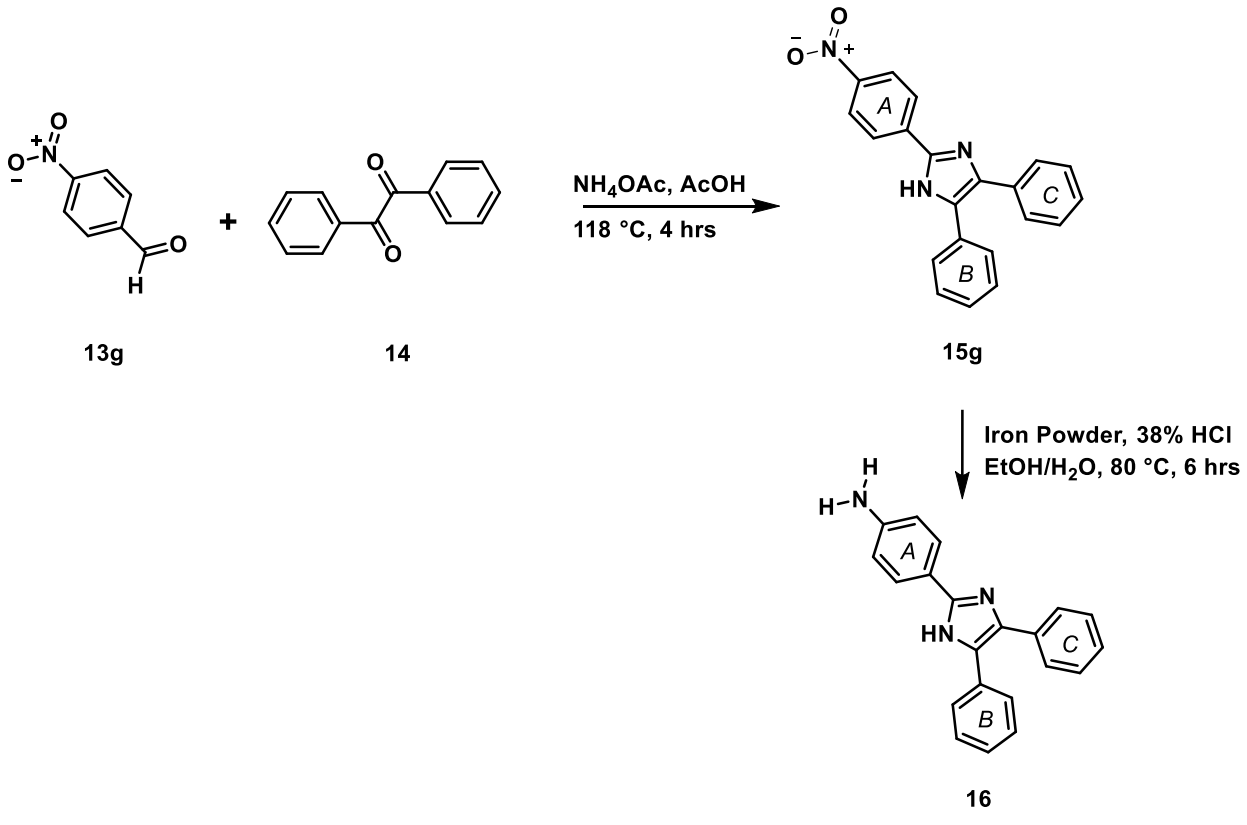


Figure 3.17 The ^{1}H NMR spectra of 4,4'-dimethoxybenzoin (**12f**; bottom) and 4,4'-dimethoxybenzil (**14f**; top) in $CDCl_3$.

Following synthesis of the benzils (**14e** and **14f**), the corresponding triarylimidazoles (**15e** and **15f**) were formed in good yields using the procedure previously described for **15a-d**.

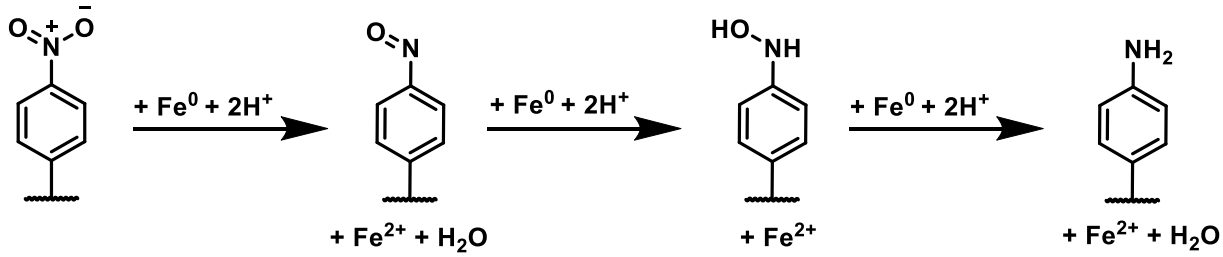
3.2.2.3 Synthesis of nitro- and amine-substituted triarylimidazoles

Attempts to synthesise target compound **16** containing a para-amino group in the A-ring from p-aminobenzaldehyde were unsuccessful. The reason for this was not further investigated since multiple spots on the TLC were observed in which the reaction formed a dark, gummy mixture that could not be extracted into an organic solvent. It is likely that the aldehyde preferentially self-condenses under these conditions to give oligomeric products. An alternative route was followed whereby the nitro derivative (**15g**) was formed and then reduced to the amine using iron powder and a catalytic amount of HCl. The conditions reported by Wan *et al.*[260] were followed for both **15g** and **16**, which involved a 4 hr glacial acetic acid reflux at 118 °C for **15g** and then a 6 hr reflux at 80 °C reflux with Fe powder in an EtOH/water (2:1) mixture for **16** *(Scheme 3.18)*.



Scheme 3.18 Preparation of the triarylimidazole derivatives **15g** and **16**.

The formation of **15g** in acetic acid occurred at a similar rate to that of **15a-d** in MeOH. Upon neutralisation of the acid with ammonia during the workup, the insoluble product, **15g**, could be extracted by filtration and then purified by recrystallisation resulting in an excellent yield of 93%. In the subsequent step, the nitro group was reduced with zero-valent elemental iron and HCl according to the stoichiometric reactions shown in *Scheme 3.19*.[280,281]



Scheme 3.19 The reduction of an aromatic nitro functionality to a phenyl amine using elemental iron and a catalytic amount of HCl.

The formation of the amine was supported in the ^{1}H NMR spectrum by the appearance of two amine protons at ~5.3 ppm as well as an upfield shift of the four protons on ring A due to the electron-releasing influence of the newly introduced amino group. *Figure 3.18* shows the overlaid spectra for the nitro (**15g**) and amine (**16**) derivatives respectively.

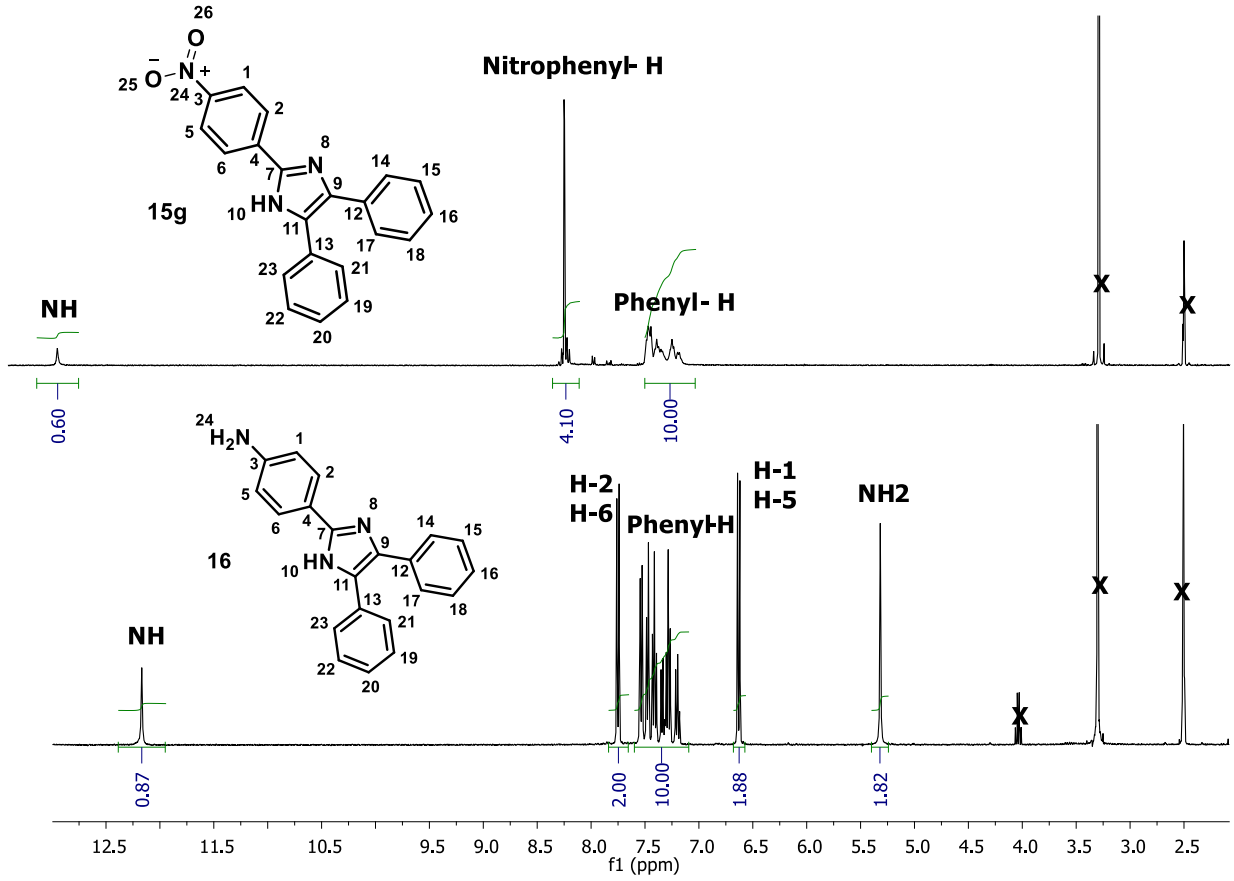


Figure 3.18 Comparison of the spectra in DMSO-d_6 for the nitro derivative (**15g**) and the amine derivative (**16**).

All the triarylimidazoles (**15a-g** and **16**) were characterised by both ^{1}H and ^{13}C NMR analysis as well as HRMS, which showed masses for the observed $[M+H]^+$ ions within 5 ppm of the theoretical values. Furthermore, all compounds were found to be adequately pure with a melting point range of ≤2 °C and >95% purity as determined by HPLC analysis.

3.3 Summary and conclusions

The formation of the benzamide derivatives was carried out from the appropriate acid chloride, which was either purchased or freshly prepared, and either aniline, an aminopyridine, or 2-aminobenzimidazole or 4-[(N-Boc)aminomethyl]aniline. In the last, the desired product, derivative **7**, was achieved via Boc-deprotection, formylation to the imine and reductive amination. All dibenzamides were symmetrical except for derivative **8**, which

was formed via the addition of the less reactive amine, the 2-aminobenzimidazole first, before the addition of aniline to the acid chloride. This involved lower temperatures and a catalytic amount of pyridine to carefully control the reaction, in contrast to the symmetrical derivatives which were conducted in neat pyridine for optimal reaction rates and yields. The addition of Et_3N to the amide coupling reaction was not necessary in the case of the phenylbenzamide derivatives but aided in solubility of the starting materials in the case of the pyridylderivatives and also neutralised the formed HCl to ensure isolation of the unprotonated product.

Synthesis of the triarylimidazoles via multiple condensation steps with benzil, an aldehyde and ammonium acetate was successfully performed using standard procedures, which required refluxing for optimal reaction times and product yields. The elimination of water in the reaction mechanism led to higher temperatures being favoured according to the Gibbs free energy equation since an increase in temperature resulted in a larger entropy term. Two of the benzils, **14e** and **14f**, were oxidised from the corresponding benzoins with the strong, hard base, sodium hydride. The reactions required diverse reaction times which agreed with the literature trends. Electron-withdrawing substituents increased the rate while electron releasing *para* substituents significantly decreased the rate of oxidation. The reaction with 4-aminobenzaldehyde was unsuccessful but the target compound could be synthesised via the nitro derivative, **15g**, which was then reduced to yield the required amino derivative, **16**.

A total of 24 mono- or di-benzamides and 8 triarylimidazoles were synthesised in moderate to excellent yields using inexpensive starting materials. Most synthetic routes required no more than two steps and minimal time for isolation and purification. All the characterisation data confirmed compound identity and high purity. The synthesised derivatives comprised a variety of electron-withdrawing and donating groups as well as hydrophobic and hydrophilic moieties. In addition, they could all be dissolved in DMSO, making them suitable for testing in the required assays for SAR analysis.

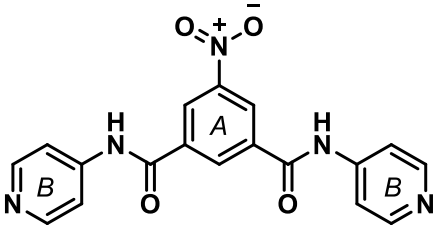
CHAPTER FOUR

BENZAMIDES: STRUCTURE-ACTIVITY RELATIONSHIPS FOR β-HAEMATIN INHIBITION AND PARASITE ACTIVITY

4.1 Introduction

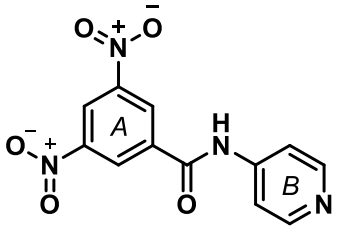
One of the major challenges of drug development is identifying appropriate research routes after HTS, particularly when it comes to selecting potential leads. In this case, HTS revealed the pyridyl-dibenzamide, 5-nitro-*N*,*N*'-di(pyridin-4-yl)benzene-1,3-dicarboxamide (*Figure 4.1a*), with nearly a ten times lower IC_{50} for βH inhibition than CQ and with a promising parasite activity of 0.19 μM against the D6 strain, the third most potent compound of all 171 parasite-active hits. Furthermore, a mono-pyridylbenzamide, 3,5-dinitro-*N*-(pyridin-4-yl)benzamide (*Figure 4.1b*) was also shown to have excellent potency against βH formation but a poorer potency against the parasite. A literature search revealed that the dibenzamide chemotype had not been previously published as an antimalarial, although other mono-benzamides had been reported to prevent parasite growth, supposedly via kinase inhibition.[213] Besides potent activity, the benzamides were attractive from a synthetic point of view, owing to the ease with which derivatives could be synthesised. Indeed, almost all the required derivatives were efficiently synthesised in short sequences from cost-effective reagents. In addition, the structural difference of the benzamides from many of the haemozoin-inhibiting antimalarials that contain fused ring systems such as the quinolines and neocryptolepines, made them attractive for probing mechanisms of action.[107-109,282] In the discussion which follows, ring A is defined as the aromatic substituent that is attached to the carbonyl of the amide and ring B to the amine of the amide.

a)



VU0001281, IC_{50} of 0.19 µM

b)



VU0358764, IC_{50} of 4.76 µM

Figure 4.1 Examples of parasite-active hits found in the HTS; a) The di-pyridylbenzamide, 5-nitro-*N*,*N*'-di(pyridin-4-yl)benzene-1,3-dicarboxamide and b) the mono-pyridyl-benzamide, 3,5-dinitro-*N*-(pyridin-4-yl)benzamide**.**

4.1.1 Specific goals

The specific goals for Chapter 4 were to:

- Test the purchased or synthesised simple benzamide derivatives in the NP-40 assay for βH inhibition to determine the minimum structural requirements for activity.
- Employ the synthesised analogues possessing substituents with varying electronic, lipophilic and spatial properties for SAR analysis by evaluating the βH inhibition activity of these analogues in the NP-40 assay and investigating mechanistic aspects of the inhibition process.
- Determine the inhibition of parasite growth activity using the LDH malaria parasite growth assay.
- Attempt to validate the biological target within parasitised red blood cells.
- Perform other relevant experiments to determine cellular accumulation or strength of haem association.

4.2 Results and discussion

4.2.1 β-haematin inhibitory activity of the benzamides

4.2.1.1 βH inhibition assay in 96-well plates

The NP40 assay for βH inhibitors employed in the HTS at Vanderbilt University (VU) was modified for use in 96-well plates at the University of Cape Town (UCT). At VU, compounds were added to a 384-well plate automatically using acoustic delivery from source plates which contained the HTS library compounds. At UCT, however, compounds and reagents had to be added manually. As a result, the use of 384-well plates became problematic, since they required very low µL volume additions. Instead, compounds were added using single channel micro pipettes. For concentration dose response experiments, this was initially carried out by preparing several stock solutions of the test compound and pipetting the relevant quantity into the well for the desired concentration. However, efficiency and accuracy was greatly improved after adopting a serial-dilution technique to achieve the required test sample concentration. Reagents were added using multichannel pipettes and incubated in either a closed water bath or a thermally controlled incubator without shaking. The IC_{50}s were consistently slightly lower than the 384-well experiment, but relatively comparable. Reproducible values were obtained (*Figure 4.2*).

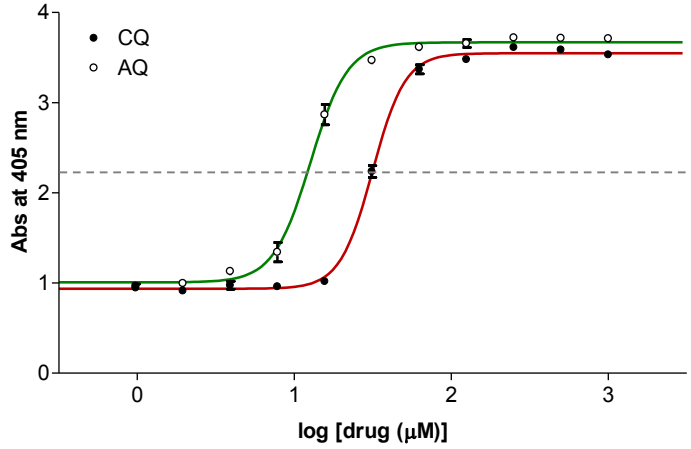


Figure 4.2 Concentration dose response curves generated from four replicates using the serial dilution technique for the NP40 assay in 96-well plates for CQ and AQ, with IC_{50}s of 31.5 ± 0.5 µM and 12.7 ± 0.3 µM respectively.

4.2.1.2 Evaluation of βH inhibition for the purchased benzamides

N-phenylbenzamide, the core structure of the hit compounds **VU0001281** and **VU0358764** (*Figure 4.1a*, *b*), as well as simple derivatives were purchased in order to ascertain the smallest fragment for which βH inhibition activity is obtained by varying substituents on the phenyl ring. Selection of the compounds was guided by the Topliss decision tree, which aids in deciding which substituents to employ for improved activity.[256] The idea behind the diagram in *Figure 4.3* is to start from a derivative with an unsubstituted ring and to evaluate the activity relative to the hit compound, which in this case contained 3,5-dinitro substituents. The branch corresponding to either less (L), equally (E) or more (M) active than the previous compound is then followed and the next substituent is evaluated. Following the guidelines for substituents within the relevant branch of the Topliss decision tree should theoretically lead rapidly to the substituents which are electronically, hydrophobically and sterically favoured for activity. In this study, the branches of the tree were not followed sequentially in the interest of time, but rather, compounds were purchased in order to cover a range of substituents throughout the tree so that SARs could be constructed using diverse derivatives. This is because the substituents selected using the Topliss scheme cover a wide range of electronic, hydrophobic and steric properties. In some cases, compounds were not commercially available, but ten were obtained (**P1** - **P10**) for βH activity measurement, as shown in *Figure 4.3* and *Table 4.1*.

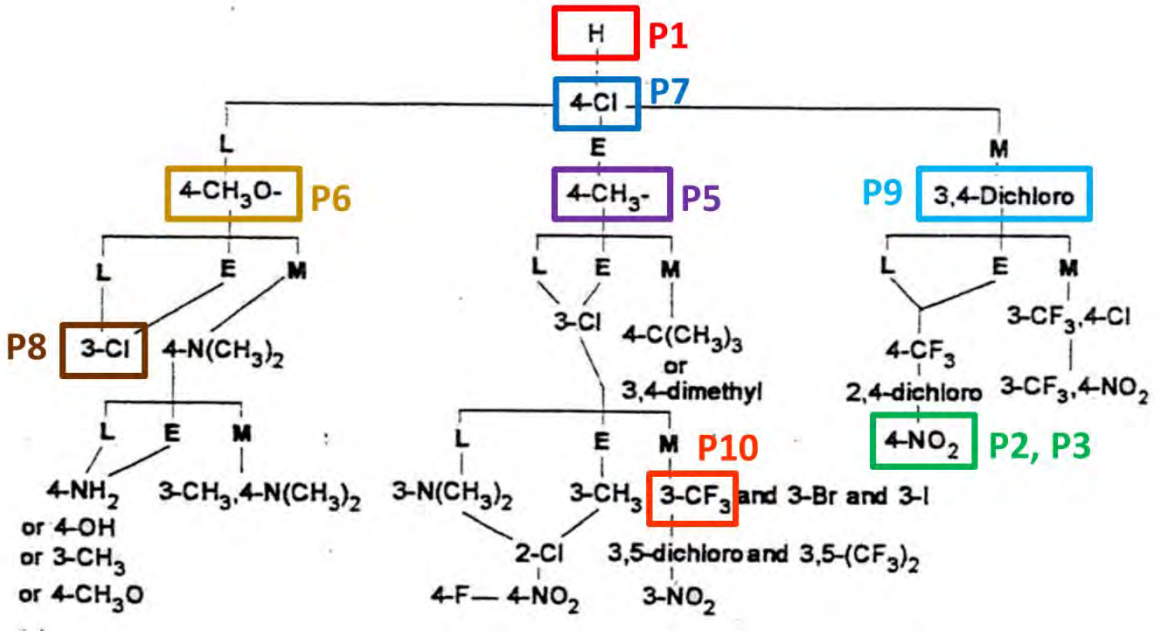
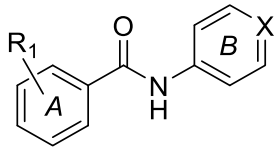


Figure 4.3 The Topliss decision tree showing aromatic substituents of purchased compounds **P1** - **P10** to evaluate SARs (L = less active, E = equally active and M = more active).[256] Numbered boxes indicate substituents of compounds purchased. Copyright © 2001, Oxford University Press.[283]

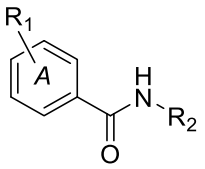
Table 4.1 Purchased benzamides for core structure analysis. All the compounds were inactive and benzamides **P7**-**P10** formed a pink precipitate at the indicated concentration range when added to a solution of haem and NP40 at pH 4.8.

	**Code**	**Concentration range at which precipitate forms (μM)**
P1: R_1 = H; X = N **P2**: R_1 = *4*-NO_2; X = CH **P3**: R_1 = *4*-NO_2; X = N **P4**: R_1 = *2,4*-di-NO_2; X = N **P5**: R_1 = *4*-CH_3; X = N **P6**: R_1 = *4*-OCH_3; X = N **P7**: R_1 = *4*-Cl; X = N **P8**: R_1 = *3*-Cl; X = N **P9**: R_1 = *3,4*-di-Cl; X = N **P10**: R_1 = *3*-CF_3; X = N	**P1 – P6**	N/A (Inactive < 5500 μM)
	P7	4500-5500
	P8	1300-2300
	P9	700-1300
	P10	1300-2300

Unexpectedly, the purchased compounds **P1** – **P10** did not display activity against βH formation (IC_{50} > 5500 μM). Compounds **P1** - **P6** contained either an unsubstituted phenyl moiety as ring A and a pyridyl as ring B, or a substituent in the *para* position of the phenyl ring and a phenyl or pyridyl as ring B (*Table 4.1*). These compounds showed no specific interaction with haematin. However, purchased compounds **P7** – **P10**, all possessing a halogen substituent in the *meta* and *para* positions, demonstrated an interesting phenomenon whereby, upon addition of haem, a pink precipitate rapidly formed, indicative of low spin Fe(III)PPIX complex formation in a similar manner to that of pyridine coordination to unconverted Fe(III)PPIX in the βH assay. Coordination of these compounds to the central Fe atom of haem was believed to be via the nitrogen of their pyridyl substituent, which has added nucleophilicity over that of pyridine due to resonance from the amide N atom to the 4-pyridyl N atom. The precipitation of this complex was likely caused by the lipophilic influence of the halogen(s) with **P9,** containing two chloro substituents, showing the effect at the lowest compound concentration and **P7,** with only one 4-chloro atom, showing the effect at five times higher concentration. In summary, this precipitation occurred at a lower concentration for *meta* substituents as opposed to *para* and for two substituents as opposed to one.

4.2.1.3 Evaluation of βH inhibition for the synthesised monobenzamides

The results for the purchased compounds showed that the Topliss decision tree was not a suitable method for selecting derivatives for SAR analysis. This was because the activity of the monobenzamides could not be “switched on” by the halogen substituents or by *para* substituents. In contrast, the hit compound with two *meta*-nitro substituents showed excellent βH inhibition activity (*Figure 4.1b*). This compound was resynthesised so that the results of the HTS could be confirmed (compound **2c**). Since both compounds **P3** (4-nitro substituent) and **P4** (*2,4*-dinitro substituents) were not βH inhibitors, the influence of the *para* nitro group was probed by synthesising and evaluating a derivative with only an *ortho* nitro substituent (**2a**). Additionally, the type of ring B (defined here as R_2) was varied in order to probe the effect of the pyridyl nitrogen. *Figure 4.4* shows the five synthesised monobenzamides.



2

2a: R_1 = *2*-NO_2; R_2 = phenyl
2b: R_1 = 3,5-di-NO_2; R_2 = phenyl
2c: R_1 = *3,5*-di-NO_2; R_2 = *4*-pyridyl
2d: R_1 = *3,5*-di-NO_2; R_2 = *3*-pyridyl
2e: R_1 = *3,5*-di-NO_2; R_2 = *2*-pyridyl

Figure 4.4 The synthesised mono-benzamide derivatives.

Activities for the monobenzamides are shown in *Table 4.2.* Compound **2a** did not show activity against βH formation up to 1000 μM, as compared to the CQ standard of 30 μM. This result led to the conclusion that mono phenyl benzamides with 2-nitro, 4-nitro or 2,6-dinitro substituents were incapable of inhibiting βH formation. However, the phenylbenzamide, **2b**, with 3,5-dinitro substituents, showed a good activity of 22 ± 1 μM, which was slightly better than the CQ standard and about half as active as AQ. Compounds **2c** and **2d**, *N*-pyridyl analogues of **2b**, showed even better activity with IC_{50} values of 10.9 ± 0.3 μM and 13 ± 1 μM respectively. The dramatic increase in activity for **2b**, **2c** and **2d** compared with **2a** showed that the *meta* nitro substituents were required for βH activity in which variation of the R_2 ring had less influence. This proved that the nitrogen in the pyridyl ring was not required for activity and therefore also demonstrated that it was unlikely that this series inhibited βH formation via coordination of the pyridyl nitrogen to the haem iron centre. Compound **2e**, containing a 2-pyridyl at R_2 also displayed some activity (IC_{50} of 134 ± 8 μM), although it was 10 fold weaker than the 3-pyridyl and 4-pyridyl analogues. Since the derivative with no nitrogen atom in the R_2 ring (**2b**) was active, the decrease in activity of **2e** was presumably not due to the loss of the nitrogen in the 3 or 4-position, but rather because of the added nitrogen in the 2-position of the ring. Estimation of the pK_a values of **2c-e** using MarvinSketch software[269] showed a decrease in basicity of the nitrogen in 4-pyridyl, 3-pyridyl and 2-pyridyl groups with predicted pK_as of 5.62, 4.37 and 2.69 respectively. The relative basicity of the derivative influences the ratio of protonated to unprotonated species present in the assay well or in the DV. At pH 4.8, using the predicted pK_as, **2c** was 87% protonated while that of **2d** was only 27%. Owing to the much weaker basicity of **2e**, more than 99% was neutral at the pH of the DV. This weaker basicity is partly

due to the inductively withdrawing effect of the *ortho* amide substituent, which has been shown to have a larger inductive effect than a resonance releasing one, for the 2-position.[284] As a result, the conjugate acid of **2e** is destabilised. This is in contrast to the 4-position, where the resonance releasing ability of the amide is much greater than its inductive withdrawing ability, leading to increased 4-pyridyl basicity or greater conjugate acid stability. Another reason for the weaker basicity for the *ortho* substituted pyridine is the steric hindrance factor which destabilises protonation on the *ortho*-nitrogen. However, the relative basicity did not appear to be the sole reason for activity differences, since both **2c** and **2d** displayed equally potent activity. Furthermore, basic groups are not required for activity since **2b**, the phenyl derivative, was also active.

Table 4.2 βH inhibition activity for the synthesised monobenzamides

Compound	βH inhibition IC_{50} (μM)
2a	>1000
2b	22 ± 1
2c	10.9 ± 0.3
2d	13 ± 1
2e	134 ± 8
CQ	31.5 ± 0.5

4.2.1.4 Factors influencing βH inhibition of mono-benzamides.

The exact mechanism by which βH inhibitors block crystal growth is still under debate. Over the past two decades it has been proposed that inhibition by quinoline antimalarials is via direct drug complexation to haematin in solution,[139,285,141] by binding to the fastest growing faces of the βH crystal,[154,207,136] or via a drug-haem complex capping the haemozoin crystal to block sites of further crystal growth.[137] Nevertheless, all the proposed mechanisms involve an interaction between the drug compound and Fe(III)PPIX. Recently, de Villiers *et al.*[158] have reported crystal structures of QD-Fe(III)PPIX and QN-Fe(III)PPIX, which demonstrate that three key interactions are involved in haem binding by these drugs, namely coordination, hydrogen bonding and π-π stacking. For QD and QN, hydrogen bonding occurs between the protonated quinuclidine nitrogen and the propionate group of haem. Another study showed that dihydroxyxanthones hydrogen bond via their hydroxyl groups, as well as π-π stack with Fe(III)PPIX.[286] Most of the benzamides possess features

which are capable of binding to Fe(III)PPIX via all three of these interactions, specifically, a pyridyl N for coordination, an amide for hydrogen bonding and aromatic rings for π-π stacking. However, **2b** does not possess a pyridyl moiety for coordination to the Fe(III)PPIX centre and yet retains a low IC_{50}, ruling out coordination as a key interaction for βH activity. Furthermore, the lack of hydrogen bonding substituents in **2b-e** suggests that this interaction can only occur via the amide. Since many of the compounds with amides, including all the purchased benzamides, were inactive, an amide hydrogen bonding interaction was considered very unlikely to be the crucial determinant of βH inhibition activity. By a logical process of elimination, π-π stacking was left as the major interaction leading to inhibitory activity. Although the exact requirements for π-π stacking are still not clear, it was hypothesised that having larger, aromatic, planar molecular surfaces would increase the strength of the interaction with either haematin or haemozoin. The ability of the molecule to lie flat would therefore be expected to influence the βH inhibition activity. This hypothesis was considered to be the only feasible explanation for the observed SARs.

The π-π stacking mechanism was consistent with the structure-activity data for pyridyl derivatives **P4** with *2*,*4*-dinitro groups (inactive) vs **2c** with *3*,*3*-dinitro substituents (active) and phenyl derivatives, **2a** with *2*-nitro groups (inactive) vs **2b** with *3*,*3*-dinitro substituents (active). *In silico* internal energy calculations carried out using Gaussian software[287] demonstrated that the barrier to rotation around the bond linking the carbonyl to the phenyl for **P4** (*Figure 4.5a*) possesses minimum energy conformers with torsion angles about the C=C-C=O bond of 70° and -110°. For the molecule to be planar, the torsional angle would be 0° or ±180°. The steric strain energies of these flat conformers are far greater than the thermally accessible energy 3kT indicted in F*igure 4.5a*. If the molecule is only able to bind to Fe(III)PPIX when in the planar conformation, this hypothesis would explain the lack of βH inhibition activity for **P4**. This hypothesis is also consistent with the positive activity data for **2c** which is able exist in a planar conformation (±180°) below 3kT (*Figure 4.5b*) owing to the low steric strain caused by *meta* as opposed to *ortho*-nitro substituents. It may also explain why **2e** with a 2-pyridyl group had lower activity than the 3-pyridyl and 4-pyridyl analogues due to the close proximity of the negative *ortho* nitrogen lone pair to the electron cloud on the carbonyl oxygen, which is likely to affect the conformational preference of **2e**.

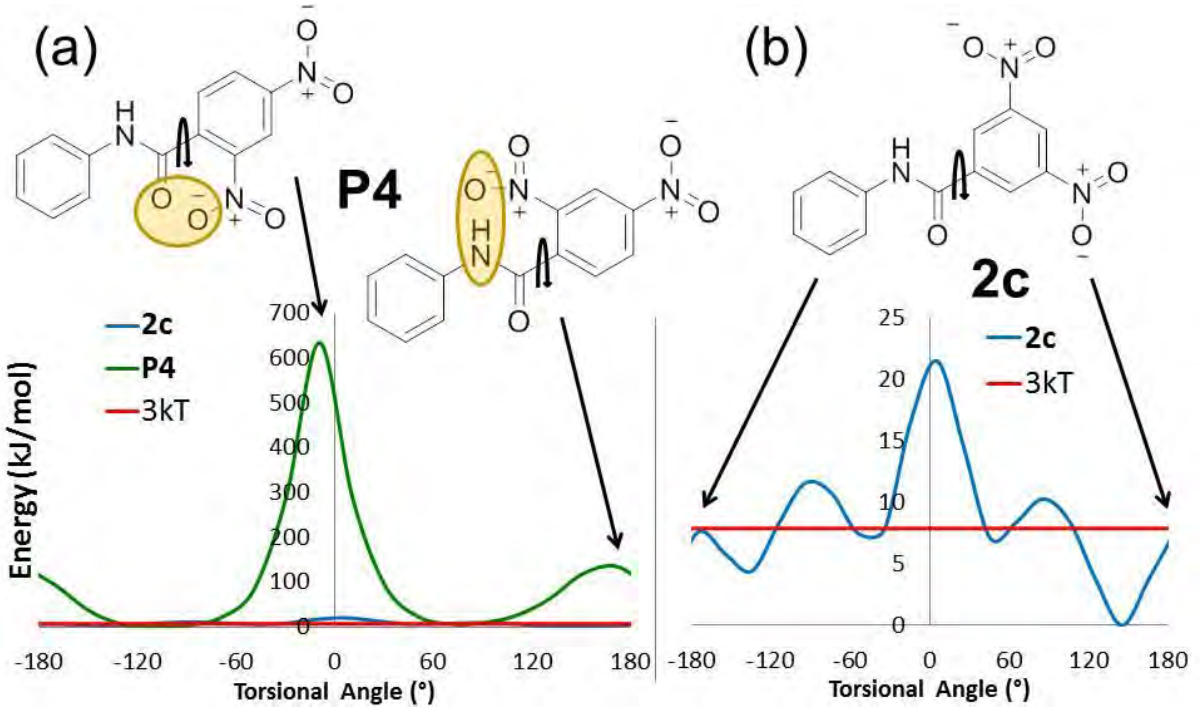


Figure 4.5 Free energy diagrams for rotation around the indicated bond where 0° or ±180° represents the flat conformation for (a) purchased compound **P4** (βH inactive) and (b) synthesised compound **2c** (βH active). The value 3kT represents the thermally accessible energy value at 37 °C. Note the difference in scale of the y-axes (a) and (b).

A summary of the SARs for the monobenzamides is given in *Figure 4.6.* These results demonstrate the minimum requirements for βH activity of the benzamides, which were later applied in the synthesis of di- and tri-benzamide derivatives (*Section 4.2.1.5*). Most importantly, only compounds with substituents in the *meta* position on ring A were chosen for further analysis, since substituents in the *ortho* and *para* positions were not tolerated for βH inhibitory activity. Furthermore, 2-pyridyl was excluded for the R_2 substituent since the results for the monobenzamides suggested that the βH inhibition activity was substantially lower in this analogue. The effect of different R_2 substituents on the SARs is further investigated for the dibenzamides in *Section 4.2.1.5*.

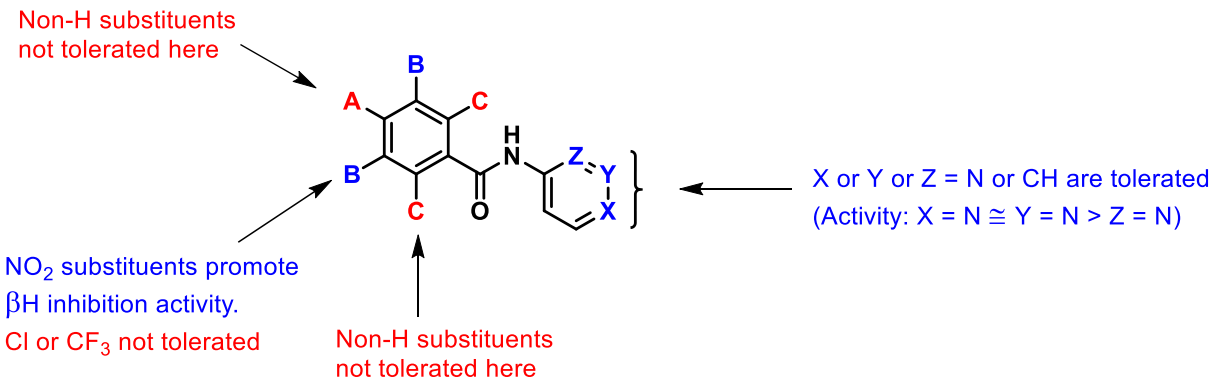
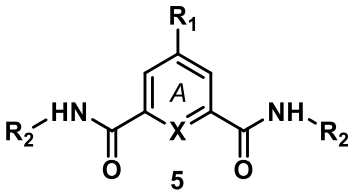


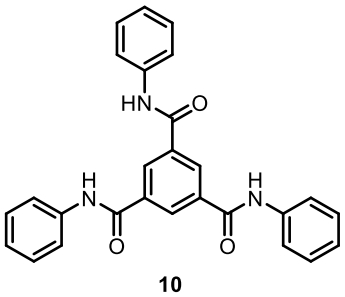
Figure 4.6 βH inhibitory SARs for the monobenzamides.

4.2.1.5 Evaluation of βH inhibition for the synthesised di- and tri-benzamides

Although two of the monobenzamides were able to inhibit βH inhibition with comparable activity to AQ, the parasite data from the HTS indicated that dibenzamides were more effective at preventing parasite growth than the monobenzamides. For this reason, a focused library of dibenzamides was synthesised and tested. In addition, the 4-pyridyl parent compound was resynthesised (**5a**) along with 3-pyridyl (**5b**) and phenyl (**5c**) derivatives. A 2-pyridyl derivative was not synthesised since the monobenzamides indicated that the 4-pyridyl and 3-pyridyl analogues were better βH inhibitors (*Table 4.2*). In addition to varying the type of ring B (R_2 substituent), the substituent on ring A (R_1) was also varied from the parent nitro to cyano, methyl, t-butyl, hydrogen or methoxy, resulting in derivatives **5b** - **5n**. These substituents were chosen for their range of electron-withdrawing vs releasing as well as hydrophobic vs hydrophilic properties, which excluded the halogens, since the chloro derivative was shown to be inactive in the purchased mono-benzamides. In addition, substituting X = CH for X = N resulted in derivatives with a pyridyl central ring as opposed to a phenyl (**5o** and **5p**) all of which are defined in *Figure 4.7,* along with the tribenzamide **10**.



5



10

5a: X = CH; R_1 = NO_2; R_2 = 4-pyridyl
5b: X = CH; R_1 = NO_2; R_2 = *3*-pyridyl
5c: X = CH; R_1 = NO_2; R_2 = phenyl
5d: X = CH; R_1 = CN; R_2 = *4*-pyridyl
5e: X = CH; R_1 = CN; R_2 = phenyl
5f: X = CH; R_1 = CH_3; R_2 = *4*-pyridyl
5g: X = CH; R_1 = CH_3; R_2 = phenyl
5h: X = CH; R_1 = *t*-butyl; R_2 = *4*-pyridyl
5i: X = CH; R_1 = *t*-butyl; R_2 = *3*-pyridyl
5j: X = CH; R_1 = *t*-butyl; R_2 = phenyl
5k: X = CH; R_1 = H; R_2 = *4*-pyridyl
5l: X = CH; R_1 = H; R_2 = phenyl
5m: X = CH; R_1 = OMe; R_2 = *4*-pyridyl
5n: X = CH; R_1 = OMe; R_2 = phenyl
5o: X = N; R_1 = H; R_2 = *4*-pyridyl
5p: X = N; R_1 = H; R_2 = phenyl

Figure 4.7 The synthesised dibenzamides, analogues of the hit compound **5a**, and the tribenzamide **10**.

Table 4.3 shows the βH inhibition values for the compounds **5a** - **5p** and **10**. Dibenzamides **5a**, **5b** and **5c** displayed similarly potent βH IC_{50} values (2.9-7.0 μM), four to ten times more active than CQ and at least three times the activity of the monobenzamide analogues. This improved activity is in agreement with the hypothesis that larger planar surfaces result in improved βH inhibition due to increased π-π stacking ability. The cyano derivatives (**5d** and **5e**) maintained excellent βH activity about six times that of CQ for both the pyridyl and phenyl rings. However, for the methyl (**5f** and **5g**), t-butyl (**5h**-**j**), hydrogen (**5k** and **5l**) and methoxy derivatives (**5m** and **5n**), only those with pyridyl rings at R_2 showed activity, while those with a phenyl ring displayed no activity below 1000 μM. Compounds **5o** and **5p** with a central unsubstituted pyridyl ring and R_1 = H showed weak activity for the pyridyl derivative and no activity below 1000 μM for the phenyl derivative. The tribenzamide **10** showed good activity, slighter better than that of CQ.

Table 4.3 βH inhibition activities for the di-benzamides **5a**-**5p** and tri-benzamide **10**.

Compound	R_1 substituent	βH inhibition IC_{50} (μM)
5a	Nitro	4.3 ± 0.1
5b		4.7 ± 0.7
5c		7.0 ± 0.2
5d	Cyano	6.34 ± 0.09
5e		4.4 ± 0.2
5f	Methyl	8 ± 1
5g		>1000
5h	*tert*-Butyl	13.3 ± 0.7
5i		6.8 ± 0.1
5j		>1000
5k	H	507 ± 33
5l		>1000
5m	Methoxy	22 ± 1
5n		> 1000
5o	H (X = N)	260 ± 7
5p		>1000
10	N/A	21 ± 1
CQ	N/A	31.5 ± 0.5

In general, a compound with a spatially flat, hydrophilic, electron withdrawing group such as NO_2 (**5a**-**c**), CN (**5d**,**e**) or benzamide (**10**)) at R_1 was able to maintain activity with either a pyridyl or phenyl ring at R_2. Conversely, when R_1 was H or a more electron releasing group, only the pyridyl derivative(s) were active and those with phenyl at R_2 showed no ability to inhibit βH formation even at a relatively high concentration. This may indicate that the molecules have two possible orientations of interaction with Fe(III)PPIX giving rise to βH inhibition. The first involves π-π stacking with central ring, which can only take place with the flat, electron withdrawing nitro and cyano groups. The second involves an interaction with the R_2 ring, which appears not to occur in the case where R_2 is phenyl as opposed to pyridyl, similar to that for naphthalene analogues of the 4-amino quinolines, which also lack βH inhibition activity.[288] This may be as a result of the relative electron richness of the phenyl as opposed to the pyridyl ring, with the latter being electron deficient owing to the electronegative nitrogen atom. Interestingly, the high IC_{50} of **5o**, the X = N derivative with

R_2 = pyridyl, can be explained in a similar manner to that of **2e**. Since both these compounds have an *ortho*-pyridyl ring B, the influence of the *ortho* nitrogen in close proximity to the amide results in a decreased probability of π-π interaction, either by causing the molecule to be less planar, or by adding too much electron density at that particular site of the molecule. The SARs for the dibenzamides are shown in *Figure 4.8.* Note that the planar and electron withdrawing nitro and cyano substituents at R_1 follow a different SAR scheme to that of the methyl, *t*-butyl, H and methoxy substituents.

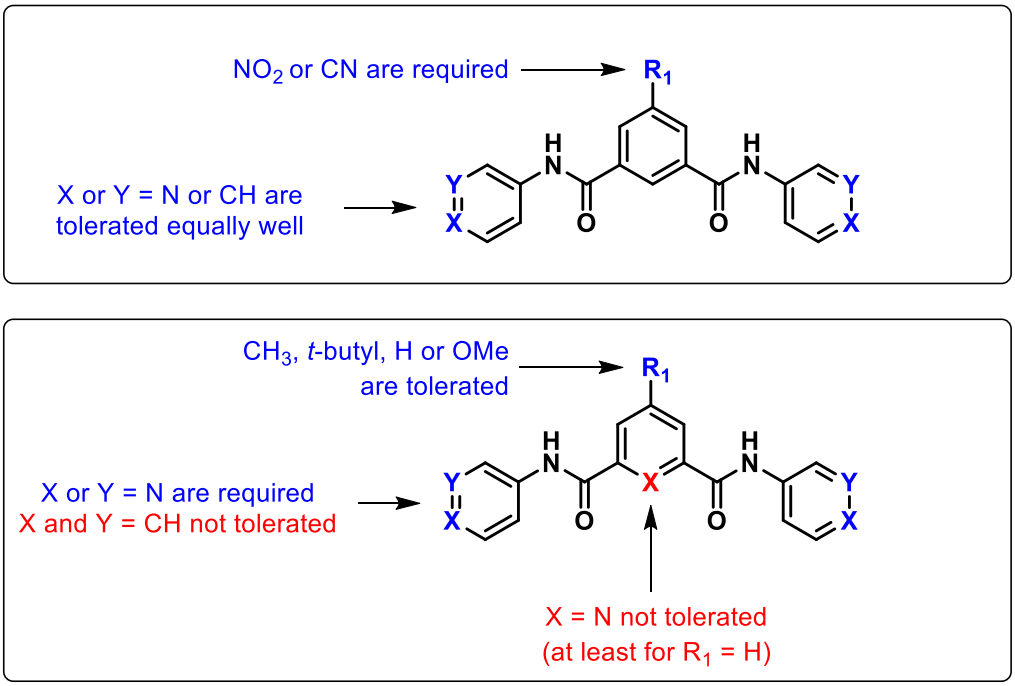


Figure 4.8 βH inhibitory SARs for the dibenzamides.

4.2.1.6 Proposed mechanism of βH inhibition for the dibenzamides

Compound **5k** showed very low activity, indicating that a non-H substituent was required at the R_1 position, although the reason for this was not clear. It does, however, show that there is a necessary role for the R_1 substituent, even when R_2 is pyridyl, since the lack of activity cannot be ascribed to hydrophobicity or electronic effects. It is possible that the substituent at R_1 aids in directing the compound into the desired orientation to inhibit βH formation. This result is qualitatively consistent with the mechanism of βH inhibition by which drug interaction is via the growing face of the crystal and not via interaction with free Fe(III)PPIX. For example, the derivatives with non-planar R_1 groups would preferentially orientate relative to the corrugated crystal face so that the substituent lies away from the crystal surface. This was observed from docking the molecule with the crystal face in

Materials Studio,[289] as shown in *Figure 4.9*. This results in the methyl, *tert*-butyl and methoxy derivatives being oriented in such a way that the R_2 ring interacts with Fe(III)PPIX molecules at the binding site. In this case, only those derivatives where R_2 is pyridyl interact strongly with Fe(III)PPIX, while those containing more electron rich phenyl rings may be able to reach the binding site, but are probably readily displaced by another Fe(III)PPIX molecule, allowing for continuation of crystal growth (*Figure 4.9*). On the other hand, **5c** and **5e** contain R_1 substituents which can fit into the crystal grooves, allowing the central ring to bind via π-π interactions to Fe(III)PPIX and for both the pyridyl and phenyl analogues to have equally potent activity. In the case of the pyridyl derivative **5k**, there is no R_1 substituent to aid in directing the molecule into the required binding groove and presumably as a result of statistically controlled binding orientations; a larger concentration is required for inhibition and a higher IC_{50} is observed.

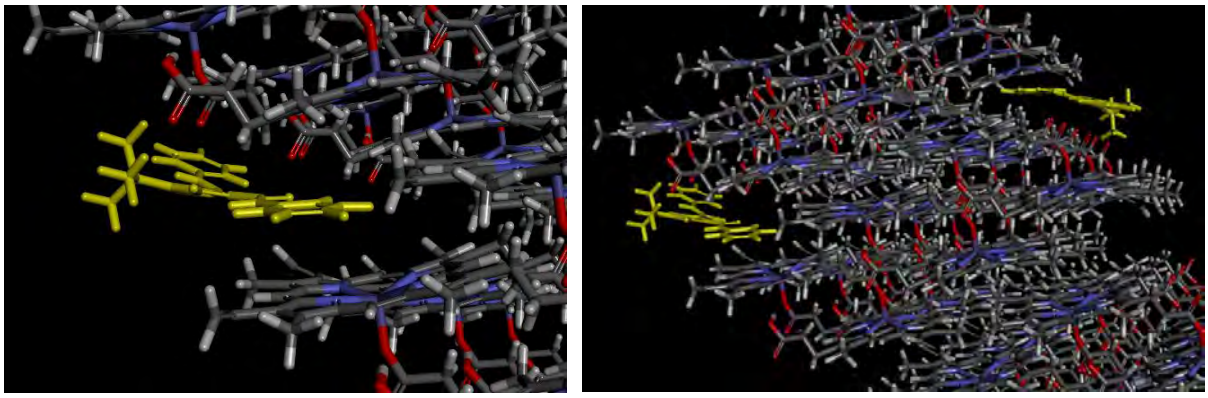


Figure 4.9 A possible mode of interaction to explain the observed βH inhibition activities. In this example, the dibenzamide, **5j** interacts with the fastest growing haemozoin crystal face {001} with the R_1 *tert*-butyl moiety directed outwards. It is likely that this substituent forces or directs the benzamide rings into the groove on the crystal face were it binds, accounting for the strong βH inhibitory activity. Images were supplied by R. Muller and generated using Materials Studio.[289]

4.2.2 Activity of benzamides against *P. falciparum* in culture

4.2.2.1 Activities of the monobenzamides

The synthesised benzamide derivatives were tested for inhibitory activity against the CQ-sensitive NF54 strain of *P. falciparum*. The monobenzamides **2b**-**e** displayed a range of IC_{50} values from 2.3 to 22 µM and **2a**, which was not a βH inhibitor, showed no activity below the maximum concentration tested, 41.3 µM (*Table 4.4*). A good correlation was identified between the inverse of the βH inhibition IC_{50} and inverse parasite growth inhibition IC_{50}, with an r^2 of 0.88 (*Figure 4.10*). By representing this relationship as the inverse of the IC_{50}, compound **2a** can be taken into account by taking its inverse IC_{50} as approximately zero. The correlation suggests that the monobenzamides do indeed exert their therapeutic action via inhibition of intracellular haemozoin formation.

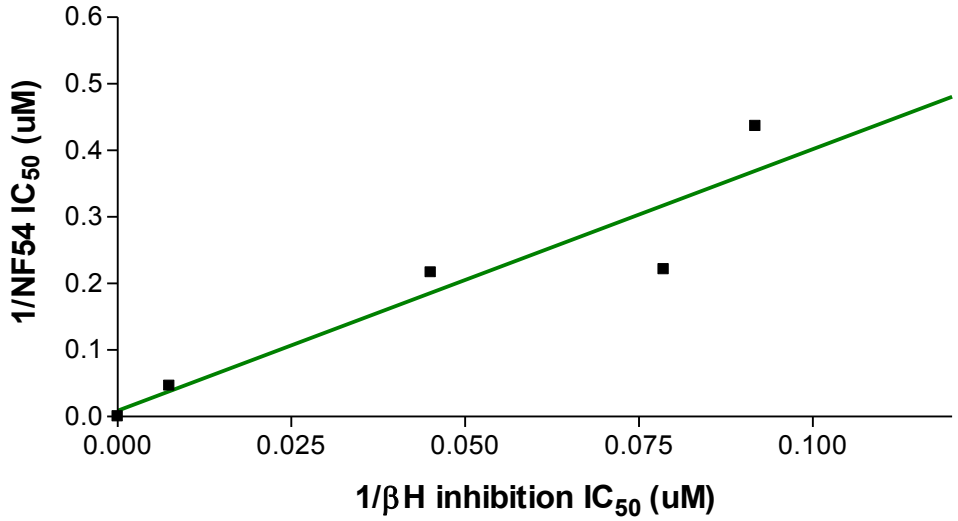


Figure 4.10 A linear correlation between the inverse of the βH and parasite growth IC_{50} values for the CQ-sensitive NF54 strain (r^2 = 0.8750; P = 0.0195).

4.2.2.2 Activities of the di- and tri-benzamides

The parasite activities of the monobenzamides were at least four-fold weaker than that of the parent compound from the HTS in the CQ sensitive D6 strain with an IC_{50} of 0.19 µM (*Table 2.3*). This hit **5a** was resynthesised and tested in the available CQ-sensitive NF54 strain. The activity was found to be three times weaker than that of the D6 strain, with an IC_{50} of 0.7 µM. The 3-pyridyl and phenyl analogues, **5b** and **5c**, exhibited a further 2-fold lower activity, similar to their decrease in βH inhibition IC_{50} values.

The other derivatives were also tested in the NF54 or D10 strains of *P. falciparum*, both of which are CQ-sensitive. Where comparisons were made, both strains gave similar IC_{50} values for tested compounds (*Table 4.4*). The strain used for a specific derivative was dictated by availability of material in combination with the strain in culture at the time of testing.

Replacing the nitro substituent (R_1) with cyano caused a 3-fold increase in IC_{50} to 1.9 µM for **5d**, the pyridyl derivative. A further IC_{50} increase to 5.2 µM was observed for the phenyl derivative **5e**. A similar reduction in activity was noticed when replacing the R_2 pyridyl ring with a phenyl group for the derivatives where R_1 is methyl (**5f** and **5g**), *t*-butyl (**5i**-**j**), hydrogen (**5k** and **5l**) or methoxy (**5m** and **5n**). Both compounds with X = N (**5o** and **5p**) were inactive up to 315 µM in the parasite, most likely because they did not possess sufficient βH inhibition activity. Only compound **5i**, the *t*-butyl derivative, showed a slight improvement in parasite activity relative to the parent compound, with IC_{50} of 0.6 µM. The trend in activities for the 4-pyridyl vs 3-pyridyl vs phenyl analogues of the t-butyl derivatives (**5h**-**j**) was different from that of the nitro derivatives (**5a**-**c**), whereby the order of potency between the 3-pyridyl vs 4-pyridyl was reversed. In all cases, however, the phenyl derivatives were less active against *P. falciparum in vitro*. The tribenzamide, **10**, showed very poor parasite activity despite being a good βH inhibitor. This may have been as a result of its excessive hydrophobicity (predicted logP = 5.21). The data are summarised in *Table 4.4*.

Table 4.4 βH inhibition and CQ-sensitive parasite activities for the synthesised mono-, di- and tri-benzamides.

Compound	Derivative info	βH inhibition IC_{50} (μM)	NF54 parasite IC_{50} (μM)	D10 parasite IC_{50} (μM)
Monobenzamides				
2a	Monobenzamides	>1000	>41.3 ± ND	ND
2b		22. ± 1	5 ± 1	ND
2c		10.9 ± 0.3	2.3 ± 0.3	2.4 ± ND
2d		13 ± 1	5 ± 1	2.3 ±ND
2e		134 ± 8	22 ± 4	24 ± 5
Dibenzamides				
5a	R_1 = Nitro	4.3 ± 0.1	0.7 ± 0.2	0.9 ± 0.1
5b		4.7 ± 0.7	1.32 ± 0.04	ND
5c		7.0 ± 0.2	1.4 ± 0.2	ND
5d	R_1 = Cyano	6.34 ± 0.09	1.9 ± 0.2	2.0 ± 0.2
5e		4.4 ± 0.2	ND	5.2 ± ND
5f	R_1 = Methyl	8 ± 1	2.5 ± 0.6	ND
5g		>1000	146 ± 1	ND
5h	R_1 = *Tert*-butyl	13.3 ± 0.7	1.6 ± 0.2	ND
5i		6.8 ± 0.1	0.6 ± 0.1	0.7 ± 0.1
5j		>1000	8.1 ± 0.5	ND
5k	R_1 = H	507 ± 33	9 ± 4	ND
5l		>1000	>314	ND
5m	R_1 = Methoxy	22 ± 1	3.0 ± 0.5	ND
5n		> 1000	37 ± 13	ND
5o	R_1 = H (X = N)	260 ± 7	>315	ND
5p		>1000	>315	ND
Tribenzamide				
10	Tri-phenylbenzamide	21 ± 1	81 ± 12	>42
4-Aminoquinoline				
CQ	Standard	31.5 ± 0.5	0.008 ± 0.002	0.021 ± 0.002

As with the monobenzamides alone, inverse activity of both the βH and NF54 parasite IC_{50}s were correlated when all the benzamides were taken into account. The relationship was statistically significant with $P < 0.0001$. Furthermore, a moderately good r^2 value of 0.68 was observed for the correlation, which increased to $r^2 = 0.84$ upon removal of the point corresponding to compound **5i**, which was an apparent outlier (*Figure 4.11*). This compound showed a better parasite activity than expected based on its βH inhibition IC_{50}. Possible reasons for this could be that its properties differ markedly from the other compounds, that there is favourable uptake in the parasite, or that the βH IC_{50} measured for this compound

in the NP-40 assay does not accurately represent its haemozoin inhibitory activity in the parasite. Nevertheless, this correlation strongly suggests that haemozoin formation is the target.

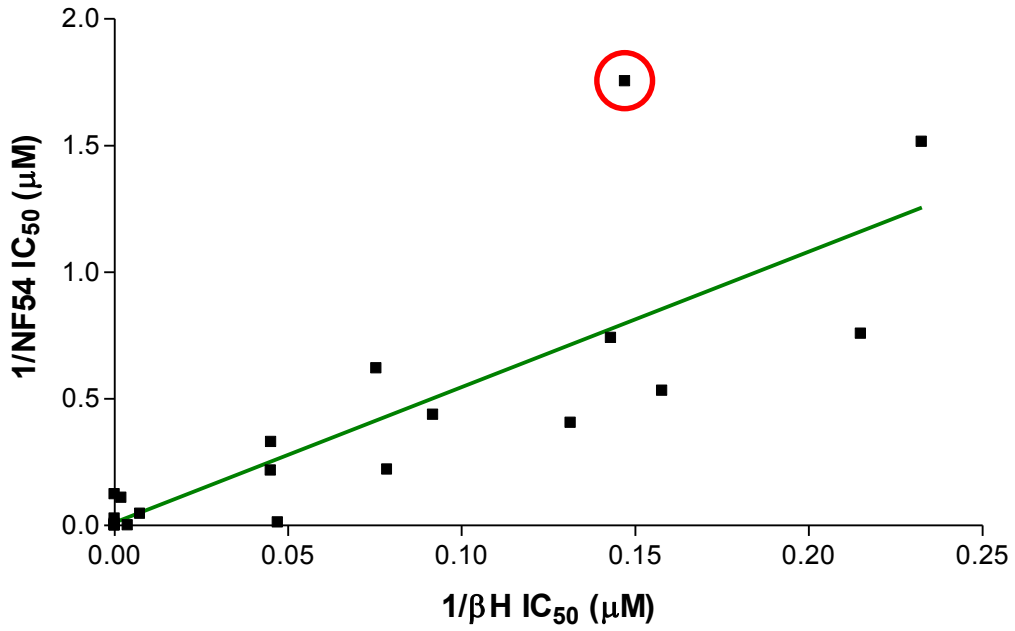


Figure 4.11 The correlation between the inverse of NF54 parasite activity and the inverse of βH inhibition activity (r^2 = 0.68, P < 0.0001). The statistics improve when derivative **5i** (red circle) is excluded (r^2 = 0.84, P < 0.0001).

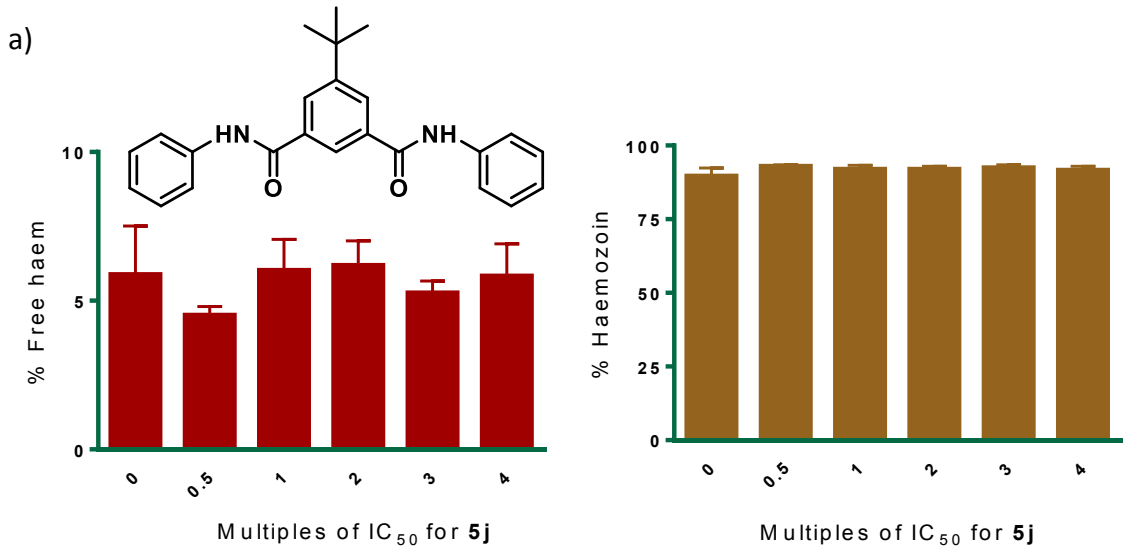
Attempts to improve the SARs further were carried out by including other parameters in the correlation analysis. A range of physicochemical properties were predicted for each compound using Molecular Modelling Pro Plus Software[290] and linear regression with all the combinations of properties was carried out *in silico*. The addition of properties such as solubility, surface area or dipole moment to the βH inhibition activity, the two-parameter linear combination of which predicted parasite activity, was found to favourably influence the statistics. These relationships suggested that compounds with lower solubility, larger surface area or larger dipole moment resulted in a higher parasite activity. Indeed, compound **5i**, with unpredictably good parasite activity, possessed one of the lowest predicted aqueous solubilities, while also having one of the largest surface areas and the biggest dipole moment of all the combined mono and dibenzamides. Including each of these parameters in the equation individually, resulted in three new correlations with r^2 values of

0.74 - 0.75, compared to the initial 0.68 with only βH inhibition activity, however, unlike the inverse βH inhibition IC_{50}, these new parameters were not significant with $P > 0.05$ (2-tailed t-test). Furthermore, when attempts were made to find correlations with three or more parameters, those included, other than the inverse βH inhibition IC_{50}, were also not significant. This was most likely since, even though the solubility, surface area and dipole moment parameters improved the r^2 values, the additional parameters primarily enabled compound **5i** to better fit the trend, while not providing improvements in the correlation for the other compounds. Indeed when **5i** was excluded, the addition of solubility, surface area or dipole moment did not significantly improve the r^2 value for the relationship involving only the βH inhibition activity. This gave further indication that βH inhibition may be the mechanism of action for these benzamides, however more advanced techniques were required to convincingly demonstrate this.

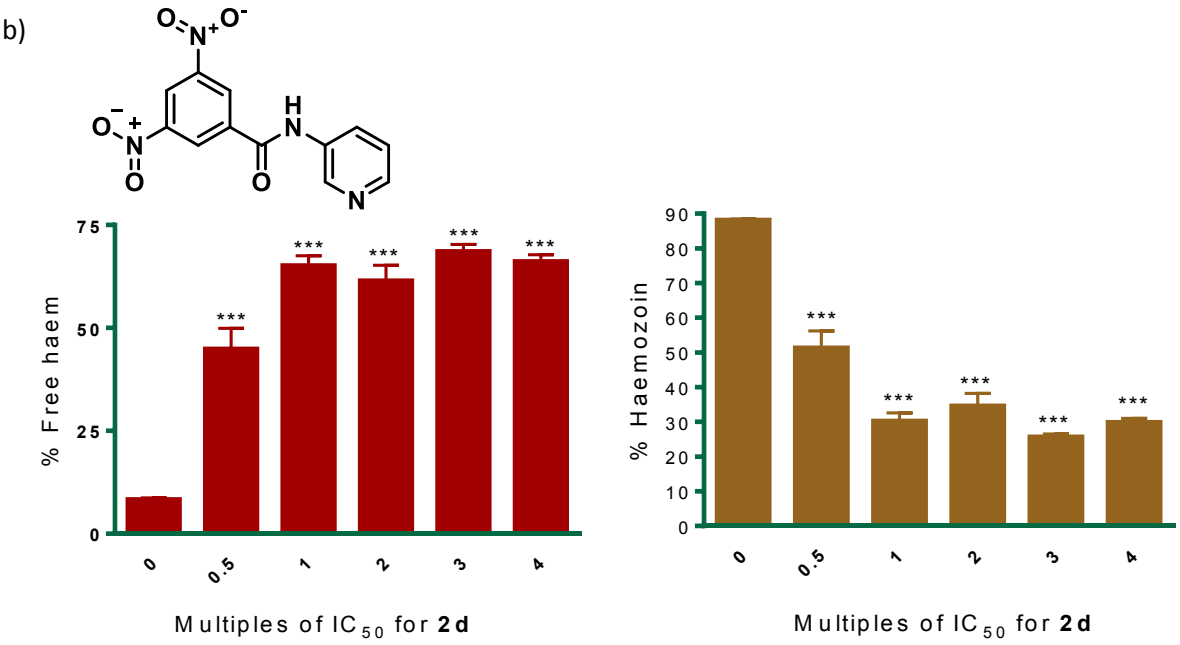
4.2.2.3 Target validation for the benzamides

Direct comparison of the benzamide derivatives with known drugs such as AQ and CQ is valid only if their mechanisms of action within *P. falciparum* cells are similar. *Section 2.2.4.1* describes the target validation for the HTS hits. In particular, it was shown that compound **5a** induces a 33% increase in free haem, accompanied by a decrease in haemozoin formation in the parasite. However, data for additional derivatives was required in order to be confident of the biological mechanism of action for this scaffold. Three benzamides were sent for analysis in the cell fractionation assay. Two of the selected compounds (**2d** and **5i**) showed excellent βH inhibition activity in the NP-40 assay, while one compound was not a βH inhibitor (**5j**). The latter served as a negative control. The monobenzamide **2d** and the dibenzamide **5i** were chosen because they have relatively low and high parasite growth inhibition activity respectively. This was desirable so as to observe whether free haem levels had an influence this activity. Compound **5j** had a relatively low activity against the parasite, probably because it is unable to inhibit βH formation. However, its ability to inhibit parasite growth, albeit at a high concentration, suggested that either **5j** did inhibit haemozoin formation in the cell at sufficiently high concentration or that it had an alternate, off-target mechanism of action. As shown in *Figure 4.12a,* there was no significant change in the free haem or haemozoin levels with increasing concentration of **5j**. This was in agreement with the NP-40 assay for βH inhibition and provided strong support for the technique for finding haemozoin inhibiting compounds used in this study. It would appear that this compound acts weakly on an alternate target. Investigation of identity of the non-haemozoin target was considered to be beyond the scope of this study and was of relatively little interest, given the low parasite activity of **5j**. In contrast, **2d** and **5i** showed a large increase in the percentage free haem (>50% at 2.5 times the IC_{50}) with a corresponding dramatic decrease in haemozoin (<40% at 2.5 times the IC_{50}) within the parasite (*Figure 4.12b,c*). The significant changes in free haem and haemozoin levels for these benzamides suggested that the precise manner by which they cause haem-related toxicity differs from that of CQ, since less accentuated changes in free haem and haemozoin levels of ~15% and ~80% respectively were observed at 2.5 times the CQ IC_{50}. Other quinoline antimalarials, such as AQ, QN and MF showed even smaller changes than CQ.[113] This finding demonstrates that higher levels of free haem are required for disrupting parasite growth in the case of the benzamides relative to CQ and other quinoline antimalarials.

a)


HN
NH
O
O
% Free haem
10
5
0
0
0.5
1
2
3
4
Multiples of IC50 for 5j
% Haemozoin
100
75
50
25
0
Multiples of IC50 for 5j

b)


O
N+
O-
O
N+
O-
O
H
N
N
% Free haem
75
50
25
0
***
Multiples of IC50 for 2d
% Haemozoin
90
80
70
60
50
40
30
20
10
0
Multiples of IC50 for 2d

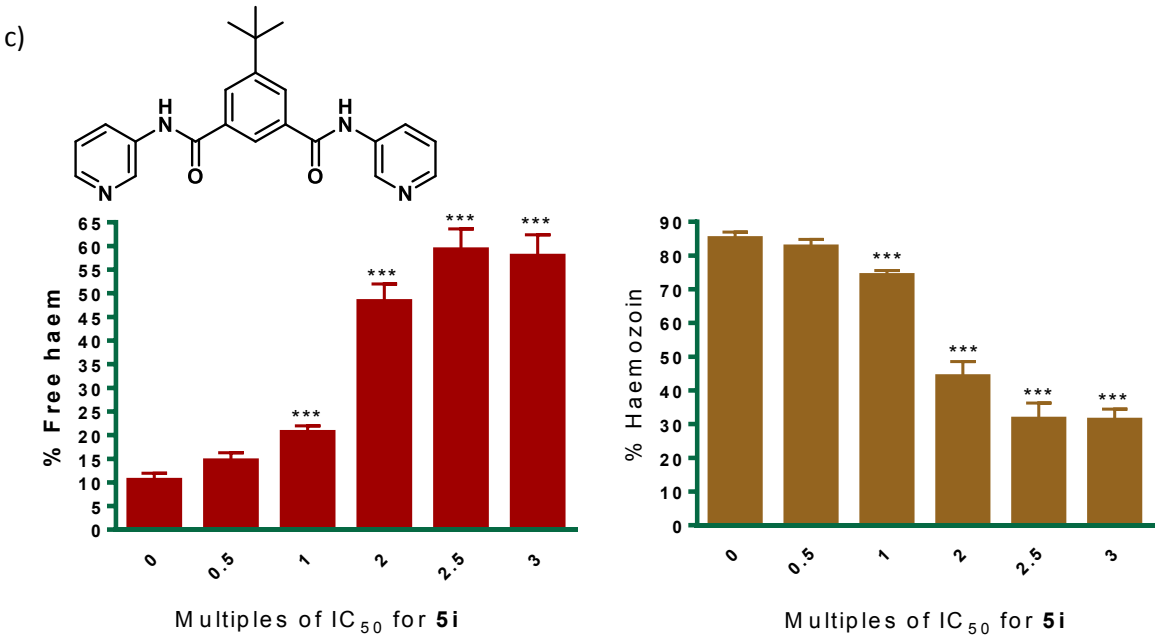


Figure 4.12 Free haem and haemozoin levels from cultured *P. falciparum* treated with varying concentrations of benzamides a) **5j**, a negative control, b) **2d**, a monobenzamide and c) **5i**, a dibenzamide. Asterisks indicate statistical significance relative to no drug (2-tailed t-test): *P < 0.05; **P < 0.01; ***P < 0.001, n = 3.

In the case of the βH inhibitors, the free haem curves and the parasite survival curves with varying test compound concentrations crossed over close to the D10 IC_{50} of **2d** (*Figure 4.13a*) and **5i** (*Figure 4.13b*), convincingly supporting the target as haemozoin formation.

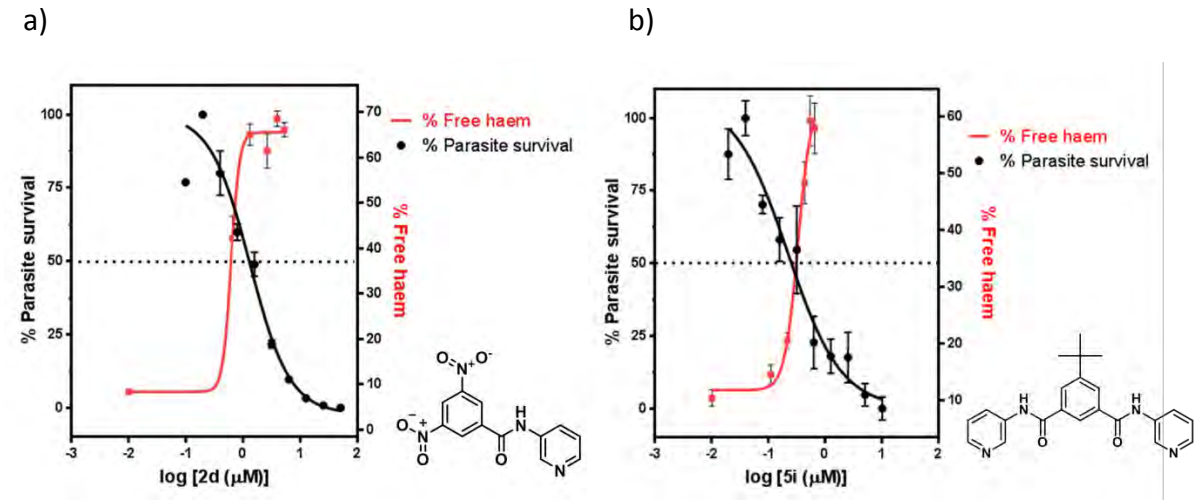


Figure 4.13 The concentration dose response curves for the % free haem and the % parasite survival, which cross near the IC_{50} for compounds a) **2d** and b) **5i**. Scales adjusted so that the IC_{50} for parasite survival coincides with the 50% value for increase in free haem (indicated by the dotted line).

4.2.3 Attempts to improve activity through pH trapping

4.2.3.1 Design of a basic analogue and pH trapping predictions

Many of the phenyl derivatives were not βH inhibitors, however, **5c** and **5e** did inhibit βH formation. The observed decrease in activity for these phenyl derivatives relative to their pyridyl analogues **5a** and **5b**, were initially thought to be the result of decreased accumulation. This was because, at pH 4.8, the basic derivatives **5a** and **5d** are predicted to be about 75% diprotonated owing to the two pyridyl rings on each molecule (pK_{a1} = 5.92, pK_{a2} = 5.32). This may be expected to result in pH trapping of these derivatives in the DV, leading to improved IC_{50}s relative to their phenyl analogues. If this were the case, designing a compound with greater basicity may be expected to enhance the pH trapping effect and result in improved parasite activity. This is believed to be the case for the 4-aminoquinolines AQ and CQ. Egan *et al.*[182] reported that 4-amino-7-chloro-quinoline is able to inhibit βH formation, but is only weakly active in the parasite with a D10 IC_{50} of 3.8 µM. Upon attachment of a variety of tertiary amino containing side chains at the 4-amino position, βH inhibition was maintained and parasite activity was greatly enhanced by approximately 100-fold. According to the study, this was probably as a result of improved accumulation in the DV. In addition, it was proposed that rational design could be employed to develop novel antimalarials from scaffolds that form π-π complexes with Fe(III)PPIX and inhibit βH formation (*Figure 4.14*).[182]

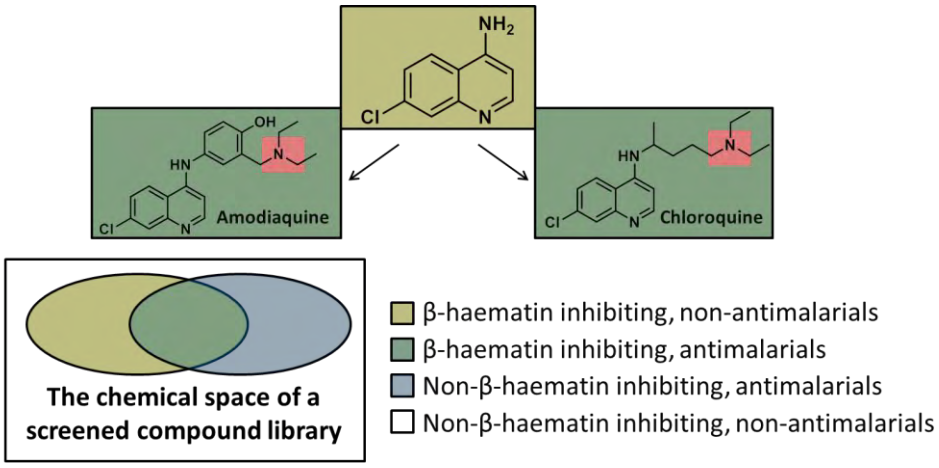
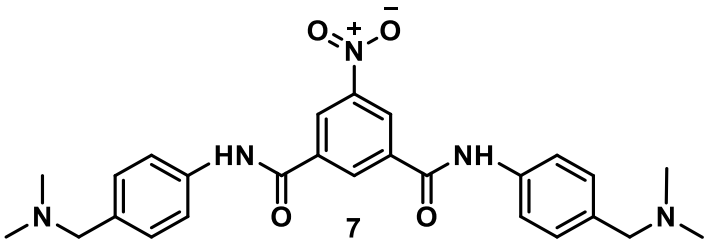


Figure 4.14 Diagram showing the potential for βH inhibiting compounds to be derivatised to potent antimalarial compounds by introducing basic tertiary amino side chains, as was the case for AQ and CQ.

This strategy was applied to the benzamide scaffold. A sufficiently basic compound required introduction of a nitrogen atom with a lone pair which was not delocalised into a π-orbital system, implying an aliphatic amine. Since the R_1 substituent was shown to have a significant influence on activity, while the R_2 rings could be varied without abolishing the activity for **5a**-**5e**, phenylmethanamines were initially envisaged for the R_2 substituents. However, a derivative with two primary amines would have resulted in a diprotonated compound, even at neutral pH, owing to the predicted high amine pK_as of 9.76 and 9.16 for the first and second protonations respectively. Further calculations showed that such a compound would be too hydrophilic, with a logD of -1.7 at neutral pH and would therefore be unlikely to cross the four parasite membranes to reach the DV. Alternatively, alkylation of the primary amine to form secondary and tertiary amines improves hydrophobicity and logD at pH 7.4. The physiochemical properties for a range of alkylated derivatives were predicted in order to identify the most suitable with molecular weights below 500 g/mol and small, but positive logD at 7.4. Derivatives with secondary amines had predicted logD values below -1 for neutral pH, making them too hydrophilic, as in the case of the primary amine. The introduction of ethyl or propyl groups onto the primary amine derivative resulted in compounds with high molecular masses that exceeded 500 g/mol. Instead, a derivative where the R_2 substituents were each *N*,*N*-dimethyl-phenylmethanamine (**7**), possessed a molecular weight of 475 g/mol and a positive logD value of 1.12 at neutral pH (*Figure 4.15*). This compound was thus selected for synthesis in order to probe the effect of greater basicity and pH trapping for the benzamide series.



7

Figure 4.15 The dibenzamide **7** selected for synthesis with basic side chains, where R_2 = *N*,*N*-dimethyl-phenylmethanamine. The tertiary amine derivative possessed an improved logD relative to a primary amine derivative at neutral pH and a molecular mass below 500 g/mol.

Compound **7** was tested for both βH inhibition and parasite activity. In the NP40 assay, the compound maintained potent βH activity with an IC_{50} of 7.9 ± 0.2 μM, appearing more active AQ and CQ. This confirmed that the R_2 rings did not greatly influence βH inhibition activity for derivatives with an electron withdrawing group at R_1. The result also agreed with the hypothesis that these analogues most likely inhibit βH formation via π-π interactions between Fe(III)PPIX and the central ring, as opposed to the R_2 ring. However, upon assessing the activity against *P. falciparum* of this more basic derivative, which was expected accumulate to a greater extent in the parasite DV than **5a** and hence possess enhanced activity, no improvement was found. An NF54 parasite IC_{50} of 1.5 ± 0.2 μM was obtained. This activity was the same as that of the phenyl derivative, **5c**, within the error margin, suggesting that no significant accumulation had taken place. This lack of effect was especially striking when compared with the predicted accumulation that should have taken place due to pH trapping. The predicted VAR due to pH trapping alone was calculated using the Henderson–Hasselbalch equation (*Eq. 4.1*), the derivation of which is shown in *Section 8.1.3.2*, at an external pH of 7.4 (pH_e) and intravacuolar pH of 4.8 (pH_v) for a compound (B) with two protonatable sites.

$$VAR = \frac{[B]_{T,v}}{[B]_{T,e}} = \frac{(1+10^{pK_{a2}-pHv}+10^{pK_{a1}+pK_{a2}-2pHv})}{(1+10^{pK_{a2}-pHe}+10^{pK_{a1}+pK_{a2}-2pHe})} \qquad \textbf{Eq. 4.1}$$

Using the predicted pK_a values of 9.12 and 8.52 for compound **7**, the expected VAR was estimated at ~147,000. This indicated that the derivative should have been about 147,000 times more concentrated in the DV compared with **5c**, which is not able to pH trap. Since the compounds possessed the same activity, two hypotheses were suggested; either compound **7** does not effectively cross the parasite membrane, resulting in a higher required concentration to inhibit parasite growth, or compound **7** accumulates but is not able to inhibit haemozoin formation in the DV.

4.2.3.2 Target validation for derivative 7

In order to evaluate whether the basic derivative, **7**, was actually capable of preventing haemozoin formation in the parasite, the compound was sent for cell fractionation studies and subjected to similar analysis as that of **5a** in the CQ sensitive D10 strain of *P. falciparum.* At a concentration of 450 nM, a 15 percentage point increase in free haem was observed for

7 relative to the control with a concentration of zero nanomolar. At a concentration of 900 nM, approximately the IC_{50} of **7** in this strain, the amount of free haem increased further, with 41% free haem observed compared to the control of 7%. A simultaneous decrease in the % haemozoin was observed from 89% for the control to 53% for **7** at 900 nM (*Figure 4.16*).

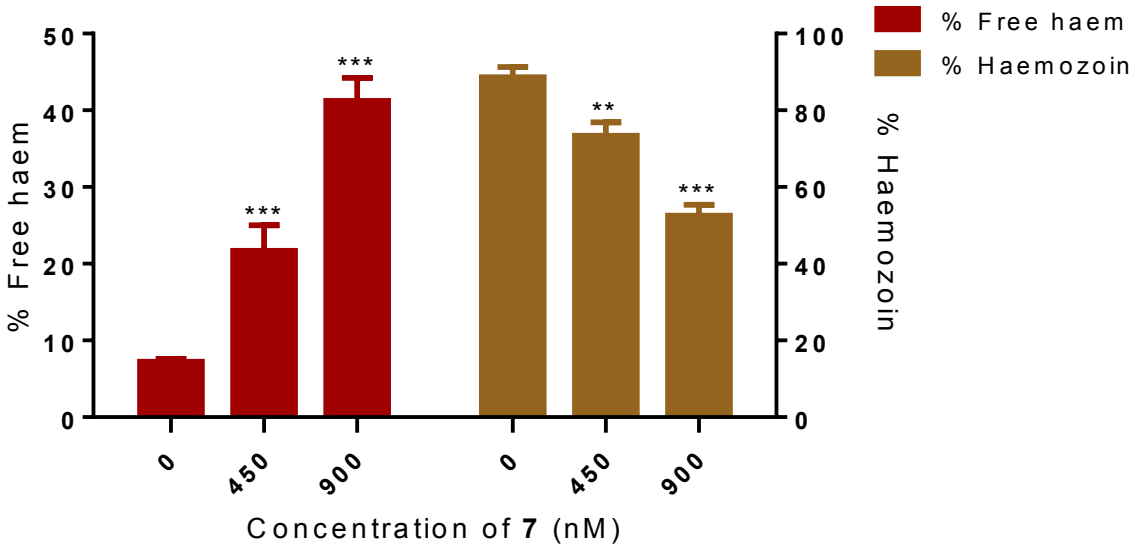


Figure 4.16 A sharp increase in the % free haem accompanied by a decrease in the % haemozoin within the CQ-sensitive D10 strain of *P. falciparum* for compound **7**. Asterisks indicate statistical significance relative to no drug (2-tailed t-test): **P < 0.01; ***P < 0.001, n = 3.

This method conclusively demonstrated that **7** inhibits cellular haemozoin formation, accompanied by an increase in the percentage free haem. This strongly suggests that the proposed haemozoin target is indeed valid. The compound showed a similarly large increase and decrease in free haem and haemozoin respectively as the pyridylbenzamides (*Section 4.2.2.3*).

4.2.3.3 The inoculum effect and cellular accumulation

With the biological target for the basic analogue strongly supported, the ability of **7** to accumulate in parasitised red cells relative to **5a-c** was investigated by measuring the inoculum effect for several derivatives. The inoculum effect is the phenomenon whereby a significant increase in the inhibitory concentration of a drug is observed when the number

of organisms inoculated is increased.[291] The effect is only detected when accumulation of the drug occurs in the organism, in this case, the parasitised erythrocyte. The apparent decrease in activity with greater parasitaemia can be rationalised by the depletion of total drug concentration in the extracellular medium upon incubation with infected red blood cells, due to a significant uptake by the parasite. By measuring the changes in IC_{50} with varying initial parasitaemia levels, the amount of drug accumulation can be quantified.[292,293]

Accumulation measurements were initially carried out for the standard drugs CQ and AQ. Significantly large differences in the IC_{50} values were found at varying concentrations of infected erythrocytes. *Figure 4.17* shows plots of the antimalarial IC_{50} obtained from the LDH assay versus the inoculum size for a particular experiment. The inoculum size is the product of the parasitaemia and haematocrit. In these experiments the haematocrit was maintained at 1% and the parasitaemia varied from 1 to 10%. A direct linear trend, characteristic of the inoculum effect, was observed for both AQ and CQ, with greater parasitaemia associated with a decrease in drug activity. This relationship was extrapolated to give the absolute IC_{50} at an inoculum size of zero in order to calculate the CAR defined in *Eq. 4.2.* This equation takes into account the fractional volume of the parasitised red blood cells ($V_{frac.PRBCs}$), which is the product of the parasitaemia and haematocrit, expressed in decimal form. See *Section 8.1.4.3* for the derivation of *Eq. 4.2*.

$$CAR = \frac{IC_{50_{measured}} - IC_{50_{absolute}}}{IC_{50_{absolute}} \times V_{frac.PRBCs}} \qquad \textbf{Eq. 4.2}$$

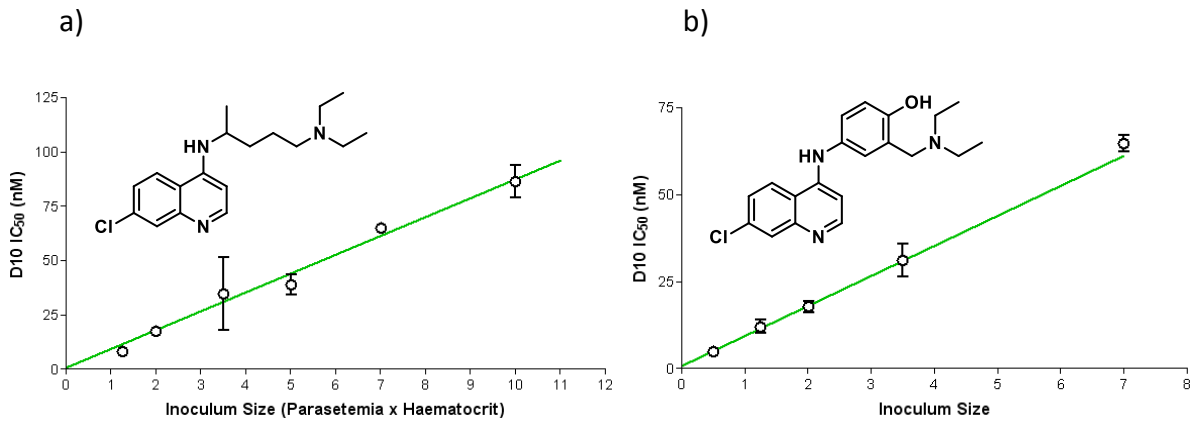


Figure 4.17 The inoculum effect observed for a) amodiaquine, with a CAR of 98024 ± 15296 and b) CQ with a CAR of 105,343 ± 3365.

The calculated CAR values for AQ and CQ were the same within the standard error and suggest that there was a 100,000-fold increase in concentration from the extracellular medium to the parasitised cell. The large measured accumulation value for CQ (105,343 ± 3365) is close to that determined by Walden of 130,081.3 ± 14,343.8 in the D10 strain.[294]

It is possible to predict the CAR due to pH trapping for CQ from the predicted VAR, which in turn can be calculated from the experimental pK_a values of CQ reported by Warhurst *et al.*[122] This molecule has two protonatable sites, namely the quinoline ring nitrogen and the tertiary diethylamine nitrogen, resulting in two measured pK_as of 8.38 and 10.18 respectively. The VAR at an external pH of 7.4 (pH_e) and intravacuolar pH of 4.8 (pH_v) for a doubly protonatable compound can be calculated directly from these pK_a values using the Henderson–Hasselbalch equation (*Eq. 4.1*).

This gives a predicted VAR of ~143,000 for a pH_v of 4.8, which can be used to calculate the CAR based on vacuolar pH trapping alone. The DV occupies an estimated 5% of the total erythrocytic volume.[14,295] The volume of the parasitised red blood cell is therefore approximately 20 times larger than that of the DV, resulting in a CAR which is 20 times smaller than the VAR. Thus, the corresponding predicted CAR for CQ is ~7000, at least 10-fold lower than the experimentally derived CAR in D10 parasites. However, the predicted CAR is close to that experimentally determined by Hawley *et al.*[99] also using inoculum effect measurements, of 5,768 for CQ in the CQ-sensitive 3D7 strain. A much lower CAR for CQ of 1859 was reported by O'Neill and co-workers in the CQ-sensitive HB3 strain.[293] This may

indicate a difference in the accumulation mechanism of these antimalarials in the D10 strain. Alternatively, differences in the DV pH would have a significant effect on the pH trapping-based accumulation ratio, since the predicted VAR is highly sensitive to this parameter. For example, an assumed pH_v of 4.5 results in a VAR of ~570,000 and a predicted CAR of ~28,000, greatly increased from the value of ~7000 calculated using a pH_v of 4.8. However, this would still not be able to account for the observed ~100,000-fold CAR for CQ. Instead, it is possible that these values are an overestimation of the CAR, due to the fact that the IC_{50} at low inoculum size is very small and extrapolation leads to a y-intercept value ($IC_{50absolute}$) which is close to zero. For example, if the true $IC_{50absolute}$ was 1.8 instead of the 0.8 observed from extrapolation, the experimentally determined CAR would only be ~40,000.

Nevertheless, the large cellular uptake of these drugs in the D10 strain appears to be due to more than pH trapping alone since the calculated and experimentally determined pH trapping ratios for CQ and AQ are not in agreement. Interestingly, only the unsaturable uptake of CQ can be taken into account by pH-trapping, while Bray *et al.*[59] have shown that a significant quantity of saturable CQ-Fe(III)PPIX binding seems to occur and can account for differences in accumulation and activity between CQ-sensitive and CQ-resistant parasite strains. This idea was originally suggested by Fitch and co-workers in in 1980s.[296,297] In fact, saturable accumulation was shown to be responsible for the four-fold greater accumulation of CQ in the sensitive HB3 strain as opposed to the resistant K1 strain. The data derived from the experiments in this project are also in support of the saturable accumulation theory, since Fe(III)PPIX-binding may account for the large CARs determined in the D10 strain. This is investigated and described in more detail in *Section 4.2.4.*

A much lower level of accumulation was expected for the benzamides as a result of their higher parasite IC_{50} values compared to CQ and AQ. This was especially since the benzamides were shown to be as good as or better than the 4-aminoquinolines at inhibiting βH formation. The resynthesised hit compound from the HTS (**5a**) along with the most potent dibenzamide (**5i**) and the derivate with basic tertiary amino side chains (**7**) were evaluated for accumulation in the parasite and the linear relationship for activity versus the inoculum size was extrapolated to obtain the absolute IC_{50} at zero percent parasitaemia (*Figure 4.18*).

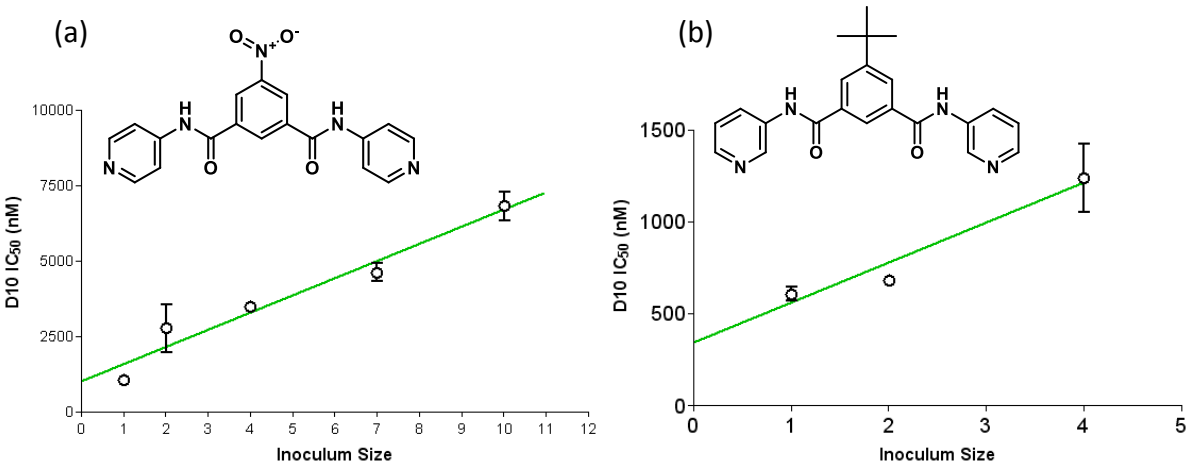


Figure 4.18 The inoculum effect observed for (a) compound **5a** with a CAR of 6046 ± 2736 and (b) compound **5i** with a CAR of 5934 ± 945.

Both compounds **5a** and **5i** showed at least 16-fold lower accumulation than CQ and AQ with CARs of ~6000. This decreased accumulation likely accounts, in part, for the 30-fold weaker activity of these dibenzamides compared with CQ. Initially this was believed to be due to the lack of highly basic functionalities for pH trapping that are present in CQ. However, upon testing the CAR for **7**, an even weaker accumulation of approximately 1000-fold was found and a relatively small change in activity with inoculum size was observed (*Figure 4.19*).

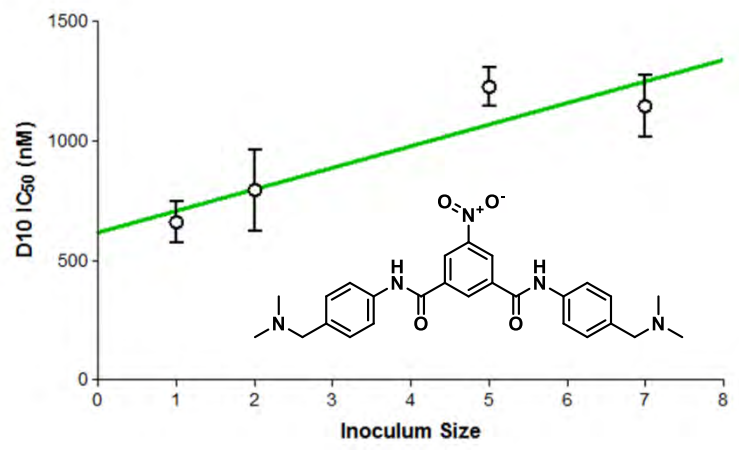


Figure 4.19 The inoculum effect observed for compound **7** with CAR of 1321 ± 533.

This 100-fold decrease in accumulation compared to that of CQ indicated that the basic benzamide derivative was not capable of pH trapping to the extent of CQ. This was intriguing given the two very basic substituents which were present, which resulted in a

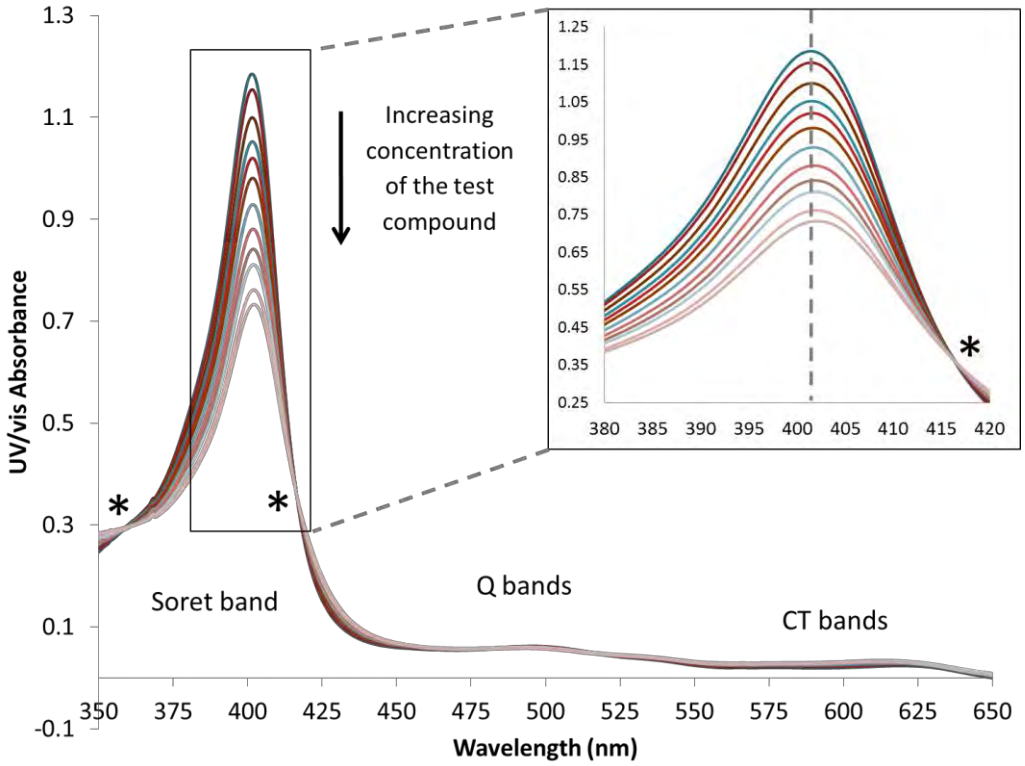
predicted pH trapping ratio (VAR) of ~147,000 using predicted pK_a values from MarvinSketch software.[269] This corresponds to a predicted pH trapping-based CAR of ~7000, similar to that of CQ. The lack of correlation between the calculated pH trapping ratio and the observed accumulation ratio for the 4-aminoquinolines and the benzamides suggested that the mechanism of drug accumulation is much more complex, at least in the CQ-sensitive D10 strain. This is especially evident when noting that, although the CARs for **5a** and **5i** are lower than CQ and AQ, they are still concentrated by over 1000-fold in the parasitised cell despite having very little pH trapping ability. The predicted CARs due to pH trapping for **5a** and **5i** were calculated to be ~2.8 and ~0.1 respectively, showing that accumulation of these pyridyl benzamides is not as a result of pH trapping. These studies further showed that the parasite activity of haemozoin inhibitors cannot necessarily be improved by adding basic moieties, as proposed by the hypothesis in *Figure 4.14*.

4.2.4 Possible role of Fe(III)PPIX association in cellular accumulation

Although pH trapping could not directly account for the poorer activities and weaker accumulation of the benzamides relative to the known drugs, some studies have suggested that the cellular uptake of aminoquinolines and bis-quaternary ammonium compounds is via binding to ferriprotoporphyrin IX.[125,59,298,299] Although there is no consensus as to the validity of the conclusions of these studies, the extent of Fe(III)PPIX binding for selected benzamides was investigated in an attempt to determine whether there is a correlation with cellular accumulation. This involved measurement of the association constants via titration of the compounds into a solution of 7.5 µM monomeric haematin in 40% v/v DMSO at pH 7.5. Upon titration of the test compound solution into the haematin solution, a hypochromism or convergent decrease in intensity of the characteristic monomeric haematin spectrum at 401 nm (Soret band) was observed owing to significant Fe(III)PPIX binding. The information gained from this experiment could be extended by monitoring the maximum absorbance at 405 nm as a function of test compound concentration. This allowed the association constant (K) to be calculated from a non-linear least squares fitting model for a 1:1 complex. A representative example is shown in *Figure 4.20*. It should be noted that the benzamide solutions were made up in 100% DMSO owing to their low solubility in 40% DMSO. This resulted in a small increase in total percent DMSO in the

working solution over the course of the titration. However, this did not appear to affect the absorbance and isosbestic points were observed as expected, confirming the 1:1 association stoichiometry.

a)



b)

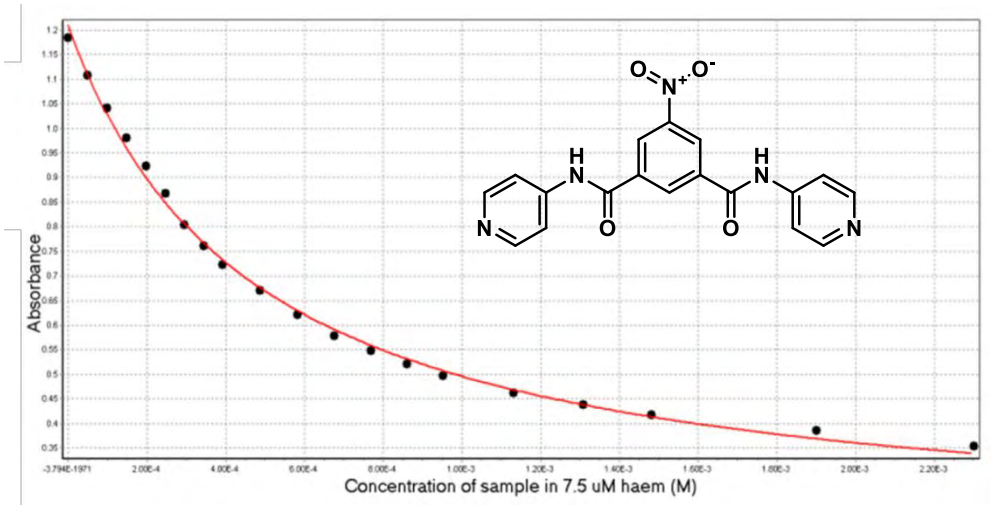
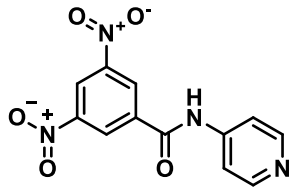
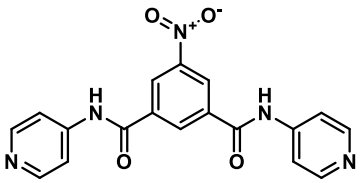
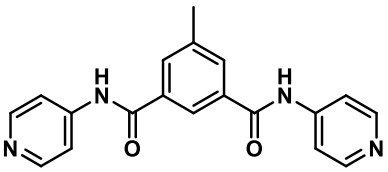


Figure 4.20 Measurement of the association constant between compound **5a** and Fe(III)PPIX at pH 7.5 in 40% v/v DMSO. Plots of a) the hypochromic effect of the Fe(III)PPIX absorbance spectra with increasing compound concentration, showing 1:1 binding confirmed by the isosbestic points (*) and b) the maximum Fe(III)PPIX absorbance vs the concentration of compound **5a**.

Haematin binding experiments were also attempted for compounds **2c**, **5f** and **7**. However, the last precipitated out of solution and useable data could not be obtained owing to its low solubility. The data relating to strength of Fe(III)PPIX binding are shown in *Table 4.5* for synthesised benzamides **2c**, **5a** and **5f**, as well as for selected 4-aminoquinolines**.** The three benzamides showed similar logK values, on the order of 100-fold lower than those of the known antimalarials and 10-fold lower than 4-amino-7-chloroquinoline. These values are consistent with the hypothesis that the benzamides associate with Fe(III)PPIX through π-π interactions alone, in accordance with the predicted LogK values reported by Kuter *et al.*[300] The linear free energy equation reported by these authors to predict the association constant due to π-π interactions with Fe(III)PPIX ($\log K_{\pi-calc}$) incorporates the number of available π electrons (nπ) by $\log K_{\pi-calc} = 0.23 \times n\pi$. Assuming fourteen available π electrons in compound **2c** are able to interact with Fe(III)PPIX (six for each aromatic ring and two for the amide), the $\log K_{\pi-calc}$ is 3.22, remarkably close to the experimentally obtained value. This is despite the fact that the predicted value does not take into account substituent effects on the strength of the interaction, suggesting that they do not play a role in this interaction. In addition, the fact that the dibenzamides **5a** and **5f** do not have significantly larger logK values than the monobenzamide **2c**, suggests that association does not take place over the entire molecule, which would result in a $\log K_{\pi-calc}$ of ~5 for the 22 π electrons in the dibenzamide.

Table 4.5 Association constants for the interaction of selected benzamides and quinolines with Fe(III)PPIX.

Compound	logK	Binding constant ratio relative to **2c** (K/K_{2c})
 **2c**	3.20 ± 0.36	1
 **5a**	3.23 ± 0.11	1.1
 **5f**	3.92 ± 0.21	5.24
QD	5.26 ± 0.01	115
CQ	5.52 ± 0.03[a]	209
4-amino-7-chloroquinoline	4.43 ± 0.01[a]	17

[a] Data from Egan *et al.*[182]

It should be noted that these values are calculated by assuming a 1:1 Fe(III)PPIX binding ratio. CQ has recently been shown to have a binding stoichiometry of 1:2 CQ:Fe(III)PPIX in aqueous solution, resulting in a $\log K_{obs}$ of 13.3 ± 0.2.[150] However, here the 1:1 complexes were compared in 40% DMSO and the value reported by Egan *et al.*[182] for CQ in this medium was used.

The concentration of test compound-haem complex (*B-Fe(III)PPIX*) relative to the concentration of free *Fe(III)PPIX* can be approximated by *Eq. 4.4* using the equilibrium binding constant (K) and the free compound concentration (B), assuming the latter remains constant and is equal to the concentration added.

$$B + Fe(III)PPIX \quad \overset{K}{\rightleftharpoons} \quad B - Fe(III)PPIX$$

$$K = \frac{[B-Fe(III)PPIX]}{[B][Fe(III)PPIX]}$$

$$K \times [B] = \frac{[B-Fe(III)PPIX]}{[Fe(III)PPIX]} \qquad \textbf{Eq. 4.4}$$

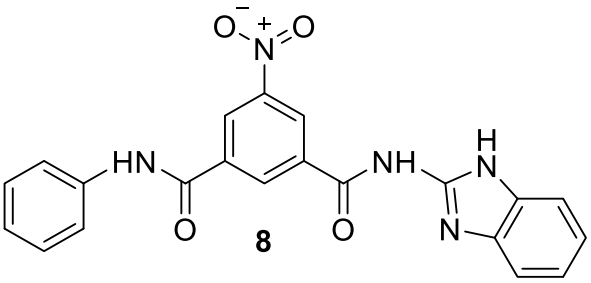
For CQ, the intracellular concentration of unbound compound at the IC_{50} is on the order of 1 mM, based on the parasite activity and CAR. This indicates that at least 99% of the free haem is bound to CQ in the parasite for a logK of 5.5. In contrast, the amount of haem bound to a benzamide compound is ~60% for a logK of 3.2.

The decreased ratio of *B-Fe(III)PPIX* to *Fe(III)PPIX* for the benzamide **5a** as a result of the 100-fold decrease in logK compared with CQ, may suggest a connection between haem binding and accumulation, since a similar 100-fold decrease was measured in the CARs.

In this case the decreased parasite activity of 4-amino-7-chloroquinoline relative to its basic derivatives, CQ and AQ (*Figure 4.14*), may relate to decreased accumulation arising from the 10-fold weaker haem binding (logK of 4.4), in addition to the lack of pH trapping.

In order to test this theory on the benzamide scaffold, an analogue expected to π-π interact with haem more effectively, was proposed for synthesis. This required conversion of the R_2 pyridyl rings into fused aromatic rings in order to increase the number of π electrons available for interaction. Since it was intended to avoid the inclusion of a quinoline moiety and because a variety of benzimidazoles were shown to be good βH inhibitors in the HTS, a derivative where R_2 was benzimidazole was suggested. This group was predicted to add a log unit to the $logK_{\pi-calc}$ value since it contains ten π electrons, as opposed to the six of the pyridyl group. Owing to synthetic problems relating to the physical properties of a compound with two benzimidazole moieties, the molecule had to be desymmetrised to an analogue containing just one benzimidazole unit (*Figure 4.21*). Although this compound (**8**) could be synthesised and isolated, its solubility in aqueous solution and most organic solvents was poor and it could not be used in haem-binding experiments. In addition, the

activities of compound **8** were found to be uninteresting. It was a moderate inhibitor of βH formation with an IC_{50} of 68 μM but displayed no parasite activity up to the maximum dose concentration of 249 μM. It was concluded that adding more fused rings does not necessarily improve parasite activity, although this needs further investigation.



8

Figure 4.21 The desymmetrised benzimidazole **8** containing a benzimidazole for an anticipated increase in the π-π interactions with haem compared with **5a**.

4.3 Summary and conclusions

A variety of purchased or synthesised benzamides were evaluated for their ability to inhibit βH formation and prevent parasite growth. The activities were related to their molecular features in order to determine their minimum structural requirements and SARs.

In general, the benzamides were only able to inhibit crystal growth when the substituents on the ring A (that adjacent to the amide carbonyl group) were meta to each other. For the dibenzamides, ring B (R_2) could be varied but only those compounds where R_1 was a planar electron withdrawing group showed βH inhibition activity with both pyridyl and phenyl R_2 groups. The observed SARs indicated that only predominantly planar molecules were able to inhibit βH formation and that this most likely takes place by interaction with the fastest growing face of the crystal, in agreement with previous studies on 4-aminoquinolines.[154,113,155] The dibenzamides were identified to have better biological activity than the corresponding mono or tri-benzamides and the pyridyl derivatives had the lowest IC_{50}s. Only one compound, where R_1 was varied to *t*-butyl and R_2 to 3-pyridyl, showed better parasite activity than the parent hit compound. However, this activity was not comparable with clinical haemozoin-inhibiting antimalarial drugs and attempts to improve the activity were unsuccessful. Nonetheless, the benzamides investigated here were useful as mechanistic probes. The basicity of the benzamide had no effect on the parasite activity or CAR, however, their 100-fold weaker Fe(III)PPIX binding constants compared with those of the 4-aminoquinonlines, indicated that haem association may have some influence on parasite accumulation and activity. Although lower than that of CQ and AQ, the experimentally determined CAR values for the benzamides were 1000-fold larger than that expected if accumulation was due to pH trapping alone. This strongly suggests that haem binding or binding to the fastest growing face of the haemozoin crystal may be involved in the accumulation of these compounds.

CHAPTER FIVE

TRIARYLIMIDAZOLES: STRUCTURE-ACTIVITY RELATIONSHIPS FOR β-HAEMATIN INHIBITION AND PARASITE ACTIVITY

5.1 Introduction

Seven triarylimidazoles were found to inhibit parasite growth in the HTS discussed in *Section 2.2.4.3*, with three showing IC_{50} values below 1 μM. Interestingly all seven were very good βH inhibitors at least twice as active as CQ, but parasite activity decreased significantly upon variation of the 3,5-dimethoxy-4-hydroxy substituents on ring A (*Figure 5.1*). This prompted investigation into the SARs for this chemotype, especially given its novelty as a potential class of antimalarials.

a) b)

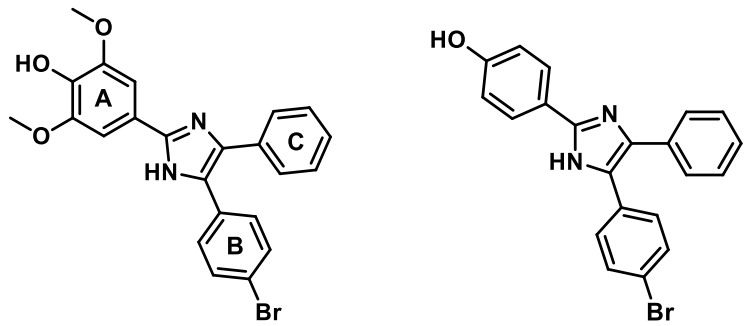


Figure 5.1 Examples of parasite-active hits found in the HTS; a) a representative 3,5-dimethoxy-4-hydroxyphenyl substituted triarylimidazole (**VU0132562** IC_{50} of 0.31 μM) and b) a 4-hydroxy substituted triarylimidazole (**VU0099289,** IC_{50} of 9.1 μM).

5.1.1 Specific goals

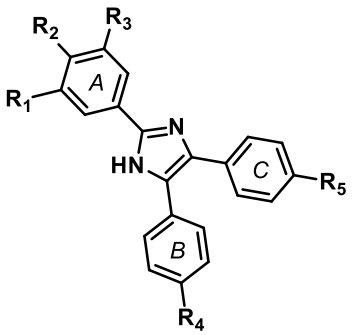
The specific goals for Chapter 5 were to:

- Test the purchased or synthesised simple triarylimidazole derivatives to determine minimum structural requirements for βH inhibition activity.
- Use the synthesised compounds containing substituents possessing a range of electronic, lipophilic and spatial properties to evaluate the βH inhibition activity for this series in the NP-40 assay.
- Determine the inhibition of parasite growth activity of the derivatives using the LDH malaria parasite assay.
- Identify SARs.
- Validate the biological target within pRBCs.

5.2 Results and discussion

5.2.1 β-Haematin inhibition activity for the triarylimidazoles

As in the case of the benzamides, the triarylimidazole scaffold was initially deconstructed in order to determine the minimum structural features for βH activity. Ring B and C in the triarylimidazole HTS hits appeared to have a less important influence on βH activity than ring A. As a result, rings B and C were removed (*Figure 5.2*) to evaluate their necessity. The monoarylimidazole, 2-phenylimidazole **P11** was obtained commercially and tested in the NP-40 assay described in *Section 1.8.5 to 1.8.6*. The compound was found to be inactive against βH formation. Additional data on the triarylimidazole chemotype, which was made available via the collaboration between UCT and Vanderbilt University (VU), indicated that 2,4,5-triphenylimidazole (**VU_IMR-1-101**) was also unable to inhibit βH formation, however, substituting ring A with pyridyl resulted in the simplest active compound (**VU_SC1_1D**) with an IC_{50} of 25.36 µM (*Table 5.1*). The next evaluated derivative was the purchased compound 4-(4,5-diphenyl-1H-imidazol-2-yl)phenol (**P12**), which was identified as the simplest triphenylimidazole to inhibit βH formation (IC_{50} of 17.7 µM). This established that at the minimum active scaffold required at least one substituent on ring A. In order to further investigate the SARs for this scaffold, derivatives of the parent compound (*Figure 5.1a*) were synthesised as shown in *Figure 5.2*.



15a: R_1 = OMe; R_2 = H; R_3 = OMe; R_4 = H; R_5 = H
15b: R_1 = OMe; R_2 = OMe; R_3 = OMe; R_4 = H; R_5 = H
15c: R_1 = OMe; R_2 = OH; R_3 = OMe; R_4 = H; R_5 = H
15d: R_1 = H; R_2 = OCF_3; R_3 = H; R_4 = H; R_5 = H
15e: R_1 = OMe; R_2 = H; R_3 = OMe; R_4 = Cl; R_5 = H
15f: R_1 = OMe; R_2 = H; R_3 = OMe; R_4 = OMe; R_5 = OMe
15g: R_1 = H; R_2 = NO_2; R_3 = H; R_4 = H; R_5 = H

16: R_1 = H; R_2 = NH_2; R_3 = H; R_4 = H; R_5 = H

Figure 5.2 The synthesised triarylimidazole derivatives

Compound **15c** resembled the parent compound with the exception of the bromo substituent on ring B. This derivative maintained potent activity against βH formation confirming that the substituents on ring B and C were not vital for inhibition of βH. The substituents on ring A were then varied by replacing the 4-hydroxy group with a hydrogen atom (**15a**) or methoxy group (**15b**). These analogues also maintained activity against βH formation. This was surprising given that **VU_IMR-1-101** was inactive, however it demonstrated that the substituents on ring A were important for βH activity. The data for the 4-trifluoromethoxy analogue (**15d**) further suggested that only specific substituents give rise to βH inhibition, since the compound was inactive. To evaluate whether this was an electronic effect, a nitro and an amino derivative were synthesised and assayed, both of which displayed good βH activity. This showed that compounds with both electron withdrawing and releasing groups on ring A were capable of inhibiting crystal growth. The data also demonstrated that both planar and non-planar substituents were tolerated. However, the predicted logD at pH 4.8 for **15d** was ~6.8, significantly higher than the active derivatives with logD values ranging from ~4.2 to ~5.1. The large hydrophobicity of **15d** may have altered the orientation of the interaction between the triarylimidazole and Fe(III)PPIX compared to the active analogues, resulting in a weaker interaction and the inability to prevent crystal growth. Finally, analogues of **15a**, which contained either a 4-chloro atom on ring B (**15e**) or 4-methoxy groups on both ring B and C (**15f**) were evaluated in order to measure the extent of ring B and C influence on βH activity. These compounds displayed similar activity to **15a**, with IC_{50} values of 16.6 μM and 13.8 μM for the chloro and dimethoxy analogues respectively. This also demonstrated that both electron withdrawing or releasing and planar or non-planar substituents were tolerated on rings B and C. The βH activity data for the triarylimidazoles are shown in *Table 5.1*.

Table 5.1 βH IC_{50} values for the triarylimidazole scaffold.

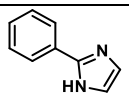
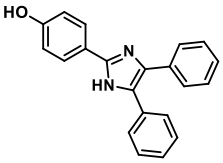
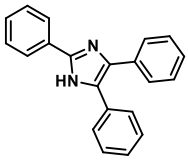
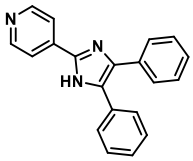
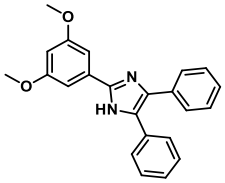
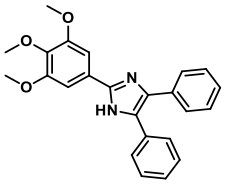
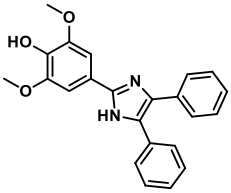
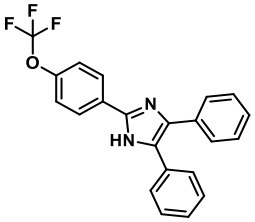
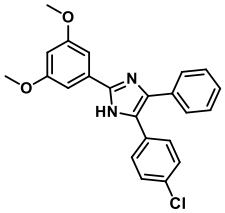
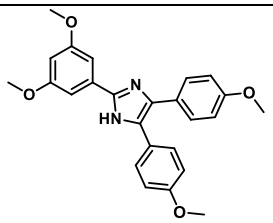
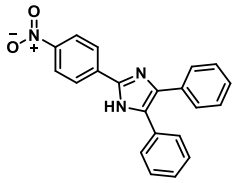
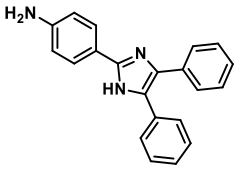
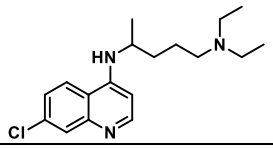
Compound	Structure	Source	βH inhibition IC_{50} (μM)
P11		Commercial	>1000
P12		Commercial	17.7 ± 0.5
VU_IMR-1-101		Vanderbilt University	>1000
VU_SC1_1D		Vanderbilt University	25.4 ± ND
15a		Synthesised	14.6 ± 0.2
15b		Synthesised	10.9 ± 0.3
15c		Synthesised	10.2 ± ND
15d		Synthesised	>1000
15e		Synthesised	16.6 ± 0.6

Table 5.1 (Continued)

Compound	Structure	Source	βH inhibition IC_{50} (μM)
15f		Synthesised	13.8 ± 0.4
15g		Synthesised	19 ± 1
16		Synthesised	16.1 ± 0.5
CQ		Commercial Standard	31.5 ± 0.5

5.2.2 Parasite growth inhibition activity

The 2,4,5-triphenylimidazole from VU which was not a βH inhibitor also showed no parasite activity up to 20 μM, while the 2-pyridyl-4,5-diphenylimidazole which was a potent βH inhibitor showed weak activity of 10.76 μM in the CQ-sensitive D6 strain. The purchased and synthesised triarylimidazoles were tested in the LDH parasite growth inhibition assay using the CQ sensitive NF54 strain (*Table 5.2*). Compound **P12** was found to be inactive at concentrations up to 32 μM, despite displaying potent βH activity. Other derivatives which performed poorly in the assay were the 4-nitro (**15g**) and 4-amino (**16**) analogues of **P12**. This showed that having substituents only in the *para* position of ring A was not sufficient for parasite activity. What was surprising however, was the IC_{50} of 4.19 μM for **15d**, the trifluoromethoxy analogue, which was identified as inactive against βH formation, but which inhibited parasite growth with moderate activity. This suggested that this compound was able to kill the parasite via an interaction with an alternate target. This might be attributable to the high lipophilicity of the molecule, which would allow it to interact favourably with parasite membrane proteins. Nonetheless, this result appeared to be specific to **15d** and not triarylimidazoles in general, since the other non βH inhibitor (**VU_IMR-1-101**) was inactive against the parasite. For the βH inhibiting compounds containing 4-hydroxy and 3,5-

methoxy groups on ring A, **15a-c**, the parasite activity improved from 6.98 to 1.3 μM with variation of the *para* substituent H < OMe < OH, illustrating that the ring A substitution pattern of the parent compound (**VU0132562,** *Figure 5.1a*) was superior for activity. Since the most active triarylimidazoles in the HTS, with submicromolar activity, contained a *para* Br or OH substituent on the ring B (*Table 2.5*), the effect of these groups, while not a significant factor for βH inhibition, appeared to increase parasite activity. This prompted the synthesis of **15e** and **15f** in an effort to determine whether the activity of triarylimidazoles with other ring A substituents could also be improved via adding ring B and ring C substituents. Owing to the availability of starting materials for synthesis, the 4-chloro and 4,4'-dimethoxy analogues were tested and found to improve the activity from 6.98 μM for **15a** to ~1.7 μM for both **15e** and **15f**. This 4-fold activity improvement for the introduction of ring B and/or C substituents was almost identical to that of **15c** vs **VU0132562** (*Figure 5.1a*). This was also true when comparing **P12** with poor activity (>32 μM) to its 4-bromo derivative (**VU0099289**, *Table 2.5*) from the HTS with IC_{50} of 9.1 μM. However, overall the activity of the synthesised compounds could not be improved relative to the HTS hits, but the study revealed the specificity for parasite growth inhibition of the 4-hydroxy-3,5-methoxy substituents on ring A. This is summarised in *Figure 5.3.*

Table 5.2 βH inhibition and *P. falciparum* IC_{50} values for the triarylimidazoles.

Compound	Source	βH inhibition IC_{50} (μM)	NF54 IC_{50} (μM)
P11	Commercial	>1000	ND
P12	Commercial	17.7 ± 0.5	>32
VU_IMR-1-101	Vanderbilt University	>1000	>20*
VU_SC1_1D	Vanderbilt University	25.4 ± ND	10.8 ± ND*
15a	Synthesised	14.6 ± 0.2	7.0 ± 0.1
15b	Synthesised	10.9 ± 0.3	3 ± 2
15c	Synthesised	10.2 ± ND	1.3 ± 0.5
15d	Synthesised	>1000	4.2 ± 0.1
15e	Synthesised	16.6 ± 0.6	1.7 ± 0.5
15f	Synthesised	13.8 ± 0.4	1.7 ± 0.4
15g	Synthesised	19 ± 1	15 ± 4
16	Synthesised	16.1 ± 0.5	42 ± 16
CQ	Commercial Standard	31.5 ± 0.5	0.008 ± 0.002

* Activity values from Vanderbilt University in the CQ-sensitive D6 strain.

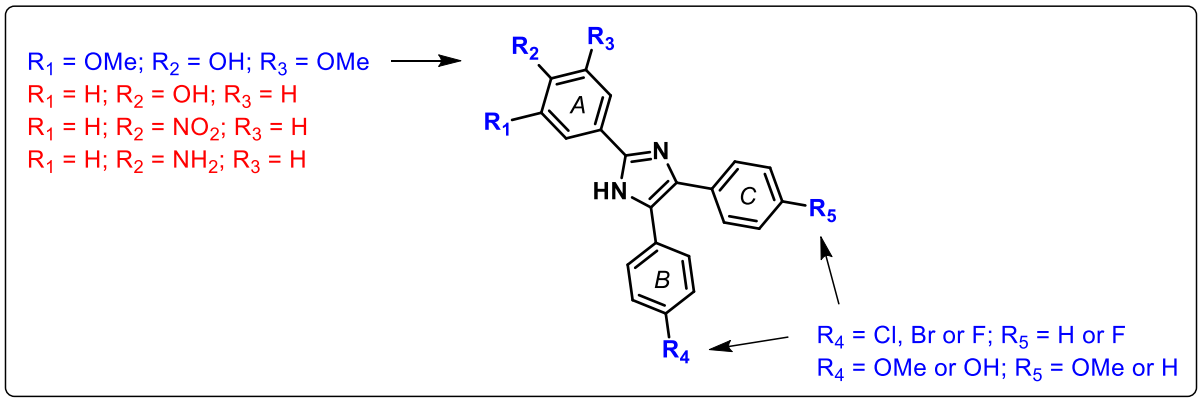


Figure 5.3 The favourable (blue) and unfavourable (red) substituents for parasite activity within the triarylimidazole HTS hits and synthesised analogues.

5.2.3 Target validation for the triarylimidazoles

Compound **15d** was found to have a moderate parasite IC_{50}, even though it did not appear to inhibit βH formation in the NP-40 assay. This result led to the need for confirmation of haemozoin inhibition within parasitised red cells. While the triarylimidazole **VU0099289** (*Table 2.5*) from the HTS was shown to cause a 43% increase in free haem, data for additional derivatives was required in order to be confident of the biological mechanism of action for this scaffold. Compounds **15a** and **15f** were sent for cell fractionation studies and both were shown to cause a very large increase in free haem in the cell (>40 fg/cell concentration corresponding to >50% free haem) at 2.5 times the relevant IC_{50} of the compound (*Figure 5.4*). This study confirmed that the triarylimidazoles inhibit haemozoin formation in the malaria parasite in a dose dependent manner that coincides with decreased parasite survival. This strongly suggests that inhibition of haemozoin formation is the mechanism of action of these compounds.

a) b)

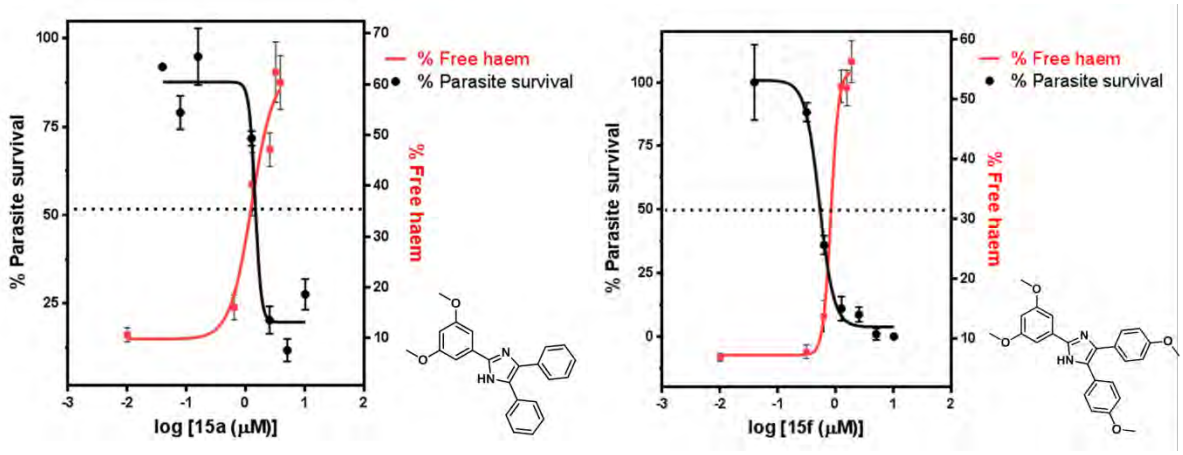


Figure 5.4 The concentration dose response curves for the amount of free haem (fg/cell) and the percent parasite survival which cross close to their respective IC_{50} values for compounds a) **15a** and b) **15f**.

5.2.4 Correlation analysis for triarylimidazole SARs

The ability of βH inhibiting derivatives with a greater number of hydroxyl and methoxy substituents to inhibit parasite growth can be seen graphically in *Figure 5.5.* This had a much greater impact on the parasite activity than the βH inhibition IC_{50}. Indeed, the relationship with βH activity was not statistically significant.

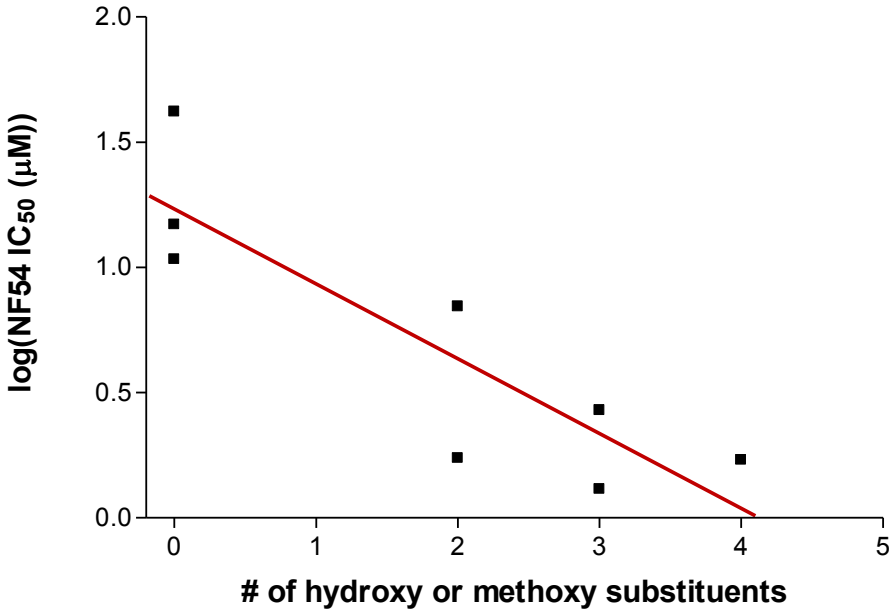


Figure 5.5 A linear correlation between the log of the parasite activity with the number of hydroxyl or methoxy substituents on the phenyl rings in a triarylimidazole given by $\log(NF54IC_{50}) = -0.30(\#\text{ of OH or OMe}) + 1.23$; ($r^2 = 0.76$, $P = 0.0048$).

Correlations between the parasite activity and other physiochemical properties were identified. The most statistically relevant correlation with the least number of parameters was that of the inverse of the NF54 IC_{50} with the water solubility (H_2O-Sol) and the molecular weight (MW) of the compound (*Figure 5.6*). The correlation equation revealed that the MW parameter had 2.5 times more weighting on the activity than that of the H_2O-Sol. This relationship showed that an increase in the H_2O-Sol as well as an increase in MW was favourable for parasite activity. This was reasonable in light of the finding regarding the need for more hydroxy and methoxy groups for improved *in vitro* antimalarial activity. Analogues with the 3,5-dimethoxy-4-hydroxy substitution pattern possessed a larger MW and were predicted to be more aqueous soluble than those without these groups. Furthermore, **15e** and **15f**, the more potent derivatives of **15a**, had larger MWs since they contained ring B and C substituents. It is important to note that a triarylimidazole with a larger MW and H_2O-Sol will probably not necessarily give rise to better activity, but rather that an increase in these features is able to explain the observed SARs for this set of data.

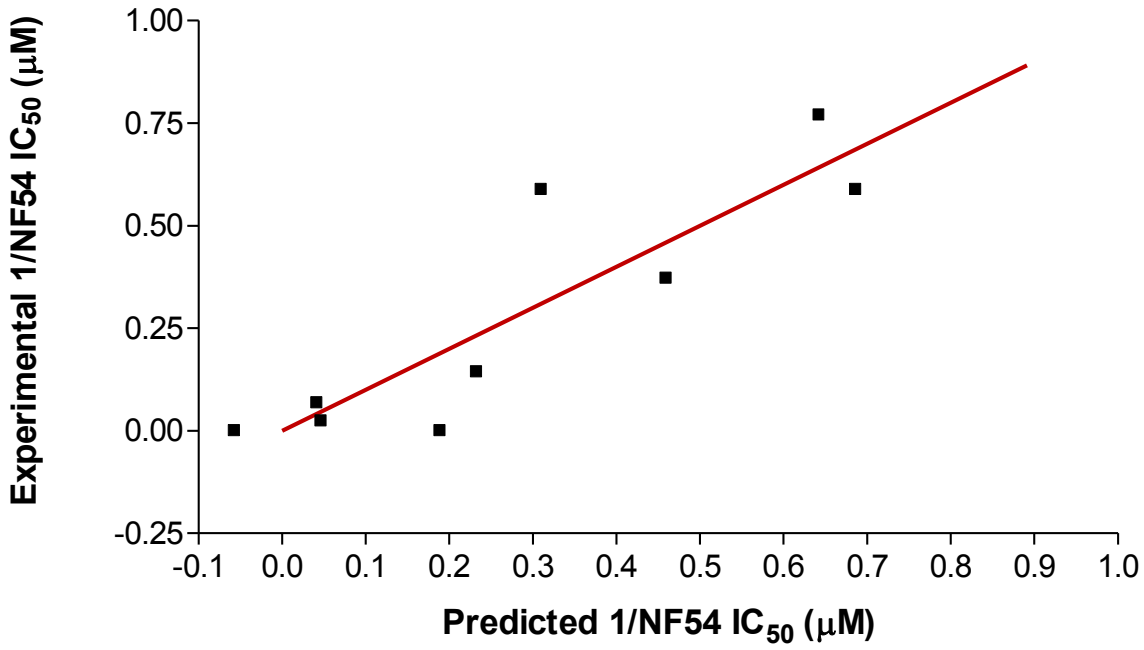


Figure 5.6 Multiple correlation analysis for the inverse of NF54 parasite activity with physiochemical properties revealed a statistically significant correlation as $1/NF54IC_{50}$ = $0.284(H_2O\text{-Sol.}) + 0.008(MW) - 1.404$; ($r^2 = 0.78$; P = 0.0016); statistically significant for both individual parameters (t = 3.10, 4.09, 3.77 > t-crit = 2.45) and overall correlation (F-stat = 10.67 > F-crit = 5.14) at the 95% confidence level.

5.3 Summary and conclusions

A variety of purchased or synthesised triarylimidazoles were evaluated for their ability to inhibit βH formation and prevent parasite growth. The activities were related to their molecular features in order to determine the minimum structural requirements and SARs.

The triarylimidazole series was much more challenging to investigate compared to the benzamide series since the trends relating to βH inhibition activity could not be explained and parasite activity relied on very specific substituents. This showed that the scaffold was not amenable to structural changes. In addition, there did not appear to be a correlation between βH inhibition and parasite activity, however, a correlation may have become apparent had the relative accumulation ratios for the triarylimidazoles been taken into account. Measurement of these CARs was beyond the scope of this project, but may be important in future if investigation of triarylimidazoles is continued. However, the scaffold was generally very aqueous insoluble, which is undesirable for medicinal compounds, and as a result, haem binding measurements in 40% aqueous DMSO could not be carried out. Correlation analysis revealed that both larger MW and aqueous solubility improved the activity of this series. The hydrophilic nature of the 3,5-methoxy-4-hydroxy substitution pattern may explain the improved solubility and therefore, the greater parasite activity of these derivatives.

Although these compounds do not possess appropriately high activities for further development as antimalarials, they are important for enriching the structure-activity data sets, particularly for use in training sets to develop predictive models for both βH inhibition and parasite activity. To this end, the structure-activity data for the triarylimidazole derivatives together with the benzamides from Chapter Four can be incorporated into training sets to create probability models based on Bayesian statistics, in order to predict the probability of activity for new test compounds.

CHAPTER SIX

BAYESIAN MODELS FOR PREDICTING β-HAEMATIN INHIBITION AND ANTIMALARIAL ACTIVITY

6.1 Introduction

The primary goal of HTS is the identification of validated hits that have potential to become chemical leads in drug discovery programmes. Determining which of these compounds has the appropriate chemical characteristics as well as pharmacodynamic and pharmacokinetic properties, is a resource and time-intensive task.[301] The increasing popularity of HTS as a starting point for drug discovery has led to a surge in the availability of both phenotypic and target-based activity data. Realising the maximum value of these data remains a challenge, since hit compounds with only moderate activity are often not further investigated. In addition, the usefulness of the data for the non-hits or negative data is not always appreciated. Large quantities of structure-activity data from HTS projects, including the negative data (non-hits), are underutilised as a result of this. One way to make use of all the screening results is by analysing the data in parallel by employing *in silico* data mining algorithms. These machine learning techniques not only help with data interpretation, but can also be used for predicting the activities of new compounds.[302] Methods that allow such data to be used for virtual screening have potential to maximise efficiency and reduce costs. Currently, this work is the only one reported in the literature[303] demonstrating Bayesian probability applied to antimalarial activity prediction. This is despite the availability of published *P. falciparum* activity datasets, including whole-cell screens from GSK (the TCAMS library) and the St Jude Children's Research Hospital.[171,172]

6.1.1 Bayesian probability

6.1.1.1 Bayes' theorem

In drug discovery, Bayesian probability provides a formal mathematical method for combining prior structure-activity data with current data to create a model for analysis of structurally important features and prediction of future data.[197] Bayesian statistics is based

on Bayes' Theorem (*Eq. 6.1*), published first in 1763 by Thomas Bayes in "An Essay Towards Solving a Problem in the Doctrine of Chances".[304,190]

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)} \qquad \textbf{Eq. 6.1}$$

Where, in the context of drug discovery, $p(h|d)$ is the probability that a molecular feature (d) contributes to activity (h) in a test molecule, $p(d|h)$ is the prior probability of the feature being present in an active compound in the training set, $p(h)$ is the probability of any compound in the training set being active and $p(d)$ is the probability of the feature being present in any molecule in the training set.

6.1.1.2 Bayesian classification in Discovery Studio

The Bayesian classification method for modelling structure-activity data in the Accelrys Discovery Studio software[305] uses the knowledge of prior data in the training set to quantify the probability of an external compound in a test set being active or inactive. The probability can then be updated in light of new evidence, which is subsequently added to the training set. Bayesian statistics also takes into consideration the complexity of the model, automatically identifying the simplest model to explain the observed data and hence, avoiding overfitting.[200]

In the Bayesian modelling component in Discovery Studio,[305] which was employed in this project, learned models are created when the user marks active sample data (chemical structures) as "good" at a chosen activity cut-off. The system then learns to distinguish the features of those active compounds from the background data (inactive or "bad"). These features are input descriptors including logP, MW, #HBD, #HBA, #RB, #AR, #R, FPSA and extended-connectivity fingerprints of depth 6 (ECFP_6). The learning process collects the frequency of occurrence of each feature in the good subset as well as in all the data samples, then assigns a weighting to each feature using a Laplacian-adjusted probability estimate (*Section 6.1.2.3*). This accounts for the different sampling frequencies of the individual features. Upon creation of a model from a particular training set with a user-defined activity cut-off, a receiver operating characteristic (ROC) score is determined. The ROC score indicates the performance of the model and is calculated from the area under the curve of the true positive (TP) rate versus the false positive (FP) rate (ROC curve). To

determine the TP rate, samples are excluded consecutively from the model and then their activities predicted using the leave-one-out model. These are compared to the real user-defined activities and each sample assigned as a TP or FP. A ROC score of 1 indicates a perfect model, predicting all molecules correctly, while that of 0.5 indicates no enrichment from random activity assignments.

The probability of activity for an external molecule X is calculated from the training set model by generating the features of the molecule X and calculating the Laplacian adjusted weight for each feature. Then, after normalisation, the weights are summed to provide a probability estimate, which is a relative predictor of the likelihood of molecule X being from the good subset, indicated by a more positive value.

6.1.1.3 The Laplacian-adjusted probability estimate

In order to account for the fact that not every generated feature has an influence or equal weighting on the activity of the sample data, a Laplacian-corrected estimator is used to adjust the uncorrected probability estimate. The derivation is given by means of an example from Discovery Studio.[305]

Assuming that there are N molecules in the training set of which M are active, the estimated baseline probability of activity for a molecule X from the test set is P(Active) = M/N. However, if A of those N molecules contained a feature F (*Section 6.1.1.2*) and only B of those A molecules were active, then the uncorrected estimate of activity for a molecule with the same feature F would be P(Active|F) = B/A. However, if there are too few molecules in the training set with feature F then the estimator is less reliable. For example, if only one molecule in the training set contains feature F, and that molecule is active, then the calculated P(Active|F) of 1 is overconfident for that feature. This is due to inadequate sampling of the feature. On the other hand, in most cases the vast majority of the features do not correlate with activity and if this is the case for feature F, P(Active|F) is equal to the baseline probability P(Active).

If an additional D molecules containing feature F are added to the original training set then the number of new molecules expected to be active is P(Active)*D. Having included these samples, the corrected estimate of the probability for activity is shown in *Eq. 6.2.*

$$P_{corr}(Active|F) = \frac{(B + P(Active)*D)}{A+D} \quad \textbf{Eq 6.2}$$

This corrected probability stabilises the estimator so that as A (and therefore B) approaches zero, the probability for feature F converges to P(Active), as expected for most features. Lastly, the probability estimator is normalised to a relative final probability estimate by dividing by P(Active) and then taking the log to give *Eq. 6.3.*

$$log\ P_{final}(Active|F) = log\ \frac{P_{corr}(Active|F)}{P(Active)} \quad \textbf{Eq 6.3}$$

These $logP_{final}$ values for each feature are summed to predict the complete probability of a molecule being active. For most features, $logP_{final}$ will be approximately zero, however, for those features which are more common in active molecules, $logP_{final}$ will contribute positively. Conversely, for those features more common in the inactives, the $logP_{final}$ will be negative.

6.1.2 Available antimalarial datasets

The green-fluorescence based assay used in the whole-cell screens by GSK and the St Jude Children's Research Hospital is phenotypic, not target specific and as a result, the active compounds cover a range of chemical and physical properties, depending on the mechanism of action of that molecule. Ekins *et al.*[200,201] used Mtb whole-cell data for model generation. Currently, their studies appear to be the only available references to Bayesian modelling of multimodal distributions. Other studies apply Bayesian probabilities to specific targets. The extent to which inhibitors of different antimalarial targets differ in chemical space is not fully understood, however it is reasonable to hypothesise that Bayesian models may perform better using compounds acting on a single target in the training set. Another reason for the lack of published models is the absence of available inactive or negative data. The exception to this is the St Jude's set where the structures and bioactivities for the entire library of 309,474 compounds have been disclosed.

Since 2010, the NP-40 based detergent mediated pyridine ferrichrome assay has been applied in HTS efforts to discover new antimalarial scaffolds. Structure-activity data from a pilot screen[161] on 38,400 compounds as well as a larger screen[204] on 144,330 remaining compounds in the Vanderbilt University (VU) compound library (*Chapter 2*) revealed a total of ~700 novel βH inhibitors, of which ~200 inhibited parasite growth by ≥90% at 23 µM. The NP-40 assay has also been employed to test ~200 parasite active compounds synthesised by the group of Inokuchi and co-workers at Okayama University (OU) in Japan. These compounds are neocryptolepine and isocryptolepine derivatives with long amine side chains for which the data are already publically available (*Section 1.7.3*).[107,108,282,109,110]

In this section of the project, a virtual screening approach for discovering βH formation and *P. falciparum* growth inhibitors has been validated by employing combined HTS data, including published data and as yet unpublished βH screening data for the 13,533 active TCAMS compounds as training sets. Two models were created, the first which uses VU, TCAMS, OU as well as in-house UCT structure-activity data, which includes the benzamide (*Chapter 4*) and triarylimidazole (*Chapter 5*) derivatives, as a training set to predict βH activity. The second model includes the biologically inactive St Jude's compounds in the training set for modelling *P. falciparum* activity. It is important to note that only known βH inhibitors were used as actives in the training set for *in vitro* antimalarial activity. Although the detergent based assay for βH formation does not give definitive proof of the mechanism of whole-cell therapeutic action, it was hypothesized that most of the parasite active molecules used in generating the model would be haemozoin inhibitors, allowing for a single-target training set. For this reason, the St. Jude's actives were excluded since their targets are, for the most part, unknown. Finally, the generated Bayesian models were employed to predict the βH inhibitory and antimalarial activity of 1510 U.S. Food and Drug Administration (FDA) approved drugs as well as purchasable compounds from Sigma-Aldrich's drug-like molecule library (Aldrich^{CPR}). A selection of 34 of these compounds were purchased and tested for both βH inhibition and *in vitro* parasite growth inhibition.[303]

6.1.3 Specific goals

The specific goals for Chapter 6 were to:

- Generate training sets as structure data files (SDFs) using published or internal HTS data.
- Create and optimise a Bayesian model for βH inhibition activity by evaluating the output ROC scores and predictions of test sets, resulting from variation of training set data and activity cut-offs.
- Create and optimise a Bayesian model for parasite growth inhibition activity by evaluating the output ROC scores and predications of test sets, resulting from variation of training set data and activity cut-offs.
- Validate the models by:
  - virtually screening structural data of the FDA approved drugs,
  - assessing the resulting Bayesian scores of the known antimalarial drugs,
  - virtually screening the Sigma-Aldrich drug-like molecule library (Aldrich^{CPR}),
  - purchasing a range of compounds, predicted to be active or inactive, and
  - finally, experimentally determining the accuracy of the Bayesian model predictions using the NP-40 βH inhibition assay and the LDH parasite growth inhibition assay.

6.2 Results and discussion

6.2.1 Optimisation of Bayesian models

6.2.1.1 Model for predicting inhibition of βH formation

It is generally accepted that Bayesian models perform more effectively when the training data cover sufficient chemical space and if meaningful molecular features can be generated.[191] This was tested by creating models based only on the VU data with increasing numbers of negative data points. A group of 339 randomly selected VU compounds was excluded from the training set to be used as a test set and the models were used to predict whether they were active or inactive. Finally, all the βH data from TCAMS, OU and UCT were also added to the training set (*Table 6.1*). As expected, the ROC score increased with increasing numbers of compounds in the training set, resulting in an excellent score of 0.915 (*Figure 6.1*). In addition, the test set containing 339 samples was evaluated. The activity predictions also improved from 78% to 83% correct as VU inactives were added, however the percentage dropped to 77 when compounds from different libraries (covering wider chemical space) were incorporated into the training set.

Table 6.1 Optimisation of βH model using a cut-off of 100 μM, with increasing numbers of inactive compounds from the VU set. The test set consisted of 339 VU samples.

Total βH actives in training set	Total βH inactives in training set	ROC score for model (leave-one-out)	ROC rating for model	VU test set correctly predicted
1,000	1,000	0.876	Good	78%
1,000	6,000	0.901	Excellent	80%
1,000	31,000	0.901	Excellent	83%
1,000[a]	51,000[a]	0.905	Excellent	83%
2,113[b]	64,118[b]	0.915	Excellent	77%

[a] These compounds were prepared at pH 5 in the calculation.

[b] This set incorporated VU, TCAMS, OU and UCT data at pH 5.

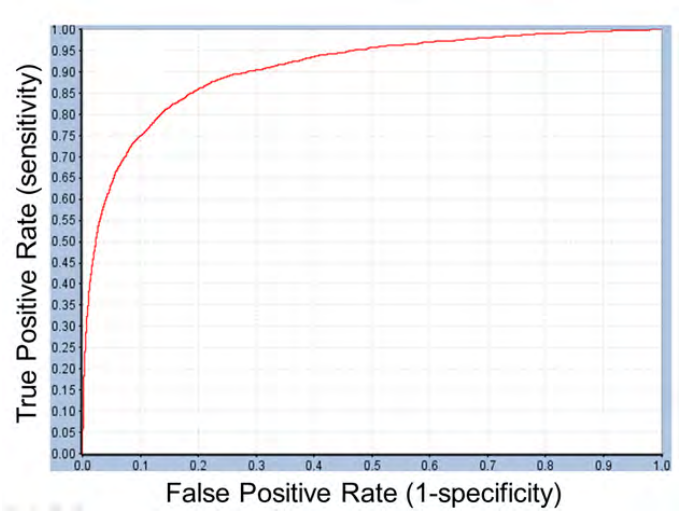


Figure 6.1 ROC curve for βH inhibition activity.

6.2.1.2 Model for predicting parasite activity

An advantage of using Bayesian models over multiple linear correlations, typical in QSAR analysis and prediction, is that the precise IC_{50} becomes less important. The data are divided into an active and inactive set at some appropriate user-defined IC_{50} cut-off and the model built on the frequency of occurrence of particular molecular features. Conversely, traditional QSAR uses the measured IC_{50} value to build a mathematical model, where fairly small differences in experimental procedures can lead to significant effects on the fitted coefficients. This was an important consideration in electing to build a Bayesian model to predict parasite activity, since data were sourced and compared from different compound sets and tested in different laboratories with non-standardised assay procedures. However, this did not appear to disadvantage the models, because for most of the samples, their assignment to the active or inactive set was not dependent on a precise IC_{50} value, since they lay either well below, or well above the cut-off value. Initially, a 2 µM cut-off with only the βH inhibiting TCAMS and VU compounds was used in the training set, where the latter compounds inhibited parasite growth by ≥90% at 23 µM. Then a selection of unpublished in-house UCT compounds were added as well as the VU <90% set in order to increase the numbers of inactive compounds. Finally, the inactive compound set from the St Jude's screen was incorporated for ROC optimisation. A test set consisting of 156 molecules from the TCAMS and VU screening data was used for basic validation of the models with excellent prediction statistics for the 2 µM cut-off model (*Table 6.2* and *Figure 6.2*).

When an additional model was developed with a 0.5 μM cut-off using the same training sets, the test set prediction percentage dropped considerably when the St Jude's inactives were incorporated in the training set (*Table 6.2*). In this case, it was found that 95% of the test set compounds with IC_{50} values <2 μM were classified as active and 88% of compounds with IC_{50} values >2 μM were classified as inactive. This caused many false positive values based on an activity cut-off of 0.5 μM. This was probably as a result of the paucity of active compounds in the training set (only 352 actives out of 42194 compounds, 0.8% of the training set) below the 0.5 μM cut-off. Interestingly, this model predicted 92% of the compounds correctly if active test compounds were reclassified as those with IC_{50} values <2 μM.

Table 6.2 Optimisation of *P. falciparum* activity models using IC_{50} cut-offs of 2 μM and 0.5 μM. The test set contained 156 molecules from the TCAMS and VU libraries. Data in bold are for the optimised model after adding all the data to the training set.

Dataset added to training set	Total parasite actives in training set	Total parasite inactives in the training set	Cut-off IC_{50} (μM)	ROC score for model (leave-one-out)	ROC rating for model	% test set correctly predicted
TCAMS (βH inhibitors only) and VU (>90% inhibition at 23 μM)	351	549	0.5	0.796	FAIR	ND
	806	94	2	0.937	EXCELLENT	85%
As above + VU (<90% inhibition at 23 μM) + UCT	352	790	0.5	0.842	GOOD	70%
	817	325	2	0.959	EXCELLENT	94%
As above + St Jude's (inactives only)	**352**	**42194**	**0.5**	**0.989**	**EXCELLENT**	**55% (at 0.5 μM) / 92% (at 2 μM)**
	817	**41729**	**2**	**0.991**	**EXCELLENT**	**92%**

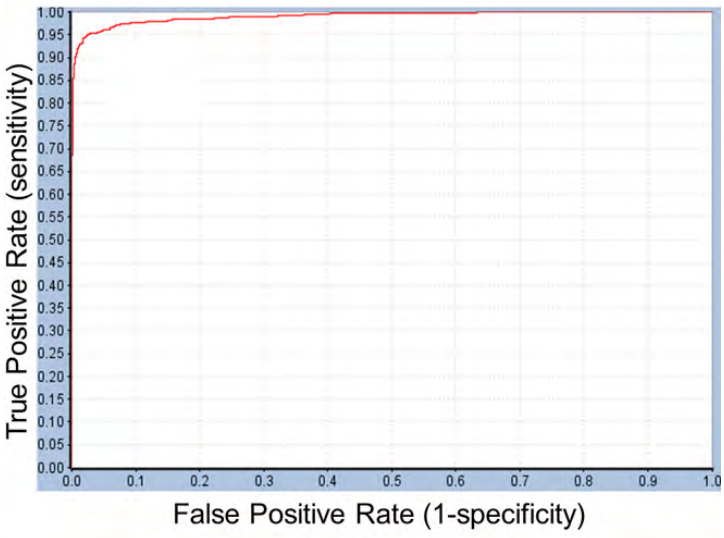


Figure 6.2 ROC curve for whole-cell parasite activity.

6.2.1.3 Comparison of molecular descriptors for βH and parasite activity

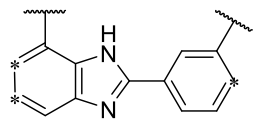
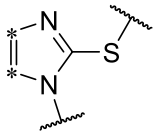
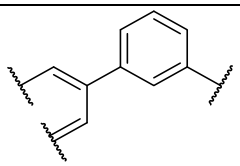
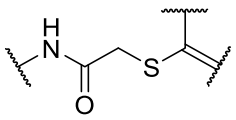
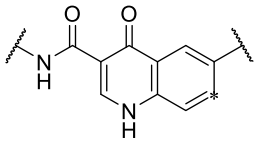
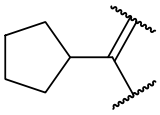
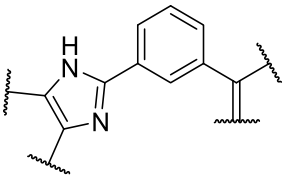
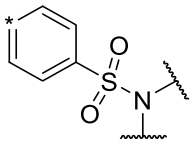
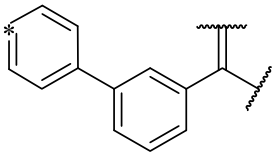
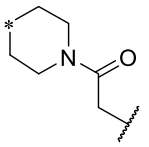
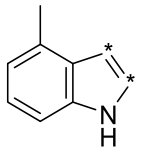
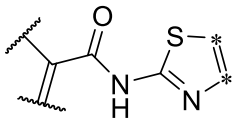
The calculated descriptors which were used to build the models were compared in order to find the optimal ranges for best activity. In *Table 6.3*, the values represent the ranges for each feature which were most favourable amongst the active training set molecules. For the majority of features, ranges are similar for the two models, suggesting that having a higher logP, MW, #HBD, #HBA, #RB and #AR would improve both activities. However, the βH model requires a lower #RB for optimal inhibition, whilst the parasite model requires a much larger #RB. This could be interpreted as a molecule needing to be planar for efficient haem interaction and requiring lipophilic saturated side chains to cross membranes for parasite activity. In addition, the βH model favours a FPSA of 0.29-0.33 whereas the parasite model prefers a lower range of 0.13-0.17. This analysis demonstrates the balance between these two features which needs to be achieved for a βH inhibiting antimalarial. However, both features had a low priority when calculating the Bayesian score, since they were ranked at or near the bottom of the probability weightings for both models. On the other hand, the features which agree between the models are more important in terms of calculating a probability for βH and parasite activity, suggesting that by optimising these features for βH activity, the likelihood of creating a βH inhibiting antimalarial is also improved. In addition to the molecular descriptors in *Table 6.3*, a large number of extended-connectivity fingerprints of depth 6 (ECFP_6) were used to create the model. These were invariably the highest ranking descriptors.

Table 6.3 Optimal feature ranges for the two activity models with the probability score rank for each feature. The word in brackets refers to the range values relative to the less favourable ranges. Those in italics exhibit contrary preferences.

Feature	Rank in βH model	Preferable value/range in βH model	Rank in parasite model	Preferable value/range in parasite model
#HBD	2	3-7 (More)	4	3-8 (More)
#AR	3	4-7 (More)	2	5-8 (More)
MW	4	460-719 (Larger)	5	459-911 (Larger)
logP	5	5-11 (Larger)	7	4-11 (Larger)
#R	6	5 (More)	3	5-8 (More)
#HBA	7	6 (More)	9	7-13 (More)
#RB	*8*	*0-2 (Fewer)*	*6*	*9-20 (More)*
FPSA	*9*	*0.29-0.33 (Larger)*	*8*	*0.13-0.17 (Smaller)*

The good ECFP_6 features for βH inhibition were 2-aryl benzimidazoles, indoles and quinolin-4(1H)-ones, while bad features included 2-thiolimidazoles and a variety of non-aromatic rings and heteroalkyl chains (*Table 6.4*). This result was consistent with the other feature observations which revealed greater #AR and fewer #RB for βH inhibition activity. Similarly, for parasite activity, fingerprints such as imidazoles, benzimidazoles and indoles dominated the good features. Interestingly, almost all the bad ECFP_6 features for parasite activity contained sulfur, either in an alkyl chain, as a sulfonamide or within a 5-membered heteroaromatic ring. A selection of the dominant fingerprints is shown in *Table 6.4*.

Table 6.4 Examples of some of the most important extended-connectivity fingerprints of depth 6 (ECFP_6) for a) the βH inhibition model (cut-off of 100 μM) and b) the parasite growth inhibition model (cut-off of 2 μM).

a) βH inhibition model		b) Parasite growth inhibition model	
Good ECFP_6	**Bad ECFP_6**	**Good ECFP_6**	**Bad ECFP_6**
			
103/115 active	0/683 active	269/276 active	0/3251 active
			
95/121 active	0/495 active	194/194 active	0/2111 active
			
127/176 active	1/982 active	141/141 active	0/2087 active

6.2.2 Chemical space analysis

Spitzmüller *et al*.[306] predicted the target space for the St Jude's and TCAMS whole-cell hit compounds and found over 200 *P. falciparum* hit proteins for 20,000 compounds. However, until recently, data for βH inhibitors has been largely unavailable. Analysis of the principle components (PC) of the TCAMS compounds used in the training set (*Figure 6.3a*) demonstrates a distinct difference in the chemical space between those that inhibit βH formation by at least 60 % (red) at 19 μM and those with <40% inhibition (blue). The βH inhibitors are shifted in space such that they have a greater first and third PC relative to the non-βH inhibitors. This shift corresponds to a lower #RB and larger FPSA, #AR, #R, #HBD, #HBA and MW. There were no significant differences in logP. Furthermore, 1029 assembly fragments were present in the TCAMS actives and only 50 were found to be common between the two sets, resulting in a low similarity score (Tanimoto distance) of 0.0486. This agrees with the Bayesian model comparison method where a large Bayesian distance of 72.3 was found, indicating that a training set consisting of only βH inhibitors may be more

specific towards predicting haemozoin inhibiting antimalarials than one incorporating all 13,533 TCAMS actives. The VU βH hits which were incorporated into the training set cover a more confined chemical space, shifted relative to the TCAMS compounds by lower PC1 (*Figure 6.3b*). This indicates a lower MW and #AR or #R based on feature weighting for PC1. The OU compounds are largely scattered between the TCAMS and VU chemical space with several molecules possessing high MW, larger #R and #RB (long amine side chains) shifted into the higher PC1 range. The analysis demonstrates the relatively large chemical space covered in the training set by βH inhibitors from the different libraries.

a) b)

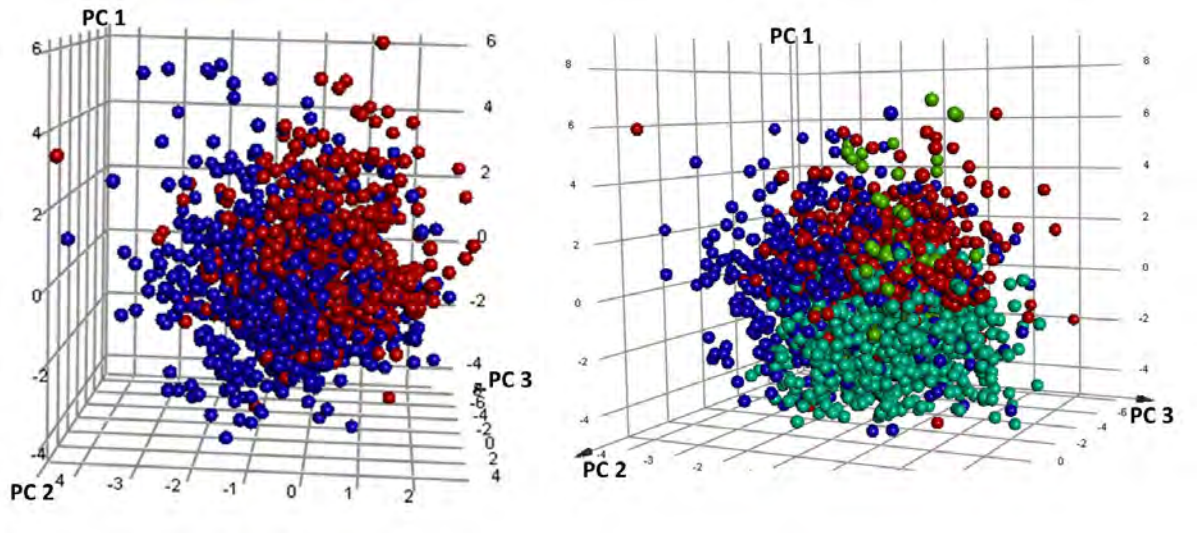


Figure 6.3 a) Plot of the three PCs for the TCAMS βH inhibitors (red-larger PC1 and PC3) vs non-βH inhibitors (blue) and b) The VU (cyan) and OU (lime green) βH inhibitors relative to the TCAMS βH inhibitors (red) and non-βH inhibitors (blue).

With the chemical space of the training set identified, the validation sets were plotted in the same PC space in order to determine how closely related the libraries are (*Figure 6.4*). As expected, the purchasable Aldrich^{CPR} drug-like compounds are contained within the space of the FDA approved drugs which are themselves more dispersed relative to the other sets. The VU compounds lie closest to the known drugs and the TCAMS and OU compounds with the largest PC1 are furthest in space from the validation sets. This diversity is important for training set verification as it demonstrates the degree of model versatility for predicting test sets which differ from the training sets.

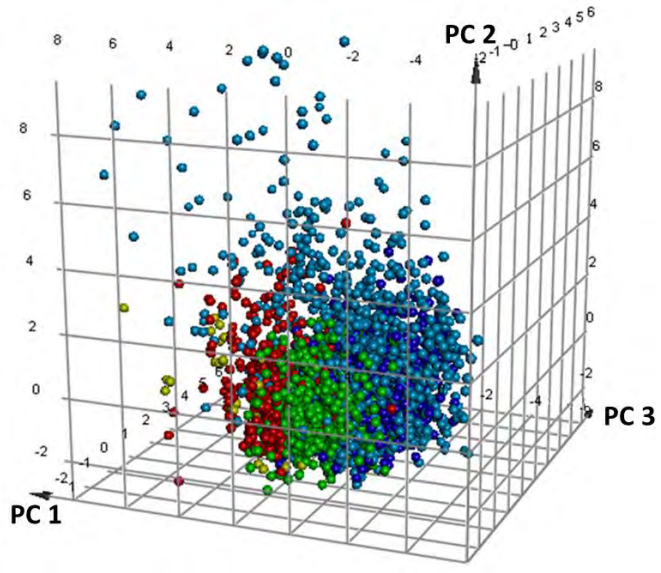


Figure 6.4 Validation sets including the FDA approved (light blue) and 1000 randomly selected molecules from the Aldrich^{CPR} drug-like compounds (dark blue) in relation to the TCAMS (red), OU (gold) and VU (green) βH inhibitors from the training set.

6.2.3 FDA approved compounds for model validation

For the purposes of validation, the optimised Bayesian model was applied to known drugs. Although βH inhibition is mostly unknown for FDA approved compounds, *in vitro* antimalarial activity is often available through PubChem (https://pubchem.ncbi.nlm.nih.gov/), even if they are not clinical antimalarials. The parasite bioactivity was predicted (using the 2 μM cut-off model) for 1510 molecules and ranked by likelihood of being bioactive from highest to lowest Bayesian score (selected portions shown in *Table 6.5*). In the sorted list, all six of the known βH inhibiting drugs that are clinical antimalarials (as well as QD barbiturate and hydroxychloroquine) were found in the top 2.1% of the 1510 compounds. The clinical antimalarials were AQ, QN, QD, CQ, QC and HF. MF was also found within the top 4%. Additionally, of the 24 compounds that are not clinical antimalarials in the top 2.1%, nine have reported *in vitro* antimalarial activity in PubChem, ten have not been tested on *P. falciparum* and only three are reported inactive below 10 μM (*Figure 6.5*). Thus, among the 14 compounds for which antimalarial activity has been reported, this represents a hit rate of 81%. On the other hand, in the bottom 2.1%, 24 are reported as inactive (below 10 μM), seven have not been determined and only two compounds are reported as having *in vitro* antimalarial activity (in one case with contradicting evidence), representing a hit rate of 8% among those tested. However, finding active compounds in the

bottom set does not negate the model since the active training set compounds were specifically βH inhibiting antimalarials. Thus, for example, the clinical antimalarials that are known to have a different mechanism of therapeutic action were also predicted at least 23 Bayesian score units lower than the βH inhibiting ones. These included the antifolates proguanil, pyrimethamine and sulfadoxine;[307] primaquine which causes redox cycling[308], the apicoplast inhibitor, doxycycline;[309] and atovaquone, which disrupts mitochondrial electron transport.[310]

Table 6.5 FDA Approved compounds ordered by parasite activity Bayesian prediction score (2 μM model). The top 2.1% of compounds showing the known βH inhibiting antimalarials (bold), a list of non-βH antimalarials with their scores and the bottom 2.1% are listed.

Rank #	Generic name	Parasite activity: Bayesian prediction score	βH activity: Bayesian prediction score	Antimalarial activity indication from PubChem
		Top ranked 2.1%		
1	Lapatinib	121.907	81.4034	Active
2	**Amodiaquine**	**55.9673**	**4.97018**	**Clinical Antimalarial**
3	Imatinib	55.1797	17.2137	Active
4	Nafarelin	50.2008	-19.0656	ND
5	Nilotinib	48.8765	18.0575	ND
6	Antrafenine	43.3054	-12.1206	ND
7	Vapreotide	41.9603	16.0751	ND
8	Gefitinib	41.7064	15.5785	Active
9	**Quinine**	**40.4845**	**-11.527**	**Clinical Antimalarial**
10	**Quinidine**	**40.4845**	**-11.527**	**Clinical Antimalarial**
11	Quinidine barbiturate	39.6127	-15.6597	Active
12	Rilpivirine	38.4368	15.1667	Active
13	Thiabendazole	37.0463	15.0677	Active
14	Erlotinib	34.2159	14.7457	Inactive <10 μM
15	Gonadorelin	33.8182	-18.7343	ND
16	Goserelin	32.5423	-17.1688	Inactive <10 μM
17	Afatinib	30.9591	11.0776	ND
18	Vandetanib	30.2873	2.8095	Active
19	Lomitapide	29.8482	-15.6564	ND
20	Hydroxychloroquine	29.579	-7.90028	Active
21	**Chloroquine**	**28.9237**	**-7.38423**	**Clinical Antimalarial**
22	**Quinacrine**	**26.9148**	**0.964427**	**Clinical Antimalarial**
23	Terazosin	24.2283	-20.9691	Inactive <10 μM

Table 6.5 (Continued)

Rank #	Generic name	Parasite activity Bayesian prediction score	βH activity Bayesian prediction score	Antimalarial activity indication from PubChem
24	Dasatinib	23.8589	3.44577	Active
25	Pazopanib	23.8397	13.2397	Inactive <10 μM
26	Lansoprazole	23.4241	-1.71826	Active
27	Cetrorelix	22.9426	-26.9078	ND
28	Sorafenib	22.8016	6.40403	Active
29	Regorafenib	22.1709	10.9474	Inactive <10 μM
30	Octreotide	21.8772	0.0651364	ND
31	Avanafil	21.125	-6.335	ND
32	**Halofantrine**	**20.8674**	**3.41694**	**Clinical Antimalarial**
Non-βH inhibiting antimalarials				
317	Primaquine	-2.66276	-8.36974	Clinical Antimalarial
469	Doxycycline	-5.95531	-5.4519	Clinical Antimalarial
718	Pyrimethamine	-10.3092	-10.695	Clinical Antimalarial
873	Sulfadoxine	-13.429	-3.24919	Clinical Antimalarial
964	Proguanil	-15.1309	-6.73212	Clinical Antimalarial
1172	Atovaquone	-19.699	-10.4775	Clinical Antimalarial
Bottom ranked 2.1%				
1479	Drotaverine	-38.3959	-22.9142	ND
1479	Ethopropazine	-38.6606	-27.2162	ND
1480	Cefonicid	-38.7062	-13.6207	ND
1481	Niclosamide	-39.3053	-22.1728	Active
1482	Cisatracurium Besylate	-39.4671	-30.1177	Inactive <10 μM
1483	Ceforanide	-39.6463	-8.66601	ND
1484	Nicardipine	-39.6929	-31.0469	Contradicting Data
1485	Chlorpromazine	-40.0968	-30.5136	Inactive <10 μM
1486	Furazolidone	-41.1267	-25.7797	Inactive <10 μM
1487	Acepromazine	-41.2932	-26.769	Inactive <10 μM
1488	Nilutamide	-42.0274	-14.8976	Inactive <10 μM
1489	Aceprometazine	-42.3818	-32.5787	ND
1490	Diltiazem	-42.5729	-17.633	Inactive <10 μM
1491	Sildenafil	-43.0273	-33.237	Inactive <10 μM
1492	Clonazepam	-43.6032	-24.5517	Inactive <10 μM
1493	Nitrofurantoin	-43.9518	-20.7281	Inactive <10 μM
1494	Nifedipine	-44.2094	-15.5129	Inactive <10 μM
1495	Propiomazine	-44.4947	-15.9686	ND
1496	Clobazam	-45.7832	-20.2821	Inactive <10 μM
1497	Nisoldipine	-46.6286	-34.8889	Inactive <10 μM
1498	Apixaban	-47.7752	-12.0871	Inactive <10 μM

Table 6.5 (Continued)

Rank #	Generic name	Parasite activity Bayesian prediction score	βH activity Bayesian prediction score	Antimalarial activity indication from PubChem
1499	Probenecid	-48.0048	-21.7798	Inactive <10 μM
1500	Cefmetazole	-48.6828	-26.0458	Inactive <10 μM
1501	Entacapone	-49.5917	-17.3079	Inactive <10 μM
1502	Vardenafil	-50.4405	-35.2596	Inactive <10 μM
1503	Nitazoxanide	-52.1584	-6.27099	Inactive <10 μM
1504	Cefazolin	-52.5087	-22.5625	Inactive <10 μM
1505	Dantrolene	-56.3814	-18.768	Inactive <10 μM
1506	Nitrendipine	-56.4016	-25.1557	Inactive <10 μM
1507	Flunitrazepam	-56.5882	-14.6742	ND
1508	Nimodipine	-56.659	-21.5478	Inactive <10 μM
1509	Nilvadipine	-56.8842	-26.4166	Inactive <10 μM
1510	Acenocoumarol	-57.1474	-29.3544	Inactive <10 μM

These activity predictions demonstrate the power of the model for specifically prioritising antimalarials that inhibit βH and are therefore likely to be haemozoin inhibitors. The 0.5 μM cut-off model was also tested and performed even better, finding the βH inhibiting antimalarials within top 1.8% of sorted compounds. Besides primaquine, which went from rank #317 using the 2 μM model to #767 with the 0.5 μM model, only minor changes to the ranking of the other compounds was observed. The βH model was also applied to the FDA compounds. AQ, QC and HF were predicted correctly to be βH inhibitors while QD, QN and CQ were falsely predicted negative. The reason for this is most likely the presence of the quinuclidine or alkyl amino side chains, absent in AQ, which are predicted to be unfavourable for βH inhibitory activity, possibly because they add too many rotatable bonds (*Table 6.3*). Although QC and HF contain alkyl side chains, they also have an extra aromatic ring which contributes favourably to the βH inhibition score. This reveals a shortcoming of the βH model for recognising inhibitors that contain these functional groups.

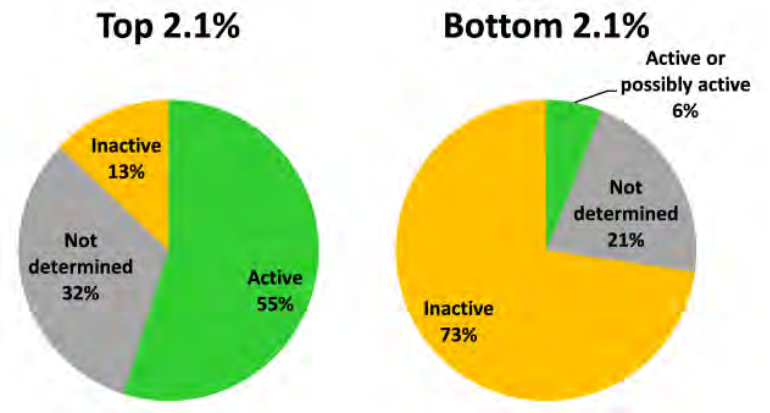


Figure 6.5 Graphical representation of the 2.1% of 1510 FDA approved drugs with highest and lowest Bayesian scores for activity against *P. falciparum* based on training sets of βH inhibiting compounds using activity data against the parasite. Active compounds consist of clinical antimalarials as well as other drugs with proven activity against *P. falciparum* reported in PubChem.

6.2.4 Prioritization of a commercial library

In order to test how the models would perform if used to prioritise compounds for HTS, 4,998 purchasable compounds were virtually screened (*Figure 6.6*). After calculating the Bayesian score using the βH model and selecting only those predicted active, about 900 compounds remained. The top 650 were chosen to screen using the whole-cell Bayesian parasite model (2 μM cut-off) which predicted 178 compounds to be active. Compounds were then selected for purchase according to the criteria specified in the methods (*Section 8.1.5.4*). By applying the βH model and only selecting predicted actives from this dataset, the bio-active and bio-inactive subsets were nested within the βH active subset (*Table 6.6*). This allowed for the most efficient and cost-effective way of validating the model.

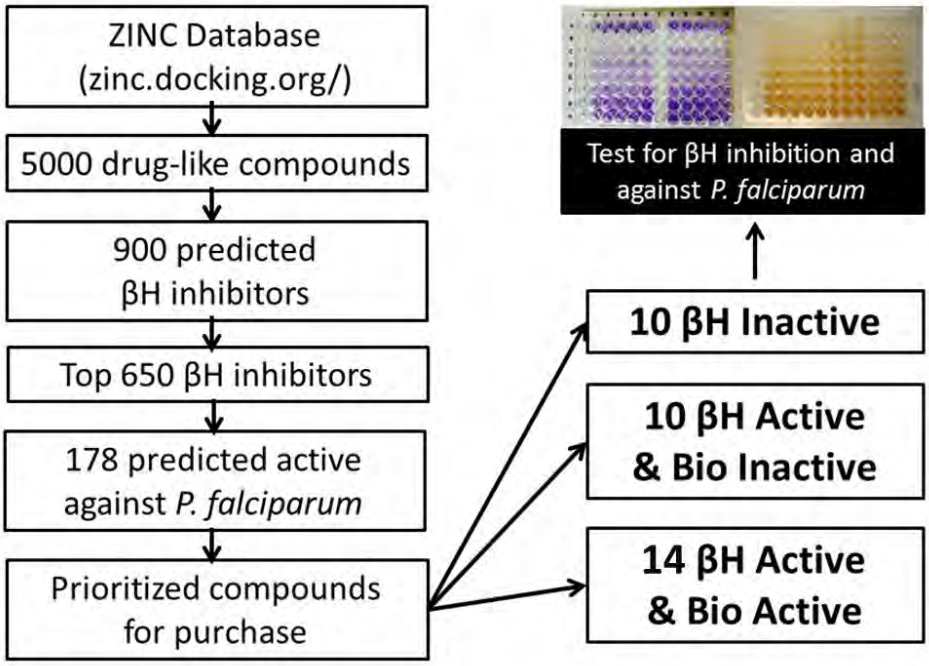


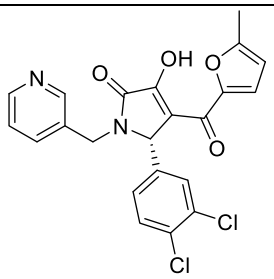
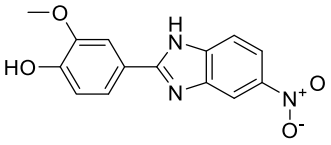
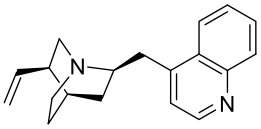
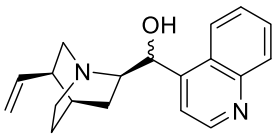
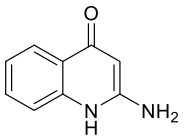
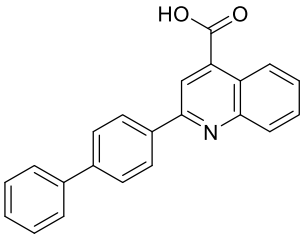
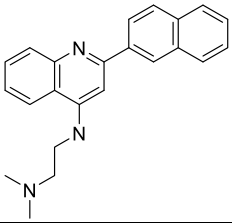
Figure 6.6 Flow chart of the strategy used to purchase and screen a small commercial library.

All of the ten purchased compounds predicted to be inactive against βH formation below 100 μM were indeed found to be inactive, even up to 500 μM. Of the 24 predicted βH inhibiting compounds, six showed IC_{50} values below 100 μM and two between 200-500 μM. This gave a hit rate for βH activity of 25%, a >25-fold enrichment over random screening for βH inhibitors at a cut-off of 100 μM (based on the hit-rate in the VU screen).[204] Of these six compounds, five were also predicted to be bioactive (**SA23**, **SA27**, **SA29**, **SA33** and **SA34**). Two were found to have whole-cell IC_{50} values below 2 μM in the CQ-sensitive NF54 strain of *P. falciparum* (**SA27**: 82 nM and **SA34**: 79 nM). Interestingly, both of these compounds were quinolines bearing structural similarities to known antimalarials QN/QD and CQ respectively, despite the drug molecules being absent from the training set. The QN/QD derivative (**SA27**) was represented in the ZINC database and sold by Sigma-Aldrich as an analogue lacking both the 9-hydroxy and 6-methoxy groups. However, Roepe and co-workers have previously shown the importance of the hydroxyl moiety for βH and parasite activity in QN.[311] This was consistent with crystal structures reported by de Villiers *et al.*[158] which demonstrated coordination of the QN and QD hydroxyl oxygen atoms to the Fe(III) centre of haem, an interaction deemed critical for inhibition of haemozoin formation. In light of this previous research, which seemed to contradict our finding, nuclear magnetic

resonance (NMR) and mass spectrometry (MS) experiments were carried out in order to confirm the structure of **SA27**. As anticipated, the spectra confirmed that the major structure of the QN/QD derivative (>90%) did indeed possess the hydroxyl moiety after all, revealing that **SA27** is either cinchonine (CN) or cinchonidine (CD). Furthermore, there was evidence of an impurity (<10%) containing the quinoline methoxy group, possibly corresponding to QN or QD (See *Figure 6.7* for spectra). This finding reemphasises the importance of proving the composition of hit compounds obtained from HTS, particularly those in milligram quantities for which structural and purity information is often not supplied by the distributor. Additional proof of identity was evident from the parasite activity (NF54) for **SA27** of 82 nM found in this study which corresponds closely to the activity of CN against the Dd2 strain.[311] We retrospectively calculated the Bayesian probability score for the actual structure of **SA27** with the hydroxyl group in place (*Table 6.6*) which showed that CN/CD was in fact more likely to be active both against βH formation (Bayesian score of -0.91 vs -3.58) and parasite growth (19.19 vs 5.74) than the 9-dehydroxy derivative.

As expected, none of the other compounds were potent parasite growth inhibitors, either because they were predicted bio-inactive or because they were false positives for βH inhibition activity. Since two of the 14 testable bioactive compounds were actually active, the hit rate was 14%, a 140-fold enrichment over random screening for compounds targeting haemozoin inhibition (based on the hit rate from the VU screen for βH inhibiting antimalarials).[204] However, in a HTS protocol for discovery of βH actives, only this subset would have been tested against the parasite, resulting in a 33% hit rate.

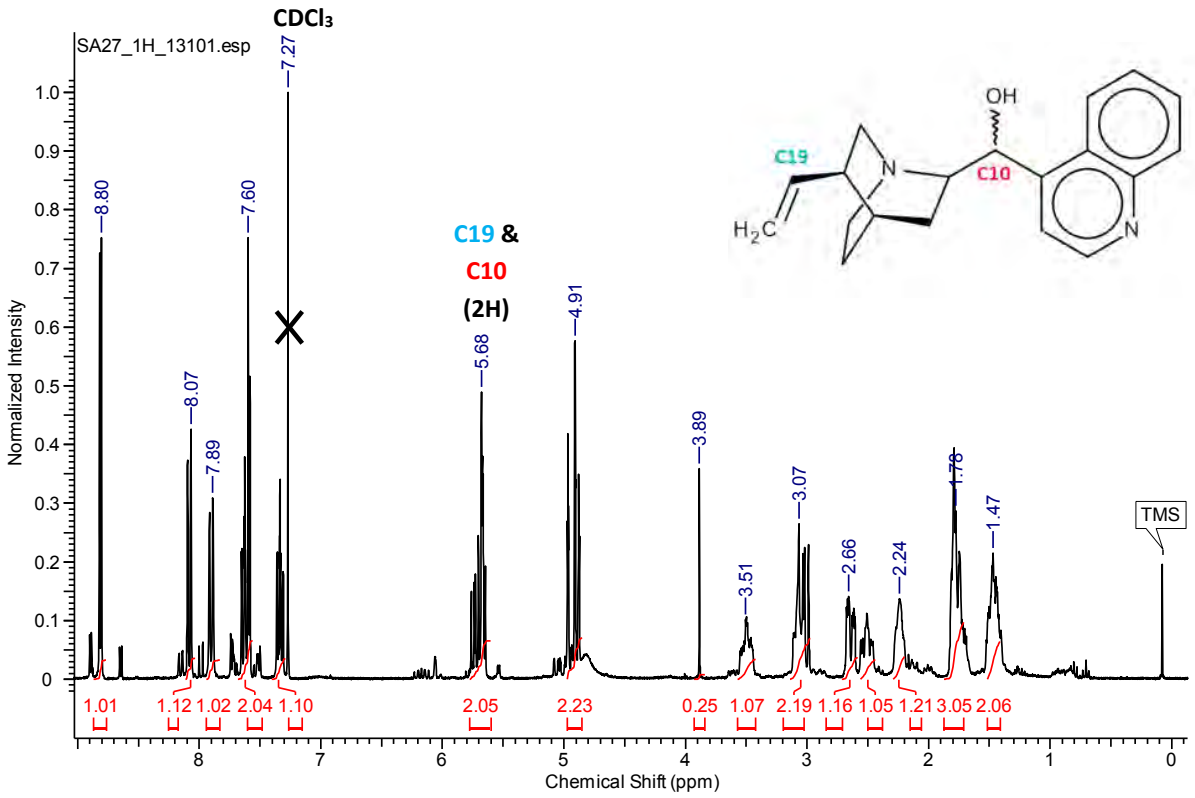
Table 6.6 Experimentally active βH inhibitors from the purchased Aldrich^{CPR} compounds, prioritised using Bayesian models.

Compound code	Structure	βH activity Bayesian prediction score	Parasite activity Bayesian prediction score	Measured βH IC_{50} (μM)	Measured NF54 IC_{50} (μM)
SA18		23.0983	-29.0591[a] -24.8557[b]	88	>10
SA23		43.3948	17.5493[a] 21.2006[b]	22	>10
SA27 Original structure from ZINC (top)		-3.58476	5.73788[a] 5.74298[b]	66	0.082 ± 0.014
Corrected Structure confirmed by NMR and MS (bottom)		-0.91	18.6400[a] 19.1900[b]		
SA29		18.4158	4.76898[a] 54.6788[b]	85	>10
SA33		-2.52201	1.07767[a] 3.6050[b]	46	>10
SA34		11.5085	-0.553378[a] 1.13127[b]	26	0.079 ± 0.026

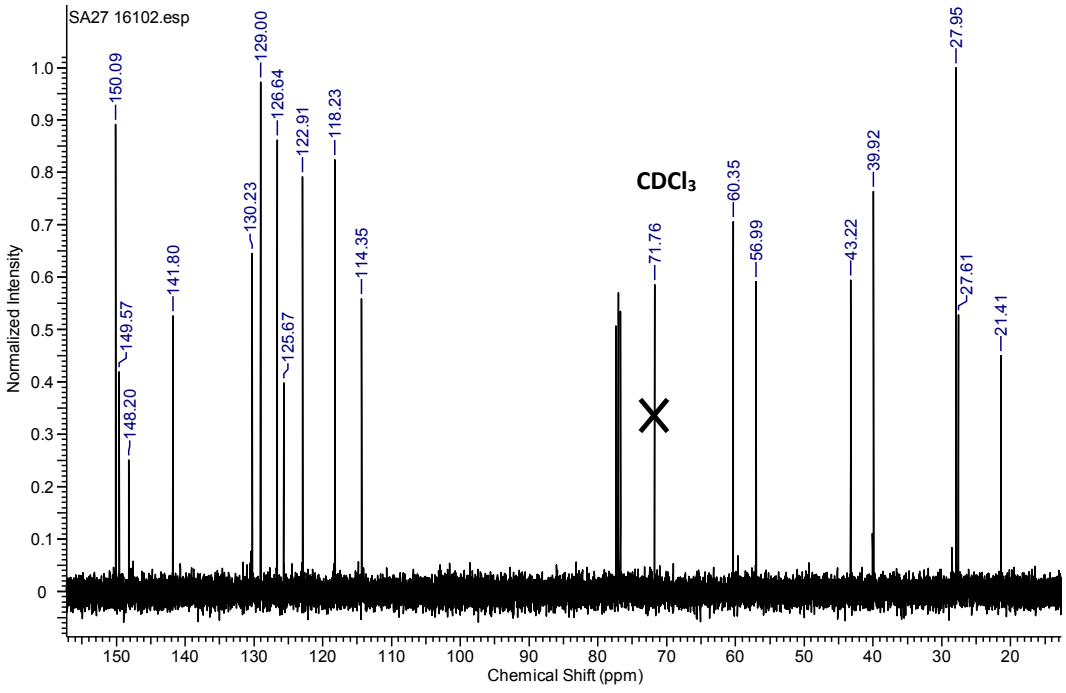
[a]0.5 μM cut-off model

[b]2 μM cut-off model

a)


CDCl₃
SA27_1H_13101.esp
Normalized Intensity
Chemical Shift (ppm)
C19 &
C10
(2H)
OH
C19
C10
N
N
H₂C
TMS
8.80
8.07
7.89
7.60
7.27
5.68
4.91
3.89
3.51
3.07
2.66
2.24
1.78
1.47
1.01
1.12
1.02
2.04
1.10
2.05
2.23
0.25
1.07
2.19
1.16
1.05
1.21
3.05
2.06

b)


SA27 16102.esp
Normalized Intensity
Chemical Shift (ppm)
CDCl₃
150.09
149.57
148.20
141.80
130.23
129.00
126.64
125.67
122.91
118.23
114.35
71.76
60.35
56.99
43.22
39.92
27.95
27.61
21.41

c)

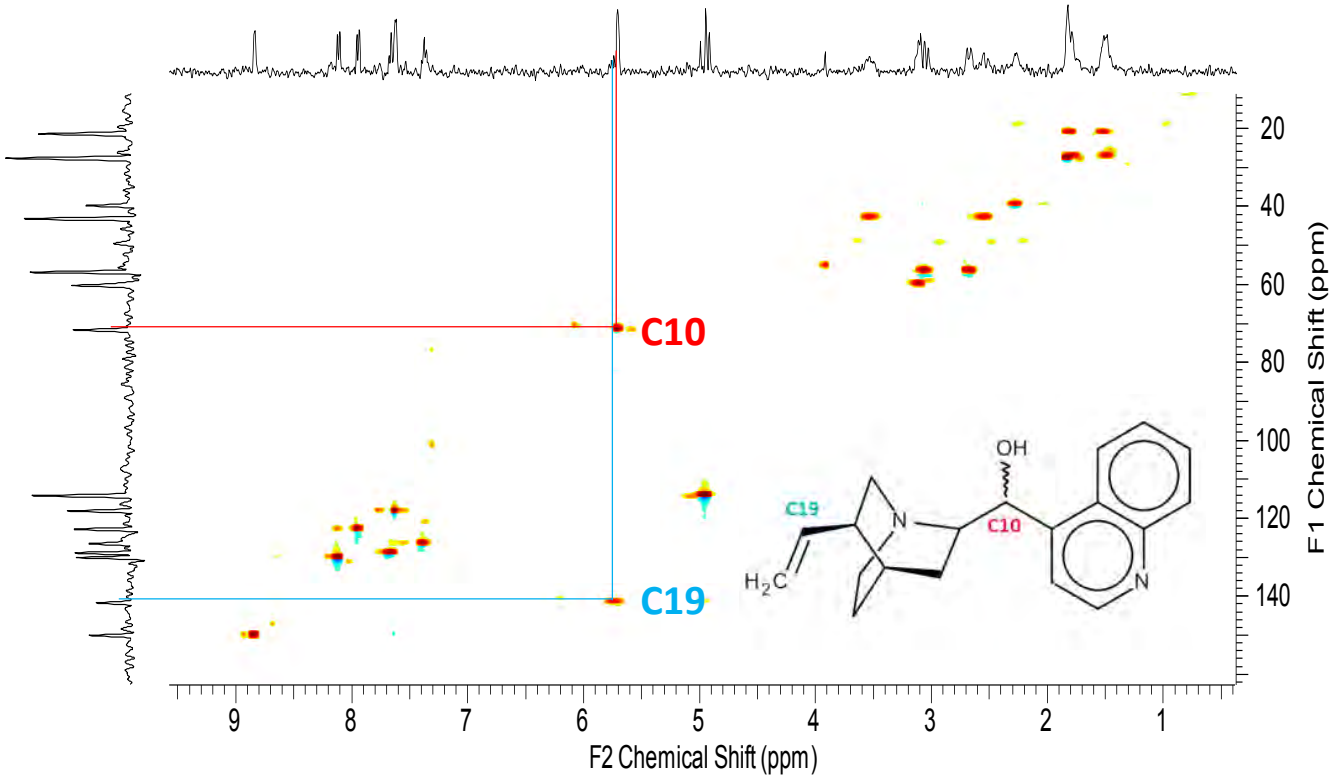


Figure 6.7 a) ^{1}H NMR spectrum of **SA27**. The protons attached to C19 and C10 overlap in the ^{1}H NMR spectrum at 5.7 ppm integrating for 2H. b) ^{13}C NMR spectrum of **SA27** and c) ^{1}H and ^{13}C two-dimensional heteronuclear correlation HSQC spectrum. The proton peak at 5.7 ppm couples to carbon C19 at 142 ppm and C10 at 71 ppm. The downfield chemical shift of C10 relative to that expected for an aliphatic carbon nucleus is a result of the electron withdrawing effects of the hydroxyl substituent. NMR spectra were collected using a Bruker Ultrashield 400 Plus (at 399.95 MHz for ^{1}H and 100.64 MHz for ^{13}C) spectrometer in deuterated chloroform ($CDCl_3$).

The parasite model built on βH inhibitors successfully identified the haemozoin inhibiting FDA approved antimalarials without first filtering out those predicted to be inactive against βH formation. The same approach was retrospectively applied to the Aldrich^{CPR} compounds to compare the procedures for prioritisation. At a cut-off of 0.5 μM (which performed best when predicting the FDA compounds), there were few differences in the prioritised compounds that would have been purchased for bioactivity (aside from small changes in the

order of their probability scores). In fact, an additional derivative of QD (a 6-ethoxyquinoline with a quinuclidine ethyl instead of a 6-methoxyquinoline with a quinuclidine vinyl group) appeared in the top three predicted bioactives which had previously been filtered out as a non-βH inhibitor. Overall, 13 of the 14 molecules purchased were also predicted active when filtering only for bioactivity.

6.3 Summary and conclusions

In this study both *in vitro* antimalarial activity and inhibitory data for βH formation, obtained from the VU HTS, synthesised analogues and the publically available GSK (TCAMS) and St Jude Children's Research Hospital HTS data sets, were used to develop Bayesian models for inhibitors of βH formation and *in vitro* antimalarial activity. The models were optimised by evaluating different activity cut-off values with increasing sample numbers in the training set. The optimised models exhibited excellent ROC scores of 0.915 and 0.991 for the βH inhibition model (cut-off of 100 μM) and parasite growth inhibition model (cut-off of 2 μM) respectively. These models were used to screen two *in silico* compound libraries. In the first, the 1510 U.S. Food and Drug Administration (FDA) approved drugs available on PubChem were ranked from highest to lowest Bayesian score based on a training set of βH inhibiting compounds active against *P. falciparum* that did not include any of the clinical antimalarials or close analogues. The six known clinical antimalarials from the FDA list that inhibit βH formation were ranked in the top 2.1% of compounds. Furthermore, the *in vitro* antimalarial hit-rate for this prioritised set of compounds was found to be 81% in the case of the subset where activity data are available in PubChem. In the second, a library of ~5000 commercially available compounds (AldrichCPR) was virtually screened for ability to inhibit βH formation and then for *in vitro* antimalarial activity. A selection of 34 compounds was purchased and tested, of which 24 were predicted to be βH inhibitors. The hit rate for inhibition of βH formation was found to be 25% and a third of these were active against *P. falciparum*, corresponding to enrichments estimated at about 25-fold and 140-fold relative to random screening, respectively.

Although HTS has become a vital tool for drug discovery, it remains an expensive, time consuming and extremely specialised process. The implementation of machine learning approaches such as Bayesian modelling is able to prioritise compounds for HTS resulting in improved hit rates using fewer test compounds. Creating *in silico* models from data contributed by previous HTS efforts is not only useful for future HTS prioritisation, but is an effective way to make use of all the available data, including inactives, from previous screens. The results of this work have shown that Bayesian models can be applied to antimalarial compounds which are βH inhibitors with impressive enrichment rates relative to random screening. The validation sets strongly suggest that the chemical space of the active compounds in the parasite model is specific for haemozoin inhibiting antimalarials. Bayesian modelling may have a role in the future identification of novel antimalarials, particularly when combined with other *in silico* techniques.

CHAPTER SEVEN

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

The rapid development of drug resistance by *P. falciparum* caused a decline in the efficacy of previously successful antimalarials, particularly CQ, MF and QN, in most parts of the world. Now, the recent reports of artemisinin resistance, which as of early 2015 was confirmed in five countries of the Greater Mekong sub-region in southeast Asia, threatens a vital component of the WHO's top recommended malaria therapeutics, ACTs.[312] This has caused a critical need for novel antimalarial chemotypes and a better understanding of intracellular drug mechanisms towards target-based, rational drug design and development. To this end, HTS has been very successful in identifying numerous compounds, both with known and novel scaffolds that are active against various strains of the malaria parasite. However, owing to the limited capacity for research on only moderately active HTS hits, many novel scaffolds exhibiting some degree of activity are not often exploited, especially by pharmaceutical companies or research groups primarily focused on delivering a new drug. These compounds can be used to probe mechanistic aspects and may have potential to be derivatised into sufficiently potent drugs further down the development line. In addition, many inactive compounds are disregarded or not reported, despite the fact that they add significantly to the structure-activity data pool.

The two types of HTS, namely phenotypic or target-based, each have advantages in certain situations. Phenotypic screening for antimalarials results in identifying all the compounds in the HTS library, regardless of their mechanism or target, which show activity against the parasite directly. This is the best way of finding the most active compounds. However, further studies are necessary to determine the drug target, which often requires multiple experiments. This is very time consuming, expensive and challenging for target validation of multiple hits.[177] On the other hand, in target-based screening, a cheaper and more efficient assay can be carried out first, to identify inhibitors of a specific pathway. Hits from the initial screen can then be tested against the parasite in a secondary whole-cell screen, resulting in

active compounds with a known target.[204] In this case, both the active and inactive data can be conveniently used to create structure-activity relationships or models for predicting activity, since the compounds, acting at the same target site in the parasite, can be directly compared. This is beneficial for establishing the important structural features of novel chemotypes and to aid in the understanding of mechanism.

In this project, target based screening in the form of βH inhibition was carried out on 43,520 compounds, as part of a larger screen of 144,330 compounds. Of the total screened, 530 compounds (0.37%) were confirmed to inhibit βH formation by ≥80%, all exhibiting IC_{50} values below that of CQ. The hits generally possessed a greater logP, MW, #AR, #HBA, #HBD, #R and FPSA than the average for the library as a whole, but preferentially contained fewer RB, which supported the hypothesis that βH inhibitors are planar, aromatic compounds capable of hydrogen bonding and π-π interactions with Fe(III)PPIX, as previously determined for several quinoline haemozoin inhibitors.[157,158] In a secondary screen against *P. falciparum*, 171 compounds were identified as inhibiting parasite growth by ≥90% at 23 μM, showing a favourable hit rate of 32.3% out of the 530 βH inhibitors. This confirmed the relative reliability of the NP-40 based assay and its ability to more accurately mimic the parasite's lipophilic environment through detergent micelles, compared with previous βH inhibition assays.[180] Moreover, 73 and 25 of the hits displayed CQ-sensitive D6 parasite activity below 5 μM and 1 μM respectively, with diverse structures. Fourteen primary scaffolds were identified and representative compounds from twelve of the fourteen were shown to cause a significant increase in free haem in parasites at their respective IC_{50}s relative to the negative control. This showed that indeed, compounds active in the parasite are haemozoin inhibitors as expected from first screening for βH inhibition, giving further validation for the hit identification protocols used in this project.

Two non-quinoline chemotypes, the benzamides and triarylimidazoles were selected for further studies to discover SARs and probe drug mechanisms. Both scaffolds lent themselves well to synthesis of small libraries of derivatives from relatively inexpensive starting materials, owing to the readily-formed amide bond for benzamide synthesis and the condensation cyclisation of the triarylimidazole synthesis.

The benzamides (Scaffold 1) were particularly interesting since SARs for βH inhibition clearly showed that only mono- and di-benzamide analogues capable of adopting planar

conformations, particularly those with electron deficient rings, could inhibit βH formation. Specifically, dibenzamides without an electron withdrawing group on the central ring required pyridyl rings attached to the amide nitrogen for βH inhibition activity, while those with an electron withdrawing group tolerated phenyl rings (*Figure 7.1*). This, together with the relatively weak Fe(III)PPIX binding of representative mono- and di-benzamides supported the hypothesis that haemozoin inhibitors prevent crystal formation via docking into the groove on the fastest growing face of the crystal.

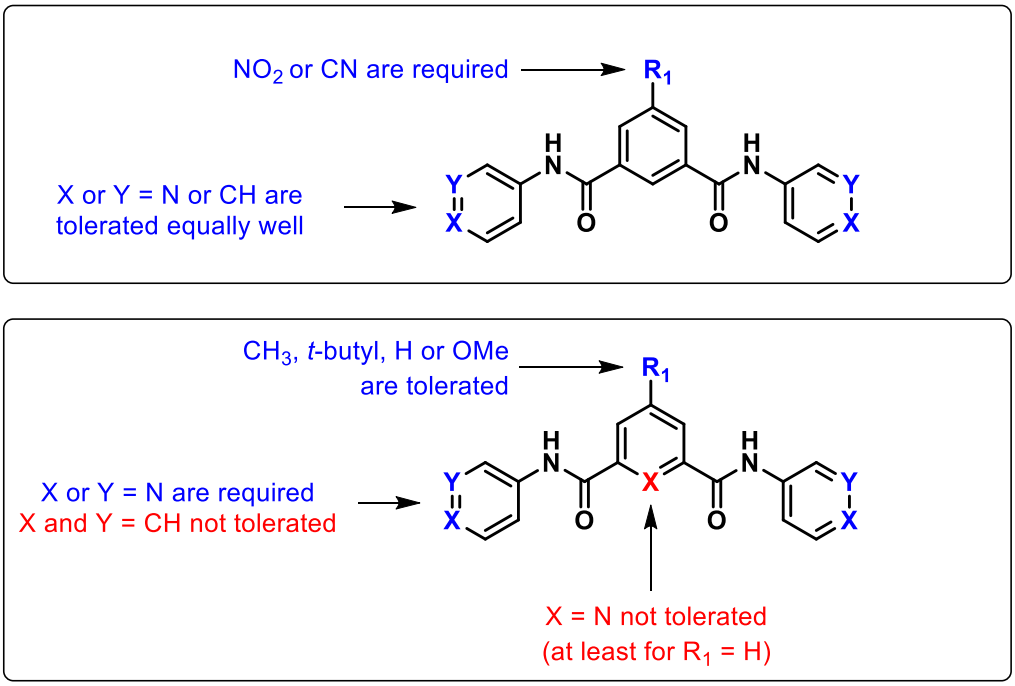


Figure 7.1 βH inhibitory SARs for the dibenzamides.

The *in vitro* parasite IC_{50}s of the benzamides correlated well with their βH inhibition IC_{50}s, apart from one particularly active analogue, which was an outlier. However, their CQ-sensitive NF54 parasite activities were at least 50 times lower than that of CQ and AQ, despite their greater or equipotent βH inhibition activities.

The minimum pharmacophore for βH inhibition in the triarylimidazole series (Scaffold 2) was identified as a triaryl-substituted imidazole where either the 2-aryl ring (ring A) was an unsubstituted pyridyl or a substituted phenyl, although 4-trifluoromethoxy was not tolerated on the phenyl ring. The SARs for βH inhibition could not be related to electronic or hydrophobic effects or to incorporation of planar or non-planar substituents. Hydroxyl, methoxy and halogen substituents were also tolerated on the other phenyl rings (rings B

and C). In fact, substitution generally improved parasite activity (*Figure 7.2*), although not in correlation with βH inhibition activity. The series suffered from poor solubility which made further experiments, such as Fe(III)PPIX association, impossible in 40% aqueous DMSO. The 5-methoxy-4-hydroxy substitution pattern assisted with solubility, which may well explain the significantly improved parasite activity for these analogues.

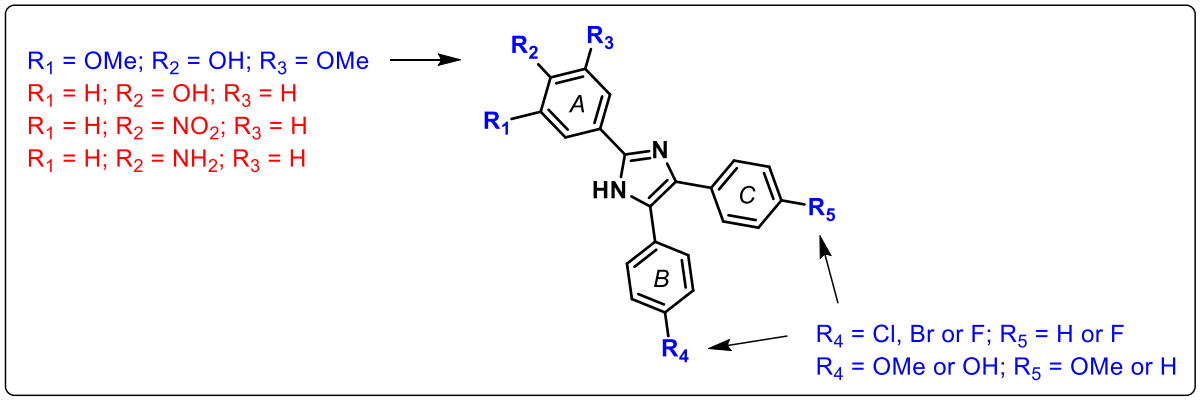


Figure 7.2 The favourable (blue) and unfavourable (red) substituents for parasite activity within the triarylimidazole HTS hits and synthesised analogues.

Representative benzamides and triarylimidazoles caused significantly higher levels of free haem in the parasite at their respective IC_{50}s, relative to the quinoline antimalarials, validating haemozoin inhibition as the target for both series. These results were intriguing, given that both series showed at least 30 to 50-fold lower activity than CQ. The *P. falciparum* cell fractionation experiment to measure free haem levels suggested that a drug-Fe(III)PPIX complex may be responsible for *in vitro* toxicity. One might expect free haem levels to be similar at the IC_{50} of diverse compounds, if free haem alone causes the toxicity which leads to parasite growth inhibition. However, the lower levels of free haem in the quinoline antimalarials compared with the benzamides and triarylimidazoles, suggests that the quinolines complex a large proportion of the free haem, forming a toxic complex. This is reasonable given the strong association between Fe(III)PPIX and the quinolines, especially in the case of CQ and QD. In the case of the benzamides, being ~100 fold weaker Fe(III)PPIX binders, larger amounts of free haem are required in order for parasite cell death to occur. Given that the triarylimidazoles, like the benzamides, are only able to interact with Fe(III)PPIX via π-π interactions, since they do not necessarily possess groups capable of

coordination or hydrogen bonding, this may well also be the case for this series. As a result, higher concentrations of benzamide and triarylimidazole compounds are required to inhibit haemozoin formation to a greater extent in order to release enough free haem for necessary toxicity levels.

The cellular accumulation of several benzamide derivatives was measured. Two pyridylbenzamides showed ~20 fold lower accumulation than CQ and AQ in the D10 strain of the parasite, partially accounting for their 50-fold weaker activity. The phenomenon of pH trapping is believed to account for the large accumulation of CQ and AQ, which contain basic moieties. To this end, a benzamide analogue with basic side chains, predicted to have a VAR of ~147,000, was synthesised with dimethyl amine side chains via reductive amination. Although the compound maintained βH inhibition activity, it displayed a slightly lower IC_{50} of ~1 μM compared with the pyridyl derivatives in the D10 strain and was subsequently shown to accumulate ~6-fold less than the pyridylbenzamides. This suggested that the basic compound could not accumulate via pH trapping in the DV. Alternatively, it was very recently shown that triethylammonium ion and protonated basic amino acids are probably transported out of the DV by *Pf*CRT.[67] In addition, triethylammonium uptake was unaffected by neutral or acidic amino acids or by organic anions, but could be strongly inhibited by verapamil and QD, as well as CQ, and nutrients, such as arginine, lysine, and peptides. Even in wild type CQ-sensitive strains, transport of CQ was observed, although less than in resistant strains. This uptake of protonated basic compounds (BH^+ and BH_2^{2+}) from low to high pH, may well explain the poor accumulation ratios measured for the basic benzamide analogue. It is reasonable to hypothesise that a basic compound, not able to bind strongly to free haem, may accumulate in the acidic compartment via pH trapping but simultaneously be effluxed via *Pf*CRT transport out of the DV. In the case of strong haem binders, such as CQ and related quinolines, the drug-haem complex is either not effluxed out of the DV (at least in the CQ-sensitive strain), allowing for a toxic build-up in the DV, or alternatively, if able to cross the DV membrane, exerts its toxic effect within the parasite cytosol. Both these hypotheses would account for the larger CARs and lower free haem levels measured for CQ and AQ relative to weaker haem binders, such as the benzamides.

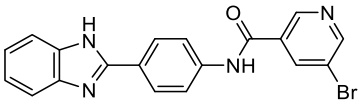
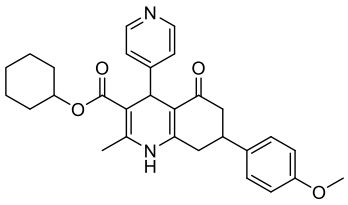
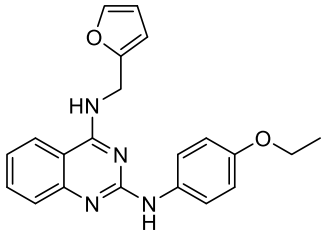
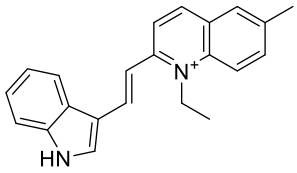
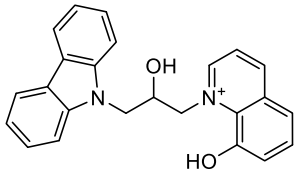
The benzamides and triarylimidazoles have been useful probes of the mechanisms by which antimalarials exert their therapeutic effects. Without target-based HTS, novel chemotypes are unlikely to be discovered. However, HTS is an expensive, time consuming and extremely specialised process, requiring dedicated resources. As a result it is logical to fully exploit the acquired HTS data, particularly if it could simplify future HTS efforts. This has been successfully demonstrated for haemozoin inhibitors identified via HTS, by the implementation of the machine learning technique using Bayesian modelling. Here, *in silico* prioritisation of compounds for HTS resulted in improved hit rates using fewer test compounds. In addition, parameter details of the created models illuminated trends between data sets, specifically βH inhibitors and parasite growth inhibitors. For example, the validation sets strongly suggested that the chemical space of the active compounds in the parasite model is specific for haemozoin inhibiting antimalarials. These approaches are especially important for combating neglected tropical diseases such as malaria, where resources and capacities for research are limited. This finding also raises the intriguing possibility that models trained from compounds with other known targets may be able to be similarly used in antimalarial drug discovery and to identify hits as probes with alternative mechanisms. Furthermore, when combined with *in silico* techniques for prioritising ADMET properties, this approach may have a role in the future identification of novel antimalarials.

In this project, HTS was successful in identifying many novel βH inhibiting scaffolds as well as compounds active against *P. falciparum* via haemozoin inhibition. The scope of the HTS hits as probes for biological mechanisms was revealed after fairly inexpensive and simple synthetic procedures were applied to generate analogues for two selected scaffolds, benzamides and triarylimidazoles. Different characteristics relating to βH and parasite growth inhibition activities, levels of free haem, haem binding affinity and parasite accumulation were revealed for the novel chemotypes compared with the relatively widely-studied quinoline antimalarials, despite also being haemozoin inhibitors. These data were then successfully used to develop a model predicting activity on the basis of Bayesian statistics, demonstrating the value of generating structural diversity via HTS, even when the hits cannot themselves be further developed as antimalarials.

7.2 Future work

The relatively large numbers of haemozoin inhibiting antimalarials identified via HTS in this project, compared with previous efforts,[180,159] has resulted in several novel chemotypes that have future potential to be used as probes in similar experiments to that of the benzamides and triarylimidazoles. *Table 2.3* in Chapter Two presents a list of the 25 most potent compounds against the CQ-sensitive D6 strain of *P. falciparum*, which were hits in the primary βH inhibition assay. Synthesis of analogues and SAR investigations into several of these chemotypes has already been initiated by other members of the Egan/Hunter research group, including the benzimidazoles (**VU0002101**), dihydropyridines (**VU0122653**) and the quinazolines (**VU0129813**) shown in *Table 7.1*, however, their Fe(III)PPIX association constants, free haem levels and CARs are unknown. While many of the other parasite active hit compounds contain quinoline moieties, they also have potential to be used as a way of demonstrating correlations. For example, compounds **VU0098755** and **VU0065708**, exhibiting high potency and containing a charged *N*-alkyl substituted quinoline may be used to probe the relationship between accumulation, free haem levels, haem binding and parasite activity, given that they do not contain basic moieties for pH trapping. Correlations between these parameters may become more obvious by eliminating the factor of pH trapping.

Table 7.1 Selected hits, other than the benzamides and triarylimidazoles, from *Table 2.3* for which investigation has been initiated or which have potential to further probe mechanisms.

Code	Structure	βH IC_{50} (μM)	D6 IC_{50} (μM)
VU0002101		14.3	0.35
VU0122653		24.54	0.5503
VU0129813		26.78	0.396
VU0098755		12.6	0.11
VU0065708		16.2	0.20

For example, if **VU0098755** and **VU0065708** are found to accumulate to a large extent, resulting in their observed good activity, but are weak Fe(III)PPIX binders, this would suggest that accumulation is via another mechanism, possibly active uptake into the DV. In this case, finding trends between different classes of haemozoin inhibitors based on physicochemical or biological properties may be impossible, if each class or derivative possesses its own unique system for therapeutic action. However, this is unlikely since the data in this project suggest a relationship exists between the observed biological parameters, specifically, haem binding affinity, free haem levels, accumulation and parasite activity for the different chemotypes. Uncovering this relationship has great potential for rational drug design. The design of an analogue possessing the correct structural features for coordination, hydrogen-bonding and π-π interactions with Fe(III)PPIX should result in a strong haem binder. With

the assistance of a basic group for initial pH-trapping-driven accumulation of the compound, this would lead to drug-haem complex formation, which could indirectly be measured via free haem levels in cell fractionation experiments. If this hypothesis is correct, saturable accumulation via haem binding will prevent *Pf*CRT efflux of the protonated base and result in improved CARs and hence higher potency against the parasite. Since this hypothesis has been proposed based on only three scaffolds, similar haem-binding, accumulation and cell-fractionation experiments need to be carried out for other classes. In addition, the selected classes should not only be the most potent but also involve moderately active compounds, such as the benzamides and triarylimidazoles, so that potential correlations can be determined.

Subsequently, the new SAR data should be added to the Bayesian training sets, in order to update the statistical models. Once the hits from the initial HTS have been explored, additional screening of new diverse libraries would enrich knowledge of the chemical space of haemozoin inhibitors. Before these libraries are screened using βH inhibition assays, they could be screened *in silico* using the continually-updated Bayesian models for haemozoin inhibiting antimalarials. This could be carried out at varying user-defined cut-offs in order to prioritise the library compounds for screening. Compounds with high Bayesian scores can be screened first to attain the largest hit rates with minimal time and economic resources, while compounds with low scores may be disregarded entirely from the experimental HTS. Indeed, the Bayesian models were 100% accurate at predicting true negatives.

The complexity of *P. falciparum* demands a holistic approach to discovering novel antimalarials. No data should be taken for granted, especially considering the limited resources for research into diseases prevalent in impoverished and developing countries. Successful techniques for directly identifying the most highly potent compounds, such as large-scale phenotypic HTS, are very specialised and expensive, and involve challenging target discovery programmes. In the future, methods for prioritising compounds for screening and thereafter, designing hit derivatives with the appropriate structural features for improved accumulation in the parasite, could have a significant impact on antimalarial drug discovery.

CHAPTER EIGHT

EXPERIMENTAL METHODS

8.1 Physical methods, assays and cell culture protocols

8.1.1 Materials and general procedures

Commercially obtained chemicals (AR or higher grade) were purchased from the suppliers listed in *Table 8.1* and used without further purification.

Table 8.1 Chemical suppliers

CHEMICAL	SUPPLIER
Haemin (≥98%)	Fluka
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
Acetone (AR grade)	Kimix
Pyridine	Sigma-Aldrich
Sodium acetate trihydrate (NaOAc)	Sigma-Aldrich
Glacial acetic acid (AcOH)	Kimix
HEPES buffer	Sigma-Aldrich
Sodium hydroxide (NaOH)	Kimix
Nitric acid (HNO_3)	Kimix
Chloroquine (CQ)	Sigma-Aldrich
Amodiaquine (AQ)	Sigma-Aldrich
Nonidet P-40 (NP-40)	Pierce Biotechnology, Rockford, IL, USA
Phosphate buffer solutions	BDH Laboratory Supplies
Compounds **P1**-**P10**	Vitas-M Laboratory
Compounds **P11** and **P12**	Sigma-Aldrich
Compounds **SA1**-**SA34**	Sigma-Aldrich

The water used in all the experiments was double distilled deionised Millipore® Direct-Q water ($D.H_2O$). All pH measurements were carried out using a Crison 2000 MicropH meter or Jenway 3510 benchtop pH meter, calibrated using standard phosphate buffer solutions at pH 4.00 ± 0.02 and 7.00 ± 0.02 prior to use. Manual micro-volume additions were delivered

using Eppendorf single-channel micropipettes or multichannel micropipettes. All glassware in contact with haematin was washed with 0.2 M NaOH, and then 1 M HNO_3, with H_2O rinsing in between washes. For moderately or poorly soluble compounds, sonication and heating were employed while dissolving solids in the relevant solvent.

8.1.2 High throughput screening methods

8.1.2.1 Manual β-haematin formation assay preparation

Deionised water ($D.H_2O$, 15 μL) was added to all the wells of a 384 well plate followed by the addition of either $D.H_2O$ (5 μL) or AQ (5 μL, 1000 μM) in a checkerboard pattern across the plate. NP40 (5 μL, 305.5 μM) was dispensed first to all AQ wells and then to all $D.H_2O$ wells after which a Fe(III)PPIX suspension an in acetate buffer (1 M, pH 4.8 ± 0.5) was prepared. Haemin (16.3 mg) was dissolved with sonication in 1 ml of DMSO (25 mM) and filtered through a 0.22 μm polyvinylidene fluoride (PVDF) membrane syringe driven filter unit. This 25 mM solution (177.8 μL) was immediately added to 20 ml of the 1 M acetate buffer and vortexed for 10 seconds. The suspension (25 μL) was then swiftly pipetted into all the wells giving a final NP40 concentration of 30.55 μM and AQ concentration of 100 μM. The plate was incubated at 37°C for 5 h followed by the addition of acetone (15 μL) to all wells. A solution of pyridine (50 % v/v) and acetone (20 % v/v) in HEPES (0.2 M) was also added to each well and the plate was left to develop in the dark for 1 h before reading the maximum UV-vis absorbance at 405 nm on a Spectramax 190 plate reader. The data were analysed using Microsoft Excel 2010.[313]

8.1.2.2 Manual concentration dose response experiments

The assay was carried out as in 2.2.1 with the following changes: 12 wells of varying AQ and CQ concentration (0 – 1000 μM final concentration) were repeated 6 times each (no checkerboard) resulting in IC_{50} curves which were analysed used GraphPad Prism v5.0.[314]

8.1.2.3 Verification of drift, Z' and edge effects

An alternating checkerboard pattern of positive and negative control wells were delivered to the first and last two columns of each plate. The positive control contained AQ at the 100% βH inhibitory concentration (IC_{100}). The quality of the assay was validated using the Z' statistical test.[175] Edge effects were analysed using established guidelines of the NIH Chemical Genomics Center.[203]

8.1.2.4 HTS Procedures

The library of test compounds was provided by the Vanderbilt University HTS Facility. The compounds were originally sourced from the commercially available ChemDiv and ChemBridge libraries. Stock compounds (10 mM stock) were delivered in singlet to columns 3 to 22 of ten 384-well plates per day by a Labcyte Echo 550 noncontact liquid handler (19.3 μM final concentration). The Labcyte Echo 550 noncontact liquid handler uses acoustic energy to accurately transfer liquids. Sound waves eject precisely sized droplets from a source onto an inverted microplate. Controls were added to the first and last two columns of every plate. The positive controls contained AQ (10 mM stock, 100 μM final concentration) while negative controls contained the same volume of DMSO without drug. These volumes were added in a checkerboard pattern using the same liquid handler. All solutions were freshly prepared on the day of the experiment. Water, the haematin suspension and the NP-40 solutions were delivered using a Thermo Scientific Multidrop Combi bulk reagent dispenser. The plates were incubated at 37°C for 5 h in a shaking water bath after which the pyridine and acetone solutions were delivered using the bulk reagent dispenser. The plates were left to shake for ~30 min before reading the absorbance values at 405 nm on the SpectraMax 190 plate reader with SoftMax Pro software.[315] The data were analysed using Microsoft Excel.[313] Compounds with absorbance values which corresponded to ≥50% βH inhibition at 19.3 μM were considered active "hits". Dose response curves were determined in duplicate for all the hits using a similar procedure to the single concentration experiment, but without checkerboard controls on the edges. GraphPad Prism v5.0 was employed for IC_{50} generation.[314]

8.1.2.5 *Plasmodium falciparum* SYBR green I-based fluorescence assay

The *P. falciparum* CQ-sensitive D6 strain (Walter Reed Army Institute of Research [WRAIR]/Sierra Leone) and multidrug resistant C235 strain (WRAIR/Thailand) were maintained using a modification of methods described by Trager and Jensen.[70] Culture medium was prepared with RPMI 1640 medium supplemented with 25 mM HEPES, 11 mM glucose, 0.24% sodium bicarbonate, 10% human A (+) plasma (heat-inactivated) and 29 µM hypoxanthine. Cultures were maintained at 5% hematocrit and <5% parasetemia in A (+) blood, which was washed two times with RPMI medium before use. Culture flasks were incubated at 37 °C in a gas mixture of 5% O_2, 5% CO_2 and N_2.

Using modified literature methods,[84] the hits from the NP-40 βH inhibition assay were tested against the D6 and C235 strains of *P. falciparum*. The compounds were dissolved in DMSO (final concentration of 0.23%) and prescreened at 23 µM in 384-well plates with 0.3% starting parasitaemia at 2% haematocrit. Compounds which inhibited >90% of parasite growth were delivered at a range of concentrations from 0 to 23 µM to generate dose-response curves from which the IC_{50} values were determined with GraphPad Prism v5.0.[314]

8.1.3 β-Haematin assays, titrations and physical analysis methods

8.1.3.1 Detergent mediated assay for β-haematin inhibition

The inhibition of βH formation assay method described by Carter *et al.*[160] was modified for manual liquid delivery. Samples were dissolved in DMSO to give 20 mM solutions and 20 µL of each was delivered to wells in the last column (column 12) of a 96-well plate together with distilled water (140 µL) and NP40 detergent (305.5 mM, 40 µL). A solution containing water/NP40 (305.5 mM)/DMSO at a v/v ratio of 70%/20%/10% respectively was prepared and then 100 µL was added to all other wells (columns 1-11). A serial dilution of each compound (100 µL) from column 12 down to column 2 was carried out. Column 1 served as a blank with 0 mM sample. A 25 mM haematin stock solution was prepared by sonicating haemin in DMSO for one minute and then suspending 178 µL of this in a 1 M acetate buffer (20 mL, pH 4.8). The homogenous suspension (100 µL) was then added to the wells to give final buffer and haematin concentrations of 0.5 M and 100 mM respectively. The plate was

covered and incubated at 37 °C for 5-6 h in a water bath or incubator. Analysis was carried out using the pyridine-ferrichrome method developed by Ncokazi and Egan.[103] A solution of 50% (v/v) pyridine, 30% (v/v) H2O, 20% (v/v) acetone and 0.2 M HEPES buffer (pH 7.4) was prepared and 32 mL added to each well to give a final pyridine concentration of 5% (v/v). Acetone (60 mL) was then added to assist with haematin dispersion. The UV-vis absorbance of the plate wells was read on a SpectraMax P340 plate reader. Sigmoidal dose-response curves were fitted to the absorbance data using GraphPad Prism v3.02[314] to obtain a 50% inhibitory concentration (IC_{50}) for each compound.

8.1.3.2 Derivation of the vacuolar accumulation ratio (VAR) equation

The equilibrium expression for a compound with two protonation sites can be written as follows:

$$BH_2^{2+} \overset{K_{a1}}{\rightleftharpoons} BH^+ + H^+$$

$$BH^+ \overset{K_{a2}}{\rightleftharpoons} B + H^+$$

Where BH_2^{2+} is the doubly protonated conjugate acid of the basic compound and B is the neutral species. The equilibrium constant can then be written as:

$$[K_{a1}] = \frac{[BH^+][H^+]}{[BH_2{}^{2+}]}, \quad [K_{a2}] = \frac{[B][H^+]}{[BH^+]}$$

The total concentration of compound compound B can be defined by the sum of concentrations of the individual species and the equation can be formulated in terms of [B], $[H^+]$ and the equilibrium constants, K_{a1} and K_{a2} :

$$[B]_T = [B] + [BH^+] + [BH_2^{2+}]$$

$$= [B] + \frac{[B][H^+]}{K_{a2}} + \frac{[B][H^+]^2}{K_{a1}K_{a2}}$$

$$= [B](1 + 10^{-log(K_{a2}) + log(H^+)} + 10^{-log(K_{a1}) - log(K_{a2}) + 2log(H^+)})$$

$$= [B](1 + 10^{pK_{a2} - pH} + 10^{pK_{a1} + pK_{a2} - 2pH})$$

The total external concentration of B at the external pH (pH_e) is given by:

$$[B]_{T,e} = [B](1 + 10^{pK_{a2} - pHe} + 10^{pK_{a1} + pK_{a2} - 2pHe})$$

The total concentration of B in the DV at the vacuolar pH (pH_v) is given by:

$$[B]_{T,v} = [B](1 + 10^{pK_{a2} - pHv} + 10^{pK_{a1} + pK_{a2} - 2pHv})$$

Assuming that the neutral species B equilibrates equally across membrane and that the membranes are totally impervious to the charged forms BH^+ and BH_2^{2+}, then the extracellular and intravacuolar [B] values are equal and the VAR for B is given by the ratio of the total external and vacuolar concentrations (*Eq. 8.1*):

$$VAR = \frac{[B]_{T,v}}{[B]_{T,e}} = \frac{(1+10^{pKa2-pHv}+10^{pKa1+pKa2-2pHv})}{(1+10^{pKa2-pHe}+10^{pKa1+pKa2-2pHe})} \quad \textbf{Eq. 8.1}$$

8.1.3.3 Haem binding

The association constant for the interaction between selected benzamides and Fe(III)PPIX was determined using a modified version of the titration method reported by Egan and co-workers.[316] Briefly, a stock solution of the test sample in either 40% aqueous DMSO and HEPES buffer (0.02 M, pH 7.4) or 100% DMSO (for the benzamides) was prepared. The initial concentrations the quinolines (CQ and QD) and benzamides were 2 mM and 20 mM respectively. A working solution of haematin (7.5 µM) and HEPES buffer (0.02 M, pH 7.4) in 40% aqueous DMSO was prepared and used immediately. Identical volumes of the stock solution were added with a Hamilton syringe into quartz cuvettes (Hellma) of 1 cm path length containing the haematin working solution or a blank solution of 40% aqueous DMSO and HEPES buffer (0.02 M, pH 7.4). The UV-vis absorbance values of the blank were automatically subtracted from that of the haem working solution. The UV-vis absorbance was monitored from 350 nm to 650 nm at 25 °C with each addition of the test sample on a Varian Cary 100 UV–Visible spectrophotometer. The data obtained were corrected for dilution and analysed using non-linear least squares fitting methods. Titrations were performed in triplicate and the average association constant (K) values were reported with the standard error of the mean (SEM). The value of logK was calculated from the best fit model, which described a 1:1 association using *Eq. 8.2*, where A_0 is the initial absorbance (no test sample), A_∞ is the final limiting absorbance and [L] is the concentration of the free ligand, which in the case of weak association is approximately the total test sample concentration in the haem working solution.

$$A = \frac{A_0 + A_\infty K[L]}{1 + K[L]} \quad \textbf{Eq. 8.2}$$

8.1.3.4 Correlation analysis

Identification of statistically significant correlations was carried out using Molecular Modelling Pro Plus (MMP+) software.[290] The MDL Molfile containing structural data for each set of test samples (benzamides or triarylimidazoles) was input into the programme and an extensive range of physicochemical properties were calculated for each molecule. The data were stored as a .txt file and then edited in Microsoft Excel[313] to include the parasite and βH activities as the inverse, therefore allowing inactive compounds to be entered with an inverse activity of zero. The resulting database was opened in MMP+ where the data were analysed by "regression with all combinations". The output correlations, which were both physically relevant and displayed the highest r^2 were considered. The predicted activity values for each molecule were then calculated from the output linear equation and plotted vs the experimental activity. Those correlations which appeared as two clusters of connected points, where no correlation existed within each cluster, were omitted.

8.1.4 *P. falciparum* cell culturing methods

8.1.4.1 LDH malaria parasite survival assay

Cultures of CQ-sensitive D10 and NF54 *P. falciparum* in the asexual erythrocyte stages were continuously maintained *in vitro* using a modified version of the method of Trager and Jensen.[70] The antimalarial activity of the purchased and synthesised benzamides and triarylimidazoles were assessed using the parasite lactate dehydrogenase (LDH) assay described by Makler *et al.*[79] All samples for NF54 testing and ten of the samples for D10 testing were sent to the Department of pharmacology at the University of Cape Town Medical School, where the experiments were carried out by others. The remaining samples were tested in D10 parasites, which were also continuously cultured, by the author of this thesis using the method described below.

The D10 strain of *P. falciparum* was maintained in culture flasks containing O+ human erythrocytes (2% haematocrit, 1-10% parasitaemia) suspended in RPMI 1640 medium (10.4 g/L) with glutamine, glucose (22 mM), HEPES buffer (25 mM), hypoxanthine (0.65 mM), Albumax (5 g/L), gentamicin (0.1 mM) and $NaHCO_3$ (32 mM). Flasks were gassed before incubation at 37 °C with a mixture of 3% O_2, 4% CO_2, and 93% N_2. CQ, tested up to 100

ng/ml maximum concentration, was used as a control on each plate. Parasitised red blood cells (pRBCs) were prepared via dilution from the culture flask to give a 2% haematocrit and 2% initial parasitaemia. The samples were prepared as 2 mg/ml stock solutions in DMSO and diluted with complete cell culture medium before testing in triplicate in a 96-well plate. Row 1 of the plate was reserved as a blank with uninfected RBCs (200 µL 1% haematocrit). Row 2 contained the positive control with no test compound (200 µL, 1% haematocrit, 1% parasitaemia). Rows 3-12 contained the inoculated pRBCs at a range of sample concentrations, prepared by adding 200 µL of the required sample stock to row 3 and performing a two-fold serial dilution across the plate to row 12. Finally, the pRBCs (100 µL) were added to rows 3-12 to give a final volume of 200 µL, haematocrit of 1% and parasitaemia of 1%. The plate was incubated at 37 °C for 48 h in a gas chamber containing 3% O_2, 4% CO_2 and 93% N_2. Following resuspension of the plate contents in a non-sterile environment using a multichannel pipette, 15 µL from each well was transferred to the corresponding well of a new plate containing Malstat (100 µL per well). Nitroblue tetrazolium (NBT, 25 µL) was added to each well and the plate was left to develop in the dark for 10 min before measuring the absorbance at 620 nm on a microplate reader. The highest concentration of solvent to which the parasites were exposed had no measurable effect on the parasite viability. The IC_{50} values were obtained from fitting analysis using a non-linear least-squares dose response curve in GraphPad Prism v.3.0 software.[314]

8.1.4.2 Cell fractionation

Target validation was carried out by measuring the increase in free haem and the decrease in haemozoin formation via cell fractionation studies. The samples were sent to the Division of Pharmacology, Department of Medicine at the University of Cape Town Medical School for testing by Jill Combrinck using the procedure described by Combrinck at al.[113]

8.1.4.3 The inoculum effect

The CAR was determined via the inoculum effect using a procedure similar to those previously reported.[99,293] The LDH assay described above was employed to measured IC_{50} values in triplicate at inoculum sizes ranging from 1 to 10 and fractional volumes of

parasitised erythrocytes from 0.0001 to 0.001. The relationship was extrapolated via linear regression to give the absolute IC_{50} from the y-intercept in Graph Pad Prism v3.0.[314] The derivation for the equation describing the CAR (*Eq. 8.3*) for compound A is shown below.

The total number of moles and volume are given by:

$$n_T = n_{int} + n_{ext}$$

$$V_T = V_{int} + V_{ext}$$

The fractional volume of pRBCs for low haematocrit and parasitaemia is defined as:

$$V_{fracPRBCs} = \frac{V_{int}}{V_T} \approx \frac{V_{int}}{V_{ext}}$$

The CAR is defined as the ratio of internal to external compound concentration:

$$CAR = \frac{[A]_{int}}{[A]_{ext}}$$

$$= \frac{n_{int}/V_{int}}{n_{ext}/V_{ext}}$$

$$= \frac{n_{int} \times V_{ext}}{n_{ext} \times V_{int}}$$

$$= \frac{n_{int}}{n_{ext}} \times \frac{1}{V_{fracRBCs}}$$

$$= \frac{n_T - n_{ext}}{n_{ext}} \times \frac{1}{V_{fracRBCs}}$$

$$= \frac{\frac{n_T}{V_T} - \frac{n_{ext}}{V_T}}{\frac{n_{ext}}{V_T} \times V_{fracRBCs}}$$

For a particular experiment with pRBCs, the measured concentration which inhibits 50% of parasite growth is based on the total number of moles and volume delivered to the system.

$$IC_{50_{measured}} = \frac{n_T}{V_T}$$

In the case where there is no parasite, the total number of moles is equivalent to the number of moles externally, since there is no depletion of compound A in the medium:

$$IC_{50_{absolute}} = \frac{n_{ext}}{V_T} \quad \text{(no parasite)}$$

Substituting these expressions into the equation gives the CAR in terms of the measured and absolute inhibitory concentrations.

$$CAR = \frac{IC_{50_{measured}} - IC_{50_{absolute}}}{IC_{50_{absolute}} \times V_{frac.PRBCs}} \qquad \textbf{Eq. 8.3}$$

8.1.5 Bayesian modelling methods

8.1.5.1 Training set data

GSK (TCAMS) and St Jude's whole-cell screening data were downloaded from the ChEMBL database (www.ebi.ac.uk/chemblntd). The βH activity data were sourced from previous HTS collaborations between VU, UCT and OU, most of which are publically available.[107,204,108,282,109,110]

8.1.5.2 Comparing the chemical space of libraries

Principle component analysis (PCA) was carried out in Discovery Studio[305] using the following descriptors: logP, MW, #RB, #R, #AR, #HBA and #HBD. The assembly method for comparing libraries decomposed the molecules into unique occurrences of ring, bridge or chain assemblies.[317] The libraries were then compared using Tanimoto similarity of the assemblies. The βH hits (taken at >60% inhibition at 19 μM) were compared to the non-hits (<40%) using two Bayesian classification models and a Bayesian distance based on the Bayesian scores of each sample in the set.

8.1.5.3 Building Bayesian models

All data were modelled using Discovery Studio's[305] built-in Bayesian categorization, based on Bayes' theorem (*Eq. 6.1*). Input structure data files (SDFs) containing structural and activity data were imported into Discovery Studio and the sample data marked as active or inactive based on a user-defined IC_{50} cut-off. In the case of the βH inhibiting model, the samples were represented at pH 5 to match the conditions of the acidic DV. Default input descriptors from which the program can learn to distinguish active from inactive compounds were chosen. These 2D parameters were calculated by the program during the simulation: logP, MW, #RB, #R, #AR, #HBA, #HBD and ECFP_6. The model allocated each feature a probability score, weighted by a Laplacian-corrected estimator based on the frequency of occurrence of that feature in the active and inactive sets. In order to predict the likelihood of activity for a test compound, weights for the different features were summed to give a probability estimate.

Internal validation of the generated models was determined by the ROC score, based on the area under the plot of true positive rate vs false positive rate (ROC curve). These rates were calculated by leaving each molecule out of the training set one at a time (leave-one-out cross validation), or by leaving one fifth of the training set out (5-fold cross validation) and predicting their activities with those remaining. A score of 1 represents a perfect prediction with no false positives while 0.5 represents no enrichment. The ROC score was optimised by combining several datasets and generating models with different IC_{50} cut-offs for activity input. Training sets excluded several compounds for use as external test sets.

8.1.5.4 Model validation using external datasets

The DrugBank database (http://www.drugbank.ca/) contains 1510 FDA approved small molecule drugs.[318,319] These compounds were used as a test set to measure the accuracy of the Bayesian models since many have reported antimalarial activity data in PubChem (https://pubchem.ncbi.nlm.nih.gov/). In addition to this validation method, the ZINC database (http://zinc.docking.org/), a free collection of commercially available compounds for virtual screening was employed.[320] Purchasable samples from the Aldrich^{CPR} catalogue in ZINC were filtered through the models, listed according to predicted activity probabilities and where similar structures occurred; only one analogue was selected. Compounds were also excluded if they were currently unavailable or if they were expensive. Three sets of compounds were then purchased from Sigma-Aldrich; ten molecules predicted to be inactive for βH inhibition (**SA1-SA10**), ten with high probability of being βH inhibitors (**SA11-SA20**) and fourteen predicted to be both βH inhibitors and biologically active (**SA21-SA34**). These compounds were then tested for βH inhibition using the NP-40 method described in *Section 8.1.3.1* [321] and against the NF54 strain of *P. falciparum* (*Section 8.1.4.1*).[79,70,208]

8.2 Synthetic methods

8.2.1 Materials and general procedures

Glassware was washed with combinations of dishwashing detergent, water and acetone and in certain cases, nitric or sulphuric acid. Calcium chloride was used as a drying agent above the reflux condenser for all reflux reactions. Heating was carried out on a IKA® RCT basic IKAMAG™ safety control hot plate with a IKA® ETS-D5 temperature controller probe. All reactions were monitored by thin layer chromatography (TLC) using silica gel plates (Merck F_{254} aluminium-backed). TLC plates were visualised with ultraviolet light (254 nm) and, where an amine was present, plates were stained with ninhydrin spray (0.2% w/w in EtOH). Silica gel flash column chromatography was carried out with Fluka 60: 70-230 mesh on either a Biotage Isolera One Flash Chromatography System or using a combination of a D-star DVW-10 variable wavelength detector with a Teledyne ISCO fraction collector. Melting points were measured using a Reichert-Jung Thermovar hot stage microscope. Proton (^{1}H) and carbon (^{13}C) NMR spectra were recorded using a Varian Mercury spectrometer (300 MHz for ^{1}H), a Bruker 300 (at 300.08 MHz for ^{1}H) or a Bruker Ultrashield 400 Plus (at 399.95 MHz for ^{1}H and 100.64 MHz for ^{13}C) spectrometer. NMR experiments were performed in deuterodimethyl sulfoxide (DMSO-d_6), deuterochloroform ($CDCl_3$), deuteromethanol (MeOD) or deuteroacetone (acetone-d_6) with internal standards at δ_H = 2.50, 7.26, 3.31 or 2.05 ppm respectively for ^{1}H NMR and δ_C = 40.05, 77.16, 49.00 or 29.8 and 206.26 ppm respectively for ^{13}C NMR. All chemical shifts were reported in ppm and *J* coupling values in Hz.

Electron ionisation mass spectrometry was recorded using a JEOL GC mate II single magnetic mass spectrometer. High resolution mass spectrometry (HRMS) was performed on a Time-of-flight (TOF) Waters Synapt G2 instrument using leucine encephalin as a standard. All mass spectra for the benzamide series were recorded using the electrospray negative (ES^-) technique and introduction of the sample was via an ESI probe injected into a stream of MeOH. Mass spectra for the triarylimidazole series were recorded using the electrospray positive (ES^+) technique and the sample injected into a stream of acetonitrile.

High performance liquid chromatography (HPLC) was carried out on an Agilent Technologies 1220 Infinity LC (G4288C) with a C18 reverse phase column in HPLC grade acetonitrile (ACN) and double distilled deionised Millipore® Direct-Q water.

Commercially available solvents, chemicals and starting materials were purchased from the suppliers listed in *Table 8.2*. EtOAc and hexane were distilled before use. Tetrahydrofuran (THF) was distilled under nitrogen and dried over sodium wire with benzophenone. Dry DCM was distilled under nitrogen. Deuterated solvents for NMR were stored in a desiccator with phosphorus pentoxide (P_2O_5) drying agent.

Table 8.2 Suppliers for the commercially available chemicals used without purification unless otherwise stated in the synthesis, workup, purification and characterisation.

ITEM	SUPPLIER
Solvents	
Dichloromethane (DCM) AR grade	Kimix
Ethyl acetate (EtOAc)	Protea Chemicals
Methanol (MeOH)	Kimix
Ethanol (EtOH) 99.9%	Kimix
Hexane	Kimix
Acetone	Protea Chemicals
Acetone AR grade	Kimix
Tetrahydrofuran (THF)	Kimix
Toluene	Kimix
Pyridine (Pyr.)	Sigma-Aldrich
Dimethylformamide (DMF)	Sigma-Aldrich
Acetonitrile (ACN) HPLC grade	Kimix
Deuterodimethylsulfoxide (DMSO-d_6)	Sigma-Aldrich
Hexadeuteroacetone (Acetone- d_6)	Sigma-Aldrich
Deuterochloroform ($CDCl_3$)	Sigma-Aldrich
Deuteromethanol (MeOD)	Sigma-Aldrich
Starting materials	
5-Nitroisophthalic acid	Sigma-Aldrich
5-Cyano-1,3-benzenedicarboxylic acid	Sigma-Aldrich
5-Methylisophthalic acid	Sigma-Aldrich
5-*Tert*-Butylisophthalic acid	Sigma-Aldrich
Isophthalic acid	Sigma-Aldrich
5-Methoxyisophthalic acid	Sigma-Aldrich
2,6-Pyridinedicarbonyl dichloride	Sigma-Aldrich
Aniline	Sigma-Aldrich
4-Amino pyridine	Sigma-Aldrich
3-Amino pyridine	Sigma-Aldrich

Table 8.2 (Continued)

ITEM	SUPPLIER
Starting materials	
2-Amino pyridine	Sigma-Aldrich
4-[(*N*-Boc)aminomethyl]aniline	Sigma-Aldrich
2-Aminobenzimidazole	Sigma-Aldrich
1,3,5-Tricarbonyltrichloride	Sigma-Aldrich
3,5-Dimethylbenzaldehyde	Sigma-Aldrich
3,4,5-Trimethoxybenzaldehyde	Sigma-Aldrich
Benzil	Sigma-Aldrich
Syringaldehyde	Sigma-Aldrich
Trifluoromethoxybenzaldehyde	Sigma-Aldrich
4-Chlorobenzoin	Sigma-Aldrich
4,4'-Dimethoxybenzoin	Sigma-Aldrich
4-Nitrobenzaldehyde	Sigma-Aldrich
Miscellaneous reagents and chemicals	
Formaldehyde (37 wt. % in H_2O)	Sigma-Aldrich
Trifluoroacetic acid (TFA)	Sigma-Aldrich
Sodium cyanoborohydride ($NaBH_3CN$)	Sigma-Aldrich
Triethylamine (Et_3N)	Sigma-Aldrich
Iron powder (Fe powder)	Unilab
Sodium hydride (NaH)	Sigma-Aldrich
Ammonium acetate (NH_4OAc)	Kimix
Celite	Sigma-Aldrich
Magnesium sulphate anhydrous ($MgSO_4$)	Kimix
Sodium carbonate (Na_2CO_3)	Kimix
Sodium chloride AR (NaCl)	Kimix
Ninhydrin	Sigma-Aldrich
Hydrochloric acid (32 wt. %)	Kimix
Sulphuric acid (H_2SO_4)	Kimix
Nitric acid (HNO_3)	Kimix
Phosphorus pentoxide (P_2O_5)	Sigma-Aldrich

8.2.2 Synthesis, purification and characterisation

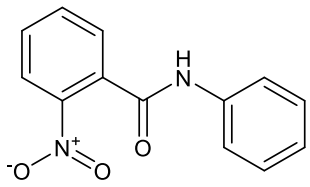
8.2.2.1 Benzamides

General procedure for the synthesis of the monobenzamides 2a-b

Aniline and the relevant acid chloride (**1a** or **1b**) were stirred in dry pyridine under N_2 at RT. Upon completion of the reaction, indicated by TLC analysis, the organic components were extracted in EtOAc (3 x 20 ml) and then washed with 1M HCl to remove excess pyridine. The organic fractions were combined, stirred with $MgSO_4$ and concentrated by evaporation

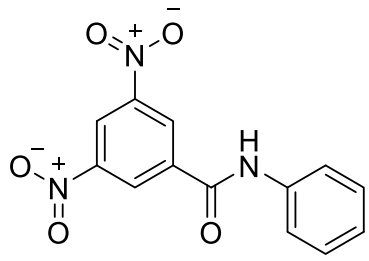
under reduced pressure. The product was dried under high vacuum for 2 h to yield the desired product.

2-Nitro-*N*-phenylbenzamide (2a)[265]



Scale: aniline (186 mg, 2 mmols) and 2-nitrobenzoyl chloride (371 mg, 2 mmols), stirred for 1 h. Yield of **2a**: 477 mg, 99%. Recrystallisation from AR-grade MeOH and AR-grade DCM afforded **2a** as off-white crystals: Mp 149-151 °C; Lit Mp[265] 155-156 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 10.62 (br. s, 1H), 8.13 (td, *J* = 0.70, 7.89 Hz, 1H), 7.86 (m, 1H), 7.76 (m, 2H), 7.66 (m, 2H), 7.35 (m, 2H), 7.11 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 164.6, 147.0, 139.3, 134.5, 133.2, 131.4, 129.8, 129.3, 124.7, 124.4, 120.2; HRMS-ES^{-}: Observed 241.0604 $(M-H)^{+}$; Calculated 241.0613 for $C_{13}H_9N_2O_3$; HPLC: 96.9%.

3,5-Dinitro-*N*-phenylbenzamide (2b)[322]

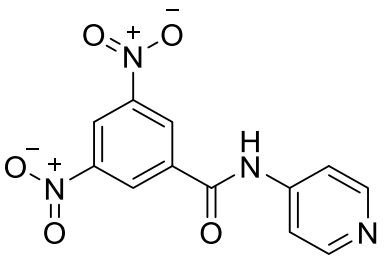


Scale: aniline (200 mg, 2.1 mmols) and 3,5-dinitrobenzoyl chloride (606 mg, 2.6 mmols), stirred for 2 h. Yield of **2b**: 505 mg, 84% as a pale-yellow solid. Recrystallisation from absolute EtOH afforded **2b** as white needles: Mp 236-238 °C; Lit Mp[322] 236 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 10.80 (s, 1H), 9.15 (d, *J* = 2 Hz, 2H), 8.98 (t, *J* = 2 Hz, 1H), 7.76 (m, 2H), 7.39 (m, 2H), 7.16 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 161.7, 148.6, 138.8, 138.0, 129.3, 128.5, 125.0, 121.6, 121.2; HRMS-ES^{-}: Observed 286.0450 $(M-H)^{+}$; Calculated 286.0464 for $C_{13}H_8N_3O_5$; HPLC: 97.3%.

General procedure for the synthesis of the monobenzamides 2c-e

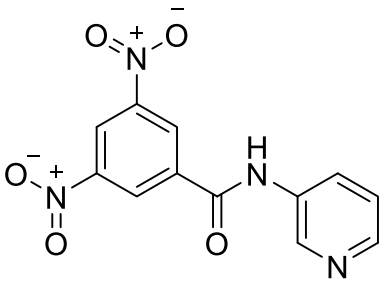
The relevant aminopyridine and acid chloride were added to dry pyridine, followed by the addition of Et_3N (1 eq.). The mixture was stirred under N_2 at RT. Upon completion of the reaction, indicated by TLC analysis, a saturated solution of Na_2CO_3 (20 ml) was added and the organic components were extracted into EtOAc (3 x 20 ml). The fractions were combined, washed with brine and then stirred with $MgSO_4$ before being concentrated by evaporation under reduced pressure. Excess pyridine was removed by an azeotropic evaporation with toluene to give a solid, which was dried under high vacuum for 2 h to yield the desired product.

3,5-Dinitro-*N*-(pyridin-4-yl)benzamide (2c)[234]



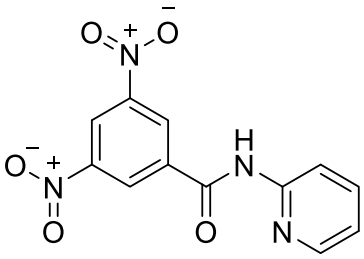
Scale: 4-aminopyridine (94 mg, 1 mmol) and 3,5-dinitrobenzoyl chloride (230 mg, 1 mmol), stirred for 2 h. Yield of **2c:** 268 mg, 93% as an off-white solid. Recrystallisation from AR grade acetone afforded **2c** as a white powder: Mp 267-268 °C; Lit Mp[234] 240-250 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 11.11 (br. s., 1H), 9.15 (d, *J* = 2 Hz, 2H), 9.02 (t, *J* = 2 Hz, 1H), 8.54 (d, *J* = 6.4 Hz, 2H), 7.78 (d, *J* = 6.4 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 162.8, 151.0, 148.7, 145.7, 137.3, 128.7, 122.0, 114.8; HRMS-ES$^-$: Observed 287.0411 $(M-H)^+$; Calculated 287.0416 for $C_{12}H_7N_4O_5$; HPLC: 99.7%.

3,5-Dinitro-*N*-(pyridin-3-yl)benzamide (2d)[233]



Scale: 3-aminopyridine (94 mg, 1 mmol) and 3,5-dinitrobenzoyl chloride (230 mg, 1 mmol), stirred for 3 h. Yield of **2d**: 234 mg, 81% as an off-white solid. Recrystallisation from AR-grade acetone afforded **2d** as a white powder: Mp 236-238 °C; Lit Mp[233] 231-232 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 11.02 (s, 1H), 9.17 (d, *J* = 2.0 Hz, 2H), 9.02 (t, *J* = 2.0 Hz, 1H), 8.94 (d, *J* = 2.4 Hz, 1H), 8.38 (m, 1H), 8.20 (m 1H), 7.45 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 162.3, 148.7, 145.9, 142.8, 137.4, 135.5, 128.6, 128.4, 124.2, 121.8; HRMS-ES^{-}: Observed 287.0403 $(M-H)^{+}$; Calculated 287.0416 for $C_{12}H_7N_4O_5$; HPLC: 99.3%.

3,5-Dinitro-*N*-(pyridin-2-yl)benzamide (2e)[323]



Scale: 2-aminopyridine (94 mg, 1 mmol) and 3,5-dinitrobenzoyl chloride (230 mg, 1 mmol), stirred for 3 h. Yield of **2e**: 286 mg, 99% as an off-white solid. Recrystallisation from AR-grade MeOH afforded **2e** as a white powder: Mp 196-198 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 11.57 (s, 1H), 9.18 (d, *J* = 2.2 Hz, 2H), 8.98 (t, *J* = 2.2 Hz, 1H), 8.42 (m, 1H), 8.19 (m, 1H), 7.89 (m, 1H), 7.23 (m, 1H); ^{13}C NMR (151 MHz, DMSO-d_6) δ ppm 162.7, 152.1, 148.6, 148.5, 138.8, 137.4, 129.0, 121.8, 121.0, 115.4; HRMS-ES^{-}: Observed 287.0402 $(M-H)^{+}$; Calculated 287.0416 for $C_{12}H_7N_4O_5$; HPLC: 99.7%.

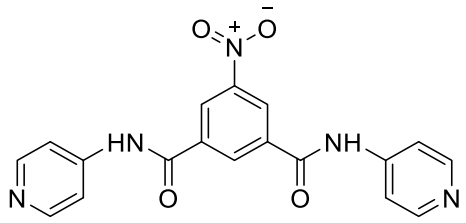
General procedure for acid chlorides (4a-n) from the corresponding isophthalic acid (3a-n)

Thionyl chloride (10 ml) was added to the appropriate isophthalic acid (**3a-n**) and refluxed at 80 °C for 4 h and then cooled and stirred for 1 h or overnight. The excess thionyl chloride was evaporated under reduced pressure. After drying the residue for 1 h on a high vacuum pump, the product (**4a-n**) was not isolated but used immediately in the next step from the same flask.

General procedure for dibenzamides (5a-5p) from the corresponding acid chloride (4a-p)

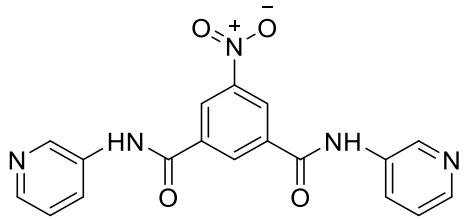
The appropriate freshly prepared acid chloride (**4a-m**) or purchased acid chloride (**4o-p**) was dissolved in 100% dry pyridine unless otherwise specified, followed by the addition of the relevant amine (aniline or aminopyridine) and Et_3N (2 equiv.). The reaction was stirred under N_2 at RT for the required time. Upon completion of the reaction as indicated by TLC analysis, one of three procedures was followed for the isolation. i) For the pyridylbenzamides (**5a**, **5b**, **5d**, **5f**, **5h**, **5i** and **5m**), a saturated solution of Na_2CO_3 (20 ml) was added and the organic components were extracted into EtOAc (3 x 20 ml). The organic fractions were combined, washed with brine and then dried with $MgSO_4$. The solutions were concentrated by evaporation under reduced pressure and excess pyridine was removed by an azeotropic evaporation with toluene to give a solid. ii) For the phenylbenzamides (**5c**, **5e**, **5g**, **5j** and **5n**), the organic components were extracted into EtOAc (3 x 20 ml) and then washed with 1M HCl to remove excess pyridine. The organic fractions were then combined, dried with $MgSO_4$ and concentrated by evaporation under reduced pressure. iii) For the compounds **5k**, **5l**, **5o** and **5p**, precipitation of the compounds from the reaction mixture allowed for isolation of the white solid by filtration. All compounds **5a**-**5p**, were dried under high vacuum for 2 h to yield the desired product.

5-Nitro-*N*,*N*-di(pyridin-4-yl)isophthalamide (5a)



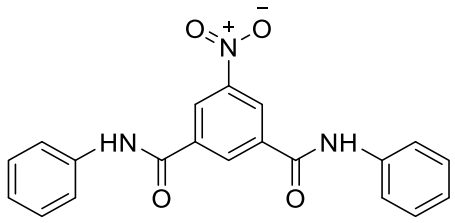
Scale: **4a** (818 mg, 3.3 mmols) and 4-aminopyridine (627 mg, 6.7 mmols), stirred for 3 h. Yield of **5a**: 795 mg, 92% as an off-white solid. Recrystallisation from AR-grade acetone afforded **5a** as small off-white crystals: Mp 303-306 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 11.93 (br. s, 2H), 9.33 (t, *J* = 1.5 Hz, 1H), 8.95 (d, *J* = 1.5 Hz, 2H), 8.69 (d, *J* = 6.9 Hz, 4H), 8.26 (d, *J* = 6.9 Hz, 4H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 164.7, 150.6, 148.6, 146.0, 135.5, 134.2, 127.1, 115.5; HRMS-ES^{-}: Observed 362.0887 $(M-H)^{+}$; Calculated 362.0889 for $C_{18}H_{12}N_5O_4$; HPLC: 96.6%.

5-Nitro-*N*,*N*-di(pyridin-3-yl)isophthalamide (5b)



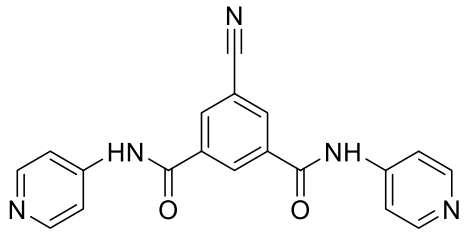
Scale: **4b** (496 mg, 2 mmols) and 3-aminopyridine (394 mg, 4.2 mmols), stirred for 4 h. Yield of **5b**: 576 mg, 79% as an off-white solid. Recrystallisation from AR-grade acetone afforded **5b** as small off-white crystals: Mp 134-136 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.91 (s, 2H), 9.00 (br.s, 3H), 8.96 (d, *J* = 2.0 Hz, 2H), 8.36 (m, 2H), 8.21 (m, 2H), 7.44 (m, 8.3 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.7, 148.4, 145.7, 142.7, 136.6, 135.8, 133.7, 128.2, 125.8, 124.1 HRMS-ES^{-}: Observed 362.0880 $(M-H)^{+}$; Calculated 362.0889 for $C_{18}H_{12}N_5O_4$; HPLC: 98.4%.

5-Nitro-*N*,*N*-diphenylisophthalamide (5c)[324]



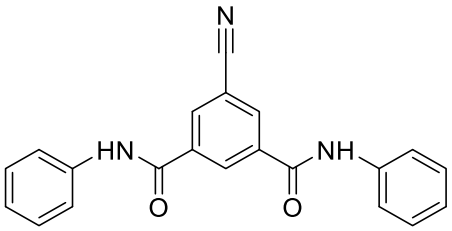
Scale: **4c** (496 mg, 2 mmols) and aniline (376 mg, 4 mmols), stirred for 3 h. Yield of **5c**: 675 mg, 93% as an off-white solid. Recrystallisation from AR-grade acetone afforded **5c** as white crystals: Mp 280-283 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.69 (s, 2H), 8.79 - 9.12 (m, 3H), 7.78 (m, 4H), 7.41 (m, 4H), 7.13 (m, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.3, 148.4, 139.1, 137.1, 133.6, 129.2, 125.5, 124.8, 121.1; HRMS-ES^{-}: Observed 360.0970 $(M-H)^{+}$; Calculated 360.0984 for $C_{20}H_{14}N_3O_4$; HPLC: 98.8%.

5-Cyano-*N*,*N*-di(pyridin-4-yl)isophthalamide (5d)



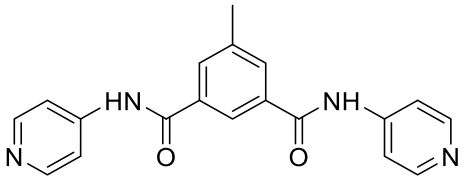
Scale: **4d** (114 mg, 0.5 mmols) and 4-aminopyridine (108 mg, 1.1 mmol), stirred for 3 h in dry DCM (20 ml) and pyridine (0.5 ml). Purification by column chromatography (gravity) with 5-10% MeOH/DCM was carried out. Yield of **5d**: 45 mg, 26%. Recrystallisation from AR-grade acetone afforded **5d** as small off-white crystals: Mp 285-288 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 11.58 (br. s, 2H), 8.99 (t, *J* = 1.6 Hz, 1H), 8.68 (d, *J* = 1.6 Hz, 2H), 8.65 (d, *J* = 6.6 Hz, 4H), 8.12 (d, *J* = 6.6 Hz, 4H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 164.5, 151.0, 145.9, 136.2, 135.0, 132.4, 118.1, 114.6, 112.5; HRMS-ES^{-}: Observed 342.0987 $(M-H)^{+}$; Calculated 342.0991 for $C_{19}H_{12}N_5O_2$; HPLC: 99.0%.

5-Cyano-*N*,*N*-diphenylisophthalamide (5e)



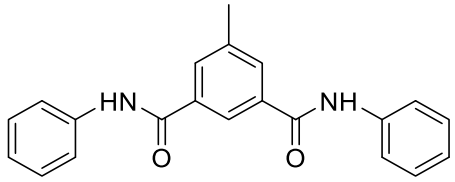
Scale: **4e** (228 mg, 1 mmol) and aniline (201 mg, 2.1 mmols), stirred for 2 h. Yield of **5e**: 308 mg, 90% as an off-white solid. Recrystallisation from AR-grade DCM gave **5e** as a white powder: Mp 300-303 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 10.55 (br. s, 2H), 8.78 (t, *J* = 1.7 Hz, 1H), 8.60 (d, *J* = 1.7 Hz, 2H), 7.80 (m, 4H), 7.42 (m, 4H), 7.17 (m, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 163.7, 139.1, 136.8, 134.3, 132.1, 129.2, 124.7, 120.9, 118.4, 112.3; MS-EI^{+} Observed 341.1129 M^{+}; Calculated 341.1164 for $C_{21}H_{15}N_3O_2$; HPLC: 96.7%.

5-Methyl-*N*,*N*-di(pyridin-4-yl)isophthalamide (5f)



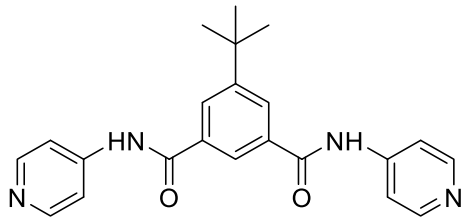
Scale: **4f** (434 mg, 2 mmols) and 4-aminopyridine (376 mg, 4 mmol), stirred for 1 h. Yield of **5f**: 557 mg, 84% as an off-white solid. Recrystallisation from absolute EtOH afforded **5f** as a white powder: Mp 209-212 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 10.70 (s, 2H) 8.49 (d, J = 6.4 Hz, 4H) 8.34 (s, 1H) 8.00 (m, 2H) 7.79 (d, *J* = 6.4, 4H) 2.49 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 166.4, 150.8, 146.3, 139.0, 135.1, 132.2, 125.1, 114.6, 21.3; HRMS-ES^{-}: Observed 331.1191 $(M-H)^{+}$; Calculated 331.1195 for $C_{19}H_{15}N_4O_2$; HPLC: 99.3%.

5-Methyl-*N*,*N*-diphenylisophthalamide (5g)



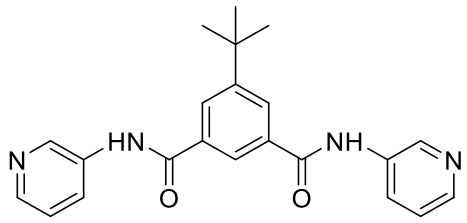
Scale: **4g** (434 mg, 2 mmols) and aniline (372 mg, 4 mmol), stirred for 2 h. Yield of **5g**: 557 mg, 84% as an off-white solid. Recrystallisation from MeOH and AR-grade DCM afforded **5g** as small white crystals: Mp 257-258 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.33 (s, 2H), 8.31 (s, 1H), 7.95 (m, 2H), 7.79 (m, 4H), 7.36 (m, 4H), 7.10 (m, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 165.7, 139.6, 138.6, 135.8, 131.6, 129.1, 124.8, 124.3, 120.9, 21.4; HRMS-ES^{-}: Observed 329.1288 $(M-H)^{+}$; Calculated 329.1290 for $C_{21}H_{17}N_2O_2$; HPLC: 99.0%.

5-(*tert*-Butyl)-*N*,*N*-di(pyridin-4-yl)isophthalamide (5h)



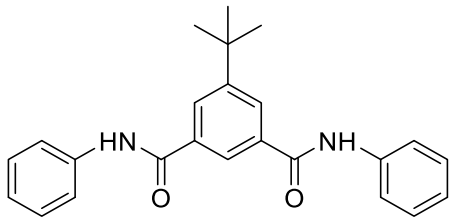
Scale: **4h** (518 mg, 2 mmols) and 4-aminopyridine (376 mg, 4 mmol), stirred for 3 h. Yield of **5h**: 467 mg, 62% as an off-white solid. Recrystallisation from AR-grade MeOH and AR-grade DCM afforded **5h** as a white powder: Mp 245-247 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.74 (s, 2H), 8.52 (d, *J* = 6.2 Hz, 4H), 8.41 (t, *J* = 1.6 Hz, 1H), 8.17 (d, *J* = 1.6 Hz, 2H), 7.82 (d, *J* = 6.2 Hz, 4H), 1.42 (s, 9H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 166.7, 152.2, 150.8, 146.3, 134.9, 128.6, 125.1, 114.7, 35.4, 31.4; HRMS-ES^{-}: Observed 373.1657 $(M-H)^{+}$; Calculated 373.1665 for $C_{22}H_{21}N_4O_2$; HPLC: 99.1%.

5-(*tert*-Butyl)-*N*,*N*-di(pyridin-3-yl)isophthalamide (5i)



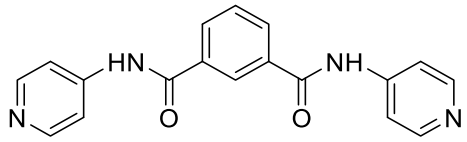
Scale: **5i** (259 mg, 1 mmol) and 3-aminopyridine (188 mg, 2 mmols), stirred for 3 h in pyridine (2 ml) and dry DCM (2 ml). Yield of **5i**: 294 mg, 79% as an off-white solid. Recrystallisation from AR-grade MeOH and AR-grade DCM afforded **5i** as a white powder Mp 227-229 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 10.57 (s, 2H), 8.94 (t, *J* = 1.3 Hz, 2H), 8.41 (t, *J* = 1.3 Hz, 1H), 8.33 (d, *J* = 4.7 Hz, 2H), 8.18 (m, 4H), 7.41 (d, *J* = 4.7, 2H), 1.4 (m, 9H); ^{13}C NMR (151 MHz, DMSO-d_6) δ ppm 166.2, 152.1, 145.3, 142.7, 136.1, 135.0, 128.3, 128.1, 125.0, 124.1, 35.4, 31.5; HRMS-ES^{-}: Observed 373.1653 $(M-H)^{+}$; Calculated 373.1665 for $C_{22}H_{21}N_4O_2$; HPLC: 99.4%.

5-(*tert*-Butyl)-*N*,*N*-diphenylisophthalamide (5j)



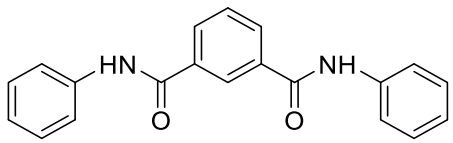
Scale: **4j** (518 mg, 2 mmols) and aniline (372 mg, 4 mmols), stirred for 2 h. Yield of **5j**: 603 mg, 81% as an off-white solid. Recrystallisation from AR-grade MeOH and AR-grade DCM afforded **5j** as white needles: Mp 285-286 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 11.26 (s, 2H) 9.26 (t, J = 1.6, 1H) 9.00 (d, *J* = 1.6 Hz, 2H) 8.67 (m, 4H) 8.25 (m, 4H) 7.99 (m, 2H); ^{13}C NMR (151 MHz, DMSO-d_6) δ ppm 165.8, 151.8, 139.5, 135.5, 129.1, 127.9, 124.8, 124.3, 121.1, 35.4, 31.5; HRMS-ES^{-}: Observed 371.1749 $(M-H)^{+}$; Calculated 371.1760 for $C_{24}H_{23}N_2O_2$; HPLC: 99.3%.

_N_,_N_-Di(pyridin-4-yl)isophthalamide (5k)[325]



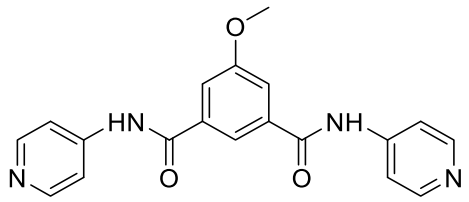
Scale: **4k** (406 mg, 2 mmols) and 4-aminopyridine (376 mg, 4 mmols), stirred for 2 h in pyridine (4 ml) and DCM (10 ml). The product was isolated by filtration of the white reaction precipitate. Yield of **5k**: 422 mg, 66%. Recrystallisation from AR-grade DCM afforded **5k** as a white powder: Mp 334-335 °C; Lit Mp[326] >350 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 10.74 (s, 2H), 8.54 (t, *J* = 1.8 Hz, 1H), 8.49 (d, *J* = 6.3 Hz, 4H), 8.18 (dd, *J* = 1.8, 7.9 Hz, 2H), 77.80 (d, *J* = 6.3 Hz, 4H), 7.73 (t, *J* = 7.9 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 166.3, 150.9, 146.3, 135.1, 131.8, 129.3, 127.8, 114.6; HRMS-ES^{-}: Observed 317.1033 $(M-H)^{+}$; Calculated 317.1039 for $C_{18}H_{13}N_4O_2$; HPLC: 95.6%.

_N_,_N_-Diphenylisophthalamide (5l)[327]



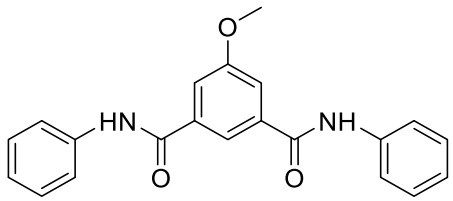
Scale: **4l** (406 mg, 2 mmols) and aniline (372 mg, 4 mmols), stirred for 2 h. Precipitation of the white product was assisted with DCM. The product was isolated by filtration of the reaction precipitate. Yield of **5l**: 455 mg, 72%. Recrystallisation from AR-grade DCM afforded **5l** as a white powder: Mp 383-384 °C; Lit. Mp[327] >300 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 10.39 (s, 2H), 8.53 (t, *J* = 1.6 Hz, 1H), 8.13 (dd, *J* = 1.6, 7.8 Hz, 2H), 7.82 (m, 4H), 7.68 (t, *J* = 7.8 Hz, 1H), 7.27 - 7.46 (m, 4H), 7.11 (m, 7.5 Hz, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 165.5, 139.5, 135.7, 131.1, 129.1, 129.1 127.5, 124.3, 120.9; HRMS-ES^{-}: Observed 315.1135 $(M-H)^{+}$; Calculated 315.1134 for $C_{20}H_{15}N_2O_2$; HPLC: 98.3%.

5-Methoxy-*N*,*N*-di(pyridin-4-yl)isophthalamide (5m)



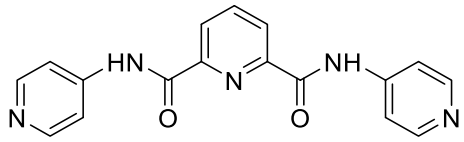
Scale: **4m** (466 mg, 2 mmols) and 4-aminopyridine (376 mg, 4 mmols), stirred for 4 h. Yield of **5m**: 483 mg, 69% as an off-white solid. Recrystallisation from AR-grade MeOH afforded **5m** as a white powder: Mp 144-145 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 10.74 (s, 2H), 8.52 (d, *J* = 6.4 Hz, 4H), 8.16 (t, *J* = 1.5 Hz, 1H), 7.82 (d, *J* = 6.4 Hz, 4H), 7.74 (d, *J* = 1.5 Hz, 2H), 3.95 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 166.1, 159.8, 150.8, 146.3, 136.4, 120.1, 117.3, 114.6, 56.4; HRMS-ES^{-}: Observed 347.1132 $(M-H)^{+}$; Calculated 347.1144 for $C_{19}H_{15}N_4O_3$; HPLC: 95.6%.

5-Methoxy-*N*,*N*-diphenylisophthalamide (5n)



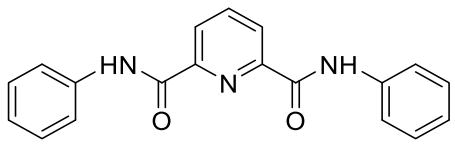
Scale: **4n** (466 mg, 2 mmols) and aniline (372 mg, 4 mmols), stirred for 2 h. Yield of **5n**: 637 mg, 92% as an off-white solid. Recrystallisation from AR-grade MeOH afforded **5n** as white needles: Mp 223-224 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 10.37 (s, 2H), 8.14 (t, *J* = 1.5 Hz, 1H), 7.81 (m, 4H), 7.70 (d, *J* = 1.5 Hz, 2H), 7.39 (m, 4H), 7.12 (m, 2H), 3.94 (s, 3H); ^{13}C NMR (151 MHz, DMSO-d_6) δ ppm 165.3, 159.7, 139.5, 137.1, 129.1, 124.3, 121.0, 119.8, 116.6, 56.3; HRMS-ES^{-}: Observed 345.1225 $(M-H)^{+}$; Calculated 345.1239 for $C_{21}H_{17}N_2O_3$; HPLC: 99.3%.

_N_,_N_-Di(pyridin-4-yl)pyridine-2,6-dicarboxamide (5o)[326]



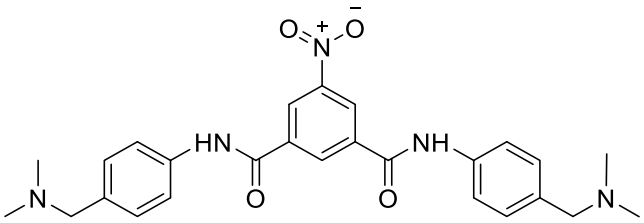
Scale: Pyridine-2,6-dicarbonyl dichloride (204 mg, 1 mmol) and 4-aminopyridine (188 mg, 2 mmols), stirred for 3 h. The product was isolated by filtration of the white reaction precipitate. Yield of **5o**: 307 mg, 92%. Recrystallisation from AR-grade MeOH and AR-grade DCM afforded **5o** as a white powder: Mp 327-328 °C; Lit. Mp[326] >300 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 12.17 (s, 2H), 8.71 (d, *J* = 6.6 Hz, 4H), 8.49 (m, 2H), 8.46 (d, *J* = 6.6 Hz, 4H), 8.39 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 164.2, 149.7, 148.6, 146.5, 140.9, 127.3, 115.8; HRMS-ES$^-$: Observed 318.0986 $(M-H)^+$; Calculated 318.0991 for $C_{17}H_{12}N_5O_2$; HPLC: 97.0%.

_N_,_N_-Diphenylpyridine-2,6-dicarboxamide (5p)[328]



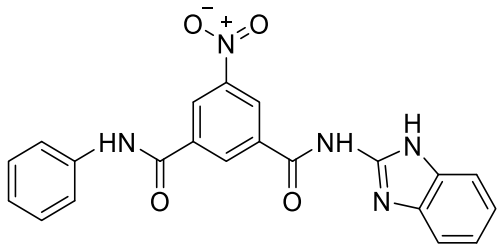
Scale: Pyridine-2,6-dicarbonyl dichloride (204 mg, 1 mmol) and aniline (186 mg, 2 mmols), stirred for 3 h. The product was isolated by filtration of the white reaction precipitate. Yield of **5p**: 302 mg, 91%. Recrystallisation from AR-grade MeOH and AR-grade DCM afforded **5p** as a white powder: Mp 271-272 °C; Lit. Mp[328] 278 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 11.00 (s, 2H), 8.39 (m, 2H), 8.29 (m, 1H), 7.91 (m, 4H), 7.43 (m, 4H), 7.18 (m, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 162.2, 149.4, 140.4, 138.5, 129.3, 125.8, 124.9, 121.7; HRMS-ES$^-$: Observed 316.1088 $(M-H)^+$; Calculated 316.1086 for $C_{19}H_{14}N_3O_2$; HPLC: 98.1%.

_N_,_N_-bis(4-((Dimethylamino)methyl)phenyl)-5-nitroisophthalamide (7)



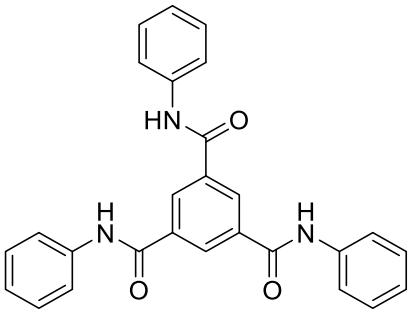
5-Nitroisophthaloyl dichloride (173 mg, 0.7 mmols) was prepared as per **4a** and then dissolved in pyridine (1 ml) and DCM (4 ml). 4-[(N-Boc)aminomethyl]aniline (318 mg, 1.4 mmols) was added and the reaction stirred at RT for 2 h. Following completion of the reaction indicated by TLC, EtOAc and 1 M HCl were added to the flask. The product was extracted into EtOAc (3 x 20 ml) and washed with cold 1 M HCl (3 x 20 ml) to remove the pyridine. The combined organic components were dried with $MgSO_4$ and the solvent removed under reduced pressure before being dissolved in dry DCM (5 ml) under inert atmosphere. TFA (1 ml) was slowly added at 0 °C and the mixture was stirred for 2.5 h. The excess TFA and DCM were removed on a rotoevaporator, followed by the addition of MeOH (20 ml) and K_2CO_3 (200 mg). The solution was stirred for 30 mins, then filtered through Celite and the MeOH evaporated off under reduced pressure. The resulting solid was dissolved in DMF (~10 ml) and formaldehyde (1.5 ml, 37% in water). Sodium cyanoborohydride (315 mg, 5 mmols) was added to the reaction stirred at RT for 5 h. Water (50 ml) was added and the product was extracted with EtOAc (3 x 50 ml), dried with $MgSO_4$ and the solvents evaporated under reduced pressure. The residual DMF was removed by heating under high vacuum on the pump for 12 h, and the resulting crude solid was recrystallized from MeOH to yield **7** as pale yellow crystals (111 mg, 53%); Mp 202-204 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 11.00 (s, 2H), 9.03 (t, *J* = 1.7 Hz, 1H), 8.61 (d, *J* = 1.7 Hz, 2H), 7.97 (d, *J* = 8.6 Hz, 4H), 7.55 (d, *J* = 8.6 Hz, 4H), 4.14 (s, 4H), 2.63 (s, 12H)**;** ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 163.5, 148.4, 140.0, 137.0, 133.6, 129.6, 126.5, 125.7, 120.5, 65.2, 49.0; MS-EI$^+$ Observed 475.1728 M^+; Calculated 475.2220 for $C_{26}H_{29}N_5O_4$; HPLC: 97.1%

N^1-(1H-Benzo[*d*]imidazol-2-yl)-5-nitro-N^3-phenylisophthalamide (8)



5-Nitroisophthaloyl dichloride (1.49 g, 6 mmols) was prepared as per **5a** and then dissolved in DCM (10 ml) under N_2 and cooled to -40 °C in an acetonitrile/liquid N_2 cooling bath. 2-Aminobenzimidazole (133 mg, 1 mmol) was added and the reaction allowed to come to RT before the addition of pyridine (2 drops) whereupon the reaction went from cloudy white to clear yellow. Once TLC monitoring indicated the disappearance of the amine starting material, the reaction was again cooled to -40 °C and another portion of 2-aminobenzimidazole (4 mmols) was slowly added. The reaction was left to come to RT and then heated to 40 °C for 30 mins, after which TLC indicated no remaining benzimidazole starting material, after which it was then brought back to RT. Aniline (500 mg, 5.4 mmols) was added and the solution stirred for a further 30 mins before commencing the work-up. The organic product was extracted into EtOAc (3 x 50ml) from a solution of brine and sodium carbonate and then the combined EtOAc fractions dried with $MgSO_4$. The EtOAc was removed on a rotoevaporator and then DCM:MeOH (1:1) was added to the yellow solid and briefly heated. The undissolved solid was filtered off and the process repeated twice using acetone to remove the dissolved impurities. The product was then recrystallized in hot pyridine to yield **8** as yellow crystals (200 mg, 10%); Mp 339-342 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ 12.63 (br. s, 2H), 10.72 (s, 1H), 9.02 - 9.18 (m, 2H), 8.91 (t, *J* = 1.9 Hz, 1H), 7.80 (dd, *J* = 1.1, 8.6 Hz, 2H), 7.42 - 7.48 (m, 2H), 7.35 - 7.41 (m, 2H), 7.18 - 7.23 (m, 2H), 7.11 - 7.17 (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6) δ 163.5 (2), 147.8, 139.7, 138.7 (2), 136.5, 133.9, 130.1, 128.8, 126.0, 124.5, 124.3, 122.6, 120.5, 112.2; HRMS-ES$^-$: Observed 400.1052 $(M-H)^+$; Calculated 400.1046 for $C_{21}H_{14}N_5O_4$; HPLC: 93.2%.

_N_,_N_,_N_-Triphenyl-benzene-1,3,5-tricarboxamide (10)[329]



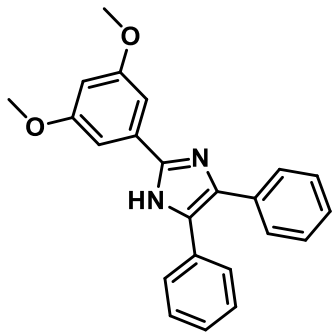
Benzene-1,3,5-tricarbonyl trichloride (**9**, 265 mg, 1 mmol) was dissolved in dry DCM:pyridine (4:1) and then aniline (279 mg, 3 mmols) was added. The reaction was stirred at RT for 1 h after which the precipitated product was filtered off and dried. The white solid was recrystallized in MeOH and DCM to give **10** as white crystals (240 mg, 55%); Mp 314-315 °C; Lit Mp[329] 327-329 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 10.55 (br. s, 3H) 8.69 (s, 3H) 7.81 (m, 6H) 7.40 (m, 6H) 7.11 (m, 3H); ^{13}C NMR (151 MHz, DMSO-d_6) δ ppm 160.7, 133.1, 130.2, 129.2, 128.1, 123.4, 121.4; HRMS-ES^{-}: Observed 434.1484 $(M-H)^{+}$; Calculated 434.1505 for $C_{27}H_{20}N_3O_3$; HPLC: 96.7%.

8.2.2.2 Triarylimidazoles

General procedure for the synthesis of triarylimidazoles (15a-d)

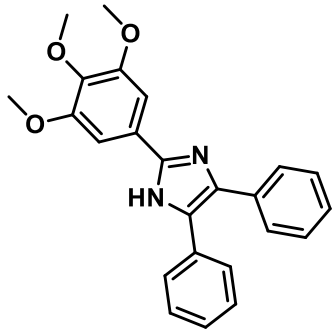
The appropriate benzaldehyde (**13**) and benzil (**14**) in a molar ratio of 1:1 were refluxed with excess ammonium acetate (NH_4OAc) in AR-grade MeOH for 4 h. Upon completion of the reaction, indicated by TLC analysis, it was cooled to RT. For compounds **15a**-**c**, the reactions produced a precipitate, which was collected by vacuum filtration to afford the desired product. For compound **15d**, the contents of the reaction flask were added to a saturated solution of Na_2CO_3 and the organic components extracted with EtOAc (3 x 10 ml). The combined organic fractions were dried with $MgSO_4$ before concentrating the solution under vacuum. All compounds (**15a**-**5d**) were dried under high vacuum for 2 h to yield the desired product.

2-(3,5-Dimethoxyphenyl)-4,5-diphenyl-1*H*-imidazole (15a)[274]



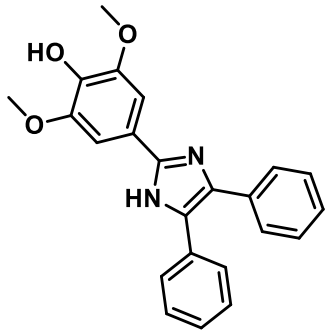
Scale: **13a** (166 mg, 1 mmol), benzil (210 mg, 1 mmol) and NH_4OAc (230 mg, 3 mmol). Yield of **15a**: 311 mg, 87%. Recrystallisation from absolute EtOH afforded **15a** as white needles: Mp 254-255 °C; Lit Mp[274] 254-256 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 12.62 (br. s., 1H) 7.02 - 7.72 (m, 12H) 6.50 (br. s., 1H) 3.80 (br. s., 6H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 161.2, 145.8, 137.5, 135.6, 132.6, 131.6, 129.1, 129.0, 128.8, 128.6, 128.3, 127.6, 127.0, 103.6, 101.1, 55.9; HRMS-ES$^+$: Observed 357.1603 $(M+H)^+$; Calculated 357.1603 for $C_{23}H_{21}N_2O_2$; HPLC: 98.5%.

4,5-Diphenyl-2-(3,4,5-trimethoxyphenyl)-1*H*-imidazole (15b)[264]



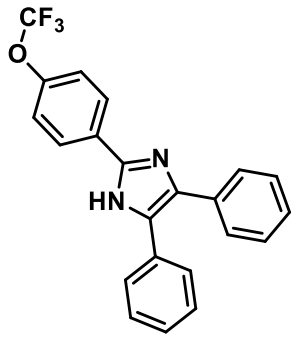
Scale: **13b** (196 mg, 1 mmol), benzil (210 mg, 1 mmol) and NH_4OAc (230 mg, 3 mmols). Yield of **15b**: 229 mg, 59%. Recrystallisation from absolute EtOH afforded **15b** as white needles: Mp 254-256 °C; Lit Mp[264] 261-262 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 12.55 (br. s., 1H) 6.60 - 7.92 (m, 12H) 3.84 (s, 6H) 3.69 (s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 153.6, 145.9, 138.3, 137.4, 135.6, 131.7, 129.2, 129.0, 128.6, 128.6, 128.3, 127.5, 127.0, 126.4, 103.2, 60.6, 56.5; HRMS-ES$^+$: Observed 387.1711 $(M+H)^+$; Calculated 387.1709 for $C_{24}H_{23}N_2O_3$; HPLC: 98.3%.

4-(4,5-Diphenyl-1H-imidazol-2-yl)-2,6-dimethoxyphenol (15c)[275]



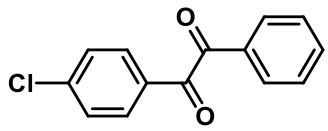
Scale: **13c** (364 mg, 2 mmol), benzil (420 mg, 2 mmol) and NH_4OAc (1.54 g, 20 mmols). Yield of **15c**: 601 mg, 81%. Recrystallisation from absolute EtOH afforded **15c** as an off-white powder: Mp 286-287 °C; Lit Mp 290-292 °C; ^{1}H NMR (300 MHz, DMSO-d_6) δ ppm 12.42 (s, 1H), 8.58 (s, 1H), 7.34 - 7.62 (m, 10H), 7.15 - 7.34 (m, 2H), 3.85 (s, 6H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 148.6, 146.5, 137.1, 136.8, 135.8, 131.9, 129.1, 128.9, 128.6, 128.1, 128.0, 127.5, 126.8, 121.3, 103.5, 56.6; HRMS-ES$^+$: Observed 373.1558 $(M+H)^+$; Calculated 373.1552 for $C_{23}H_{21}N_2O_3$; HPLC: 99.1%.

4,5-Diphenyl-2-(4-(trifluoromethoxy)phenyl)-1H-imidazole (15d)



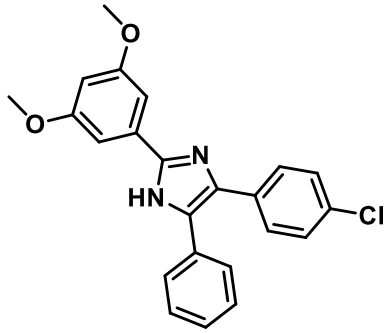
Scale: **13d** (190 mg, 1 mmol), benzil (210 mg, 1 mmol) and NH_4OAc (770 mg, 10 mmols). Yield of **15d**: 321 mg, 84%. Recrystallisation from absolute EtOH afforded **15d** as a white powder: Mp 234-235 °C; ^{1}H NMR (400 MHz, Acetone) δ ppm 11.66 (br. s, 1H) 8.11 (d, *J* = 8.5, 2H) 7.54 (d, *J* = 8.5, 2H) 7.03 - 7.47 (m, 10H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 148.6, 144.7, 137.9, 135.5, 129.3 (q, *J*CF = 144.9 Hz, CF_3), 129.2, 128.9, 128.7, 127.6, 127.5, 121.8, 119.3; HRMS-ES$^+$: Observed 381.1218 $(M+H)^+$; Calculated 381.1215 for $C_{22}H_{16}N_2OF_3$; HPLC: 98.55%.

4-Chlorobenzil (14e)



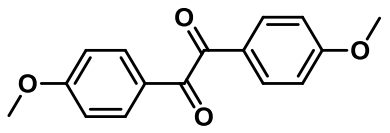
4-Chlorobenzoin (**12e**, 123 mg, 0.5 mmols) was stirred in dry THF (5 ml) and cooled to 0 °C before sodium hydride (60% w/w dispersion in mineral oil, 40 mg, 1 mmol) was added. The dark-blue solution was heated slowly to RT for 15 min upon which the mixture went yellow and TLC indicated consumption of the benzoin. The mixture was quenched with $NaHCO_3$ solution (10 ml, 0.1 M), then extracted with EtOAc (3 x 10 ml) and washed with brine. The organic layer was dried with $MgSO_4$ before being concentrated under vacuum. Purification by column chromatography was carried out in 30% EtOAc/hexane to yield **14e** which was then recrystallised from hot hexane, resulting in a white powder (37 mg, 30%). ^{1}H NMR (300 MHz, Acetone-d_6) δ ppm 8.23 - 8.51 (m, 2H), 7.74 - 8.13 (m, 4H), 7.58 - 7.72 (m, 3H).

5-(4-Chlorophenyl)-2-(3,5-dimethoxyphenyl)-4-phenyl-1H-imidazole (15e)



4-Chlorobenzil (**14e**, 29 mg, 0.12 mmols), 3,5-dimethoxybenzaldehyde (**13e**, 20 mg, 0.12 mmol) and NH_4OAc (184 mg, 2.4 mmols) were refluxed at 70 °C in MeOH (3 ml) with stirring overnight. After cooling to RT, the MeOH was removed under reduced pressure followed by the addition of EtOAc (5 ml) and brine (5 ml). The product was extracted into EtOAc (3 x 5 ml) and the organic layer dried with $NaSO_4$. The remaining solution was purified by column chromatography in 30% EtOAc/hexane and then recrystallized from DCM/hexane to give **15e** as a white powder (30 mg, 65%); Mp 218-219 °C; ^{1}H NMR (400 MHz, MeOD) δ ppm 7.41 - 7.55 (m, 4H), 7.35 (m, 5H), 7.21 (d, *J* = 2.3 Hz, 2H), 6.53 (t, *J* = 2.3 Hz, 1H), 3.85 (s, 6H); ^{13}C NMR (101 MHz, MeOD) δ ppm 161.4, 146.9, 145.3, 145.2, 144.7, 132.8, 131.5, 129.4, 128.3, 128.2, 128.2, 128.1, 127.4, 103.5, 101.1, 54.6; HRMS-ES$^+$: Observed 391.1222 $(M+H)^+$; Calculated 391.1213 for $C_{23}H_{20}N_2O_2Cl$; HPLC: 97.89%.

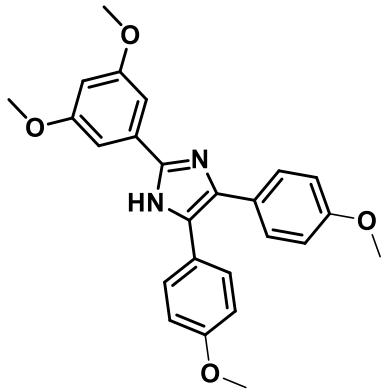
4,4'-Dimethoxybenzil (14f)



4,4'-Dimethoxybenzoin (**12f**, 545 mg, 2 mmol) was stirred in dry THF (20 ml) at 0 °C before sodium hydride (60% w/w dispersion in mineral oil, 160 mg 2.4 mmols) was added. The dark-blue solution was allowed to reach RT over 10 mins and then stirred for a further 3 h, by which time TLC showed complete consumption of the starting material. The reaction was quenched with $NaHCO_3$ (40 ml, 0.1 M), extracted into EtOAc (3 x 40 ml) and washed with brine before drying of the organic layer over $MgSO_4$. The solution was concentrated under

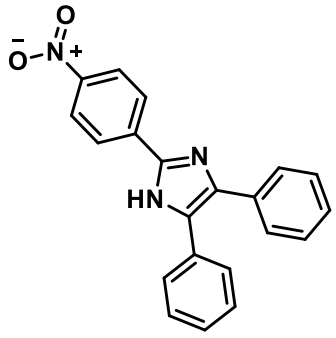
vacuum and then recrystallised from hot EtOAc/hexane to yield **14f** as a white powder (359 mg, 66%), ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 7.94 (d, *J* = 8.8 Hz, 4H), 6.97 (d, *J* = 8.8 Hz, 4H), 3.88 (s, 6H).

2-(3,5-Dimethoxyphenyl)-4,5-bis(4-methoxyphenyl)-1H-imidazole (15f)



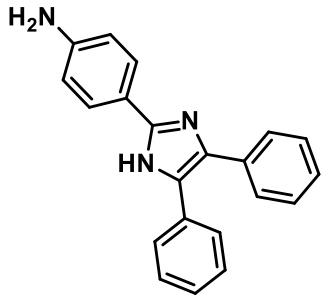
4,4'-Dimethoxybenzil (**14f**, 135 mg, 0.5 mmols), 3,5-dimethoxybenzaldehyde (**13f**, 83 mg, 0.5 mmols) and NH_4OAc (385 mg, 5 mmols) were refluxed at 80 °C in EtOH (7 ml) with stirring overnight. After cooling to RT, the EtOH was removed under reduced pressure and Na_2CO_3 (10 ml) was added to the remaining mixture. The product was extracted into EtOAc (3 x 10 ml) and washed with brine (10 ml). The organic layer was dried over $MgSO_4$ before concentrating under vacuum. Purification of the residue was carried out using column chromatography in 40% EtOAc/hexane followed by recrystallization from EtOAc/hexane to give **15f** as a white powder (167 mg, ≥), Mp 190-191 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 12.44 (s, 1H), 7.45 (d, *J* = 8.4 Hz, 2H), 7.39 (d, *J* = 8.4 Hz, 2H), 7.24 (d, *J* = 2.3 Hz, 2H), 7.00 (d, *J* = 8.6 Hz, 2H), 6.86 (d, *J* = 8.6 Hz, 2H), 6.47 (t, *J* = 2.3 Hz, 1H), 3.80 (s, 6H), 3.79 (br. s., 3H), 3.73 (br. s, 3H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 161.2, 159.3, 158.5, 145.1, 136.9, 132.8, 130.3, 128.7, 128.3, 127.7, 124.0, 114.6, 114.1, 103.5, 100.9, 55.8, 55.7, 55.5; HRMS-ES$^+$: Observed 417.1814 $(M+H)^+$; Calculated 417.1814 for $C_{25}H_{25}N_2O_4$; HPLC: 98.0%.

2-(4-Nitrophenyl)-4,5-diphenyl-1*H*-imidazole (15g)[260]



To a well stirred mixture of 4-nitrobenzaldehyde (**13g**, 600 mg, 3.9 mmol) and benzil (**14g**, 800 mg, 3.8 mmol) in acetic acid (25 ml), ammonium acetate (5 g, 0.06 mols) was added. The reaction was refluxed at 118 °C for 4 h. The solution was poured into ice-water and ammonia added to the acidic solution while stirring, until the pH was neutral. The dark orange precipitate was vacuum-filtered off and washed with water (3 x 20 ml) whilst on the vacuum. The precipitate was then stirred in hot absolute EtOH and an insoluble brown precipitate was filtered off and discarded. On cooling, **15g** (1.2 g, 93%) was obtained as a bright-orange powder, Mp 228-230 °C; Lit Mp[330] 250-251 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 13.11 (br. s, 1H) 8.35 (d, *J* = 7.9 Hz, 4H), 8.35 (d, *J* = 7.9 Hz, 4H), 7.18 - 7.63 (m, 10H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 173.6, 134.6, 133.0, 131.3, 129.8, 129.7, 129.6, 129.5, 129.2, 129.0, 128.7, 128.2, 127.1, 127.0, 126.6; HRMS-ES^{+}: Observed 342.1243 $(M+H)^{+}$; Calculated 342.1243 for $C_{21}H_{16}N_3O_2$; HPLC: 98.75%.

4-(4,5-Diphenyl-1H-imidazol-2-yl)aniline (16)[260]



2-(4-Nitrophenyl)-4,5-diphenyl-1*H*-imidazole (**15g**, 610 mg, 1.8 mmols), was dissolved in EtOH /water (2:1 v/v, 25 ml) before the addition of iron filings (785 mg, 7.8 equiv) and HCl (32% w/w, 4 drops). The mixture was refluxed at 80 °C for 6 h under anaerobic conditions, cooled and then the iron scruff filtered off. Removal of excess solvents was carried out on a rotoevaporator under reduced pressure at 60 °C to yield a pale brown residue. Recrystallisation from hexane, DCM and a few drops of EtOH resulted in **16** (398 mg, 71%) as an off-white solid, Mp 202-203 °C; Lit Mp[330] 211-212 °C; ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 12.11 (s, 1H), 7.69 (d, *J* = 8.7 Hz, 2H), 7.47 (m, 2H), 7.41 (m, 2H), 7.35 (m Hz, 2H), 7.27 (m, 1H), 7.2 (m, 2H), 7.14 (m, 1H), 6.57 (d, *J* = 8.7 Hz, 2H), 5.26 (m, 2H); ^{13}C NMR (101 MHz, DMSO-d_6) δ ppm 149.0, 146.8, 136.3, 135.6, 131.4, 128.5, 128.1, 128.0, 127.2, 127.0, 126.7, 126.4, 126.1, 118.2, 113.5; HRMS-ES^{+}: Observed 312.1501 $(M+H)^{+}$; Calculated 312.1501 for $C_{21}H_{18}N_3$; HPLC: 97.92%.

REFERENCES

1. Bruce-Chwatt, L. J., Chemotherapy of Malaria, *World Health Organisation*, Geneva, **1981**, pp. 9-19.
2. Despommier, D. D.; Gwadz, R. W.; Hotez, P. J.; Knirsch, C. A., Parasitic Diseases, *Apple Trees Productions*, New York, **2005**, pp. 363.
3. Cox, F. E. G. *Parasite. Vector.* **2010**, *3*, 1.
4. Bruce-Chwatt, L. J., History of malaria from prehistory to eradication. In Malaria: Principles and Practice of Malariology, WH, W., I, M., *Churchill Livingstone*, Edinburgh, **1988**, pp. 1-59.
5. Laveran, C. L. *Rev. Infect. Dis.* **1982**, *4*, 908.
6. Lancisi, G. M., De noxiis paludum effluviis eorumque remediis, *J. M. Salvioni*, Rome, **1717**.
7. Ross, R. *Br. Med. J.* **1897**, *2*, 1786.
8. Perkins, S. L.; Austin, C. C. *J. Parasitol.* **2009**, *95*, 424.
9. Cox-Singh, J.; Davis, T. M. E.; Lee, K.-S.; Shamsul, S. S. G.; Matusop, A.; Ratnam, S.; Rahman, H. A.; Conway, D. J.; Singh, B. *Clin. Infect. Dis.* **2008**, *46*, 165.
10. World Health Organisation World Malaria Report 2014, **2014**, Switzerland, http://www.who.int/malaria/publications/world_malaria_report_2014/wmr-2014-no-profiles.pdf?ua=1 (January 2015).
11. Gething, P. W.; Patil, A. P.; Smith, D. L.; Guerra, C. A.; Elyazar, I. R.; Johnston, G. L.; Tatem, A. J.; Hay, S. I. *Malar. J.* **2011**, *10*, 378.
12. White, N. J. *J. Clin. Invest.* **2004**, *113*, 1084.
13. Greenwood, B. M.; Fidock, D. A.; Kyle, D. E.; Kappe, S. H. I.; Alonso, P. L.; Collins, F. H.; Duffy, P. E. *J. Clin. Invest.* **2008**, *118*, 1266.
14. Foley, M.; Tilley, L. *Pharmacol. Therapeut.* **1998**, *79*, 55.
15. Sherman, I. W. *Bull. World Health Organ.* **1977**, *55*, 265.
16. Lew, V. L.; Tiffert, T.; Ginsburg, H. *Blood* **2003**, *101*, 4189.
17. Krugliak, M.; Zhang, J.; Ginsburg, H. *Mol. Biochem. Parasitol.* **2002**, *119*, 249.
18. Egan, T. J. *Targets* **2003**, *2*, 117.
19. Goldberg, D. E.; Slater, A. F. G.; Cerami, A.; Henderson, G. B. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 2931.
20. Egan, T. J. *J. Inorg. Biochem.* **2008**, *102*, 1288.
21. Hitchings, G. H. *Angew. Chem. Int. Ed.* **1989**, *28*, 879.
22. Nirmalan, N.; Wang, P.; Sims, P. F.; Hyde, J. E. *Mol. Microbiol.* **2002**, *46*, 179.
23. Looareesuwan, S.; Chulay, J. D.; Canfield, C. J.; Hutchinson, D. B. *Am. J. Trop. Med. Hyg.* **1999**, *60*, 533.
24. Vinetz, J. M.; Clain, J.; Bounkeua, V.; Eastman, R. T.; Fidock, D., Chemotherapy of malaria In: Goodman & Gilman's the pharmacological basis of therapeutics., Brunton, L. L., Ed., *The McGraw-Hill Companies Inc*, New York, **2011**, pp. 1383–1418.
25. Faurant, C. *Parasite.* **2011**, *18*, 215.
26. Olliaro, P. L.; Haynes, R. K.; Meunier, B.; Yuthavong, Y. *Trends Parasitol.* **2001**, *17*, 122.
27. Haynes, R. K.; Cheu, K.; N'Da, D.; Coghi, P.; Monti, D. *Infect. Disord. Drug Targets* **2013**, *13*, 217.
28. Posner, G. H.; Cumming, J. N.; Ploypradith, P.; Oh, C. H. *J. Am. Chem. Soc.* **1995**, *117*, 5885.
29. Asawamahasakda, W.; Ittarat, I.; Pu, Y.-M.; Ziffer, H.; Meshnick, S. R. *Antimicrob. Agents Chemother.* **1994**, *38*, 1854.
30. Pandey, A. V.; Tekwani, B. L.; Singh, R. L.; Chauhan, V. S. *J. Biol. Chem.* **1999**, *274*, 19383.
31. Meshnick, S. R.; Thomas, A.; Ranz, A.; Xu, C.-M.; Pan, H.-Z. *Mol. Biochem Parasitol.* **1991**, *49*, 181.
32. Meshnick, S. R.; Taylor, T. E.; Kamchonwongpaisan, S. *Microbiol. Mol. Biol. Rev.* **1996**, *60*, 301.
33. Asawamahasakda, W.; Ittarat, I.; Chang, C.-C.; McElroy, P.; Meshnick, S. R. *Mol. Biochem. Parasitol.* **1994**, *67*, 183.
34. Eckstein-Ludwig, U.; Webb, R. J.; Van Goethem, I. D. A.; East, J. M.; Lee, A. G.; Kimura, M.; O'Neill, P. M.; Bray, P. G.; Ward, S. A.; Krishna, S. *Nature* **2003**, *424*, 957.
35. Li, W.; Mo, W.; Shen, D.; Sun, L.; Wang, J.; Lu, S.; Gitschier, J. M.; Zhou, B. *PLoS Genet.* **2005**, *1*, 36.
36. Olliaro, P. *Pharmacol. Therapeut.* **2001**, *89*, 207.

37. Song, C. E., An Overview of Cinchona Alkaloids in Chemistry, in Cinchona Alkaloids in Synthesis and Catalysis: Ligands, Immobilization and Organocatalysis, *Wiley-VCH Verlag GmbH & Co. KGaA*, Weinheim, Germany, **2009**, Chapter 1.
38. Coatney, G. R. *Am. J. Trop. Med. Hyg.* **1963**, *12*, 121.
39. Foote, S. J.; Cowman, A. F. *Acta Trop.* **1994**, *56*, 157.
40. Le Bras, J.; Deloron, P.; Ricour, A.; Andrieu, B.; Savel, J.; Couland, J. P. *Exp. Parasitol.* **1983**, *56*, 9.
41. Laufer, M. K.; Thesing, P. C.; Eddington, N. D.; Masonga, R.; Dzinjalamala, F. K.; Takala, S. L.; Taylor, T. E.; Plowe, C. *N. Engl. J. Med.* **2006**, *355*, 1959.
42. Ecker, A.; Lehane, A. M.; Clain, J.; Fidock, D. A. *Trends Parasitol.* **2012**, *28*, 504.
43. Achan, J.; Talisuna, A. O.; Erhart, A.; Yeka, A.; Tibenderana, J. K.; Baliraine, F. N.; Rosenthal, P. J.; D'Alessandro, U. *Malar. J.* **2011**, *10*, 144.
44. Giboda, M.; Denis, M. B. *J. Trop. Med. Hyg.* **1988**, *91*, 205.
45. Hüttinger, F.; Satimai, W.; Wernsdorfer, G.; Wiedermann, U.; Congpuong, K.; Wernsdorfer, W. H. *Wien. Klin. Wochenschr.* **2010**, *122*, 52.
46. White, N. J. *Br. Med. J.* **1994**, *308*, 286.
47. Cowman, A. F.; Galatis, D.; Thomson, J. K. *Proc. Nati. Acad. Sci. USA* **1994**, *91*, 1143.
48. Lim, A. S.; Galatis, D.; Cowman, A. F. *Exp. Parasitol.* **1996**, *83*, 295.
49. Noedl, H.; Se, Y.; Schaecher, K.; Smith, B. L.; Socheat, D.; Fukuda, M. M. *N. Engl. J. Med.* **2008**, *359*, 2619.
50. Talisuna, A. O.; Karema, C.; Ogutu, B.; Juma, E.; Logedi, J.; Nyandigisi, A.; Mulenga, M.; Mbacham, W. F.; Roper, C.; Guerin, P. J.; D'Alessandro, U.; Snow, R. W. *Lancet Infect. Dis.* **2012**, *12*, 888.
51. Yayon, A.; Cabantchik, Z. I.; Ginsburg, H. *EMBO J.* **1984**, *3*, 2695.
52. Fitch, C. D.; Yunis, N. G.; Chevli, R.; Gonzales, Y. *J. Clin. Invest.* **1974**, *54*, 24.
53. Krogstad, D. J.; Gluzman, I. Y.; Kyle, D. E.; Oduola, A. M.; Martin, S. K.; Milhous, W. K.; Schlesinger, P. H. *Science* **1987**, *238*, 1283.
54. Fojo, A.; Akiyama, S.; Gottesman, M. M.; Pastan, I. *Cancer Res.* **1985**, *45*, 3002.
55. Gottesman, M. M.; Pastan, I. *Annu. Rev. Biochem.* **1993**, *62*, 385.
56. Martin, S. K.; Oduola, A. M.; Milhous, W. K. *Science* **1987**, *235*, 899.
57. Ginsburg, H.; Krugliak, M. *Biochem. Pharmacol.* **1992**, *43*, 63.
58. Martiney, J. A.; Cerami, A.; Slater., A. F. G. *J. Biol. Chem.* **1995**, *270*, 22393.
59. Bray, P. G.; Mungthin, M.; Ridley, R. G.; Ward, S. A. *Mol. Pharm.* **1998**, *54*, 170.
60. Fidock, D. A.; Nomura, T.; Talley, A. K.; Cooper, R. A.; Dzekunov, S. M.; Ferdig, M. T.; Ursos, L. M. B.; Sidhu, A. S.; Naude, B.; Deitsch, K. W.; Su, X.; Wootton, J. C.; Roepe, P. D.; Wellems, T. E. *Mol. Cell* **2000**, *6*, 861.
61. Shrivastava, S. K.; Gupta, R. K.; Mahanta, J.; Dubey, M. L. *PLoS One* **2014**, *9*, 1.
62. Picot, S.; Olliaro, P.; de Monbrison, F.; Bienvenu, A. L.; Price, R. N.; Ringwald, P. *Malar. J.* **2009**, *8*, 89.
63. Lekostaj, J. K.; Natarajan, J. K.; Paguio, M. F.; Wolf, C.; Roepe, P. D. *Biochem.* **2008**, *47*, 10394.
64. Gligorijevic, B.; Bennett, T.; McAllister, R.; Urbach, J. S.; Roepe, P. D. *Biochemistry* **2006**, *45*, 12411.
65. Reed, M. B.; Saliba, K. J.; Caruana, S. R.; Kirk, K.; Cowman, A. *Nature* **2000**, *403*, 906.
66. Cowman, A. F. *Int. J. Parasitol.* **2001**, *31*, 871.
67. Juge, N.; Moriyama, S.; Miyaji, T.; Kawakami, M.; Iwai, H.; Fukui, T.; Nelson, N.; Omote, H.; Moriyama, Y. *Proc. Natl. Acad. Sci.* **2015**, *112*, 3356.
68. Martin, R. E.; Marchetti, R. V.; Cowan, A. I.; Howitt, S. M.; Bröer, S.; Kirk, K. *Science* **2009**, *325*, 1680.
69. Sanchez, C. P.; Dave, A.; Stein, W. D.; Lanzer, M. *Int. J. Parasitol.* **2010**, *40*, 1109.
70. Trager, W.; Jensen, J. B. *Sci. Cult.* **1976**, *193*, 673.
71. LeRoux, M.; Lakshmanan, V.; Daily, J. P. *Trends Parasitol.* **2009**, *25*, 474.
72. Maier, A. G.; Rug, M., In Vitro Culturing Plasmodium falciparum Erythrocytic Stages in Malaria: Methods in Molecular Biology, Walker, J. M.; Ménard, R., Eds., *Springer Science*, London, **2013**, Vol. 923, pp. 3-15.
73. Giemsa, G. *Centralblatt für Bakteriol* **1094**, *32*, 307.
74. Orjih, A. U. *Exp. Biol. Med.* **2008**, *233*, 1359.
75. Lambros, C.; Vanderberg, J. P. *J. Parasitol.* **1979**, *65*, 418.

76. Desjardins, R. E.; Canfield, C. J.; Haynes, J. D.; Chulay, J. D. *Antimicrob. Agents Chemother.* **1979**, *16*, 710.
77. Elabbadi, N.; Ancelin, M. L.; Vial, H. J. *Antimicrob. Agents Chemother.* **1992**, *36*, 50.
78. Iber, P. K.; Pavanand, K.; Wilks, N. E.; Colwell, E. J. *J. Med. Assoc. Thai.* **1975**, *58*, 559.
79. Makler, M. T.; Ries, J. M.; Williams, J. A.; Bancroft, J. E.; Piper, R. C.; Gibbins, B. L.; Hinrichs, D. J. *Am. J. Trop. Med. Hyg.* **1993**, *48*, 739.
80. Druilhe, P.; Moreno, A.; Blanc, C.; Brasseur, P. H.; Jacquier, P. *Am. J. Trop. Med. Hyg.* **2001**, *64*, 233.
81. Baniecki, M. L.; Wirth, D. F.; Clardy, J. *Antimicrob. Agents Chemother.* **2007**, *51*, 716.
82. Bennett, T. N.; Paguio, M.; Gligorijevic, B.; Seudieu, C.; Kosar, A. D.; Davidson, E.; Roepe, P. D. *Antimicrob. Agents Chemother.* **2004**, *48*, 1807.
83. Corbett, Y.; Herrera, L.; Gonzalez, J.; Cubilla, L.; Capson, T.; Coley, P.; Kursar, T.; Romero, L.; Ortega-Barria, E. *Am. J. Trop. Med. Hyg.* **2004**, *70*, 119.
84. Johnson, J. D.; Dennull, R. A.; Gerena, L.; Lopez-Sanchez, M.; Roncal, N. E.; Waters, N. C. *Antimicrob. Agents Chemother.* **2007**, *51*, 1926.
85. Hahn, F. E.; O'Brien, R. L.; Ciak, J.; Allison, J. L.; Olenick, J. G. *Mil. Med.* **1966**, *131*, 1071.
86. Cohen, S. N.; Yielding, K. *J. Biol. Chem.* **1965**, *240*, 3123.
87. Peters, W., Chemotherapy and Drug Resistance in Malaria., *Academic Press*, London., **1970**, Vol. 2, pp. 1100.
88. Vander Jagt, D. L.; Hunsaker, L. A.; Campos, N. M. *Mol. Biochem. Parasitol.* **1986**, *18*, 389.
89. Zarchin, S.; Krugliak, M.; Ginsburg, H. *Biochem. Pharmacol.* **1986**, *35*, 2435.
90. Rosenthal, P. J.; McKerrow, J. H.; Aikawa, M.; Nagasawa, H.; Leech, J. H. *J. Clin. Invest.* **1988**, *82*, 1560.
91. McKerrow, J. H.; Sun, E.; Rosenthal, P. J.; Bouvier, J. *Annu. Rev. Microbiol.* **1993**, *47*, 821.
92. Hellerstein, S.; Spees, W.; Surapathana, L. O. *J. Lab. Clin. Med.* **1970**, *76*, 10.
93. Peters, W. *Ann. Soc. Belg. Med. Trop.* **1965**, *45*, 365.
94. Fitch, C. D. *Parasitol. Today* **1986**, *2*, 330.
95. McChesney, E. W.; Fitch, C. D., 4-Aminoquinolines. In: Antimalarial Drugs II: Current Antimalarials and New Drug Developments, *Springer-Verlag*, Berlin, **1984**, pp. 30–60.
96. Slater, A.; Cerami, A. *Nature* **1992**, *355*, 167.
97. Dorn, A.; Vippagunta, S. R.; Matile, H.; Jaquet, C.; Vennerstrom, J. L.; Ridley, R. G. *Biochem. Pharma.* **1998**, *55*, 727.
98. Hawley, S. R.; Bray, P. G.; O'Neill, P. M.; Park, B. K.; Ward., S. A. *Biochem. Pharmacol.* **1996**, *52*, 723.
99. Hawley, S. R.; Bray, P. G.; Mungthin, M.; Atkinson, J. D.; O'Neill, P. M.; Ward, S. A. *Antimicrob. Agents Chemother.* **1998**, *42*, 682.
100. Vippagunta, S. R.; Dorn, A.; Matile, H.; Bhattacharjee, A. K.; Karle, J. M.; Ellis, W. Y.; Ridley, R. G.; Vennerstrom, J. L. *J. Med. Chem.* **1999**, *42*, 4630.
101. Kaschula, C. H.; Egan, T. J.; Hunter, R.; Basilico, N.; Parapini, S.; Taramelli, D.; Pasini, E.; Monti, D. *J. Med. Chem.* **2002**, *45*, 3531.
102. Parapini, S.; Basilico, N.; Pasini, E.; Egan, T. J.; Olliaro, P.; Taramellia, D.; Monti, D. *Exp. Parasitol.* **2000**, *96*, 249.
103. Ncokazi, K. K.; Egan, T. J. *Anal. Biochem.* **2005**, *338*, 306.
104. Zishiri, V. K.; Joshi, M. C.; Hunter, R.; Chibale, K.; Smith, P. J.; Summers, R. L.; Martin, R. E.; Egan, T. J. *J. Med. Chem.* **2011**, *54*, 6956.
105. Kalkanidis, M.; Klonis, N.; Tilley, L.; Deady, L. W. *Biochem. Pharma.* **2002**, *63*, 833.
106. Kelly, J. X.; Winter, R.; Peyton, D. H.; Hinrichs, D. J.; Riscoe, M. *Antimicrob. Agents Chemother.* **2002**, *46*, 144.
107. Lu, W.-J.; Wicht, K. J.; Wang, L.; Imai, K.; Mei, Z.-W.; Kaiser, M.; Sayed, I. E. T. E.; Egan, T. J.; Inokuchi, T. *Eur. J. Med. Chem.* **2013**, *64*, 498.
108. Wang, N.; Wicht, K. J.; Wang, L.; Lu, W.-J.; Misumi, R.; Wang, M.-Q.; Gokha, A. A. A. E.; Kaiser, M.; Sayed, I. E. T. E.; Egan, T. J.; Inokuchi, T. *Chem. Pharm. Bull.* **2013**, *61*, 1282.
109. Wang, N.; Wicht, K. J.; Imai, K.; Ngoc, T. A.; Wang, M.-Q.; Kaiser, M.; Egan, T. J.; Inokuchi, T. *Bioorg. Med. Chem.* **2014**, *22*, 2629.
110. Wang, N.; Wicht, K. J.; Shaban, E.; Ngoc, T. A.; Wang, M.-Q.; Hayashi, I.; Hossain, M. I.; Takemasa, Y.; Kaiser, M.; Sayed, I. E. T. E.; Egan, T. J.; Inokuchi, T. *Med. Chem. Commun.* **2014**, *5*, 927.
111. Gligorijevic, B.; McAllister, R.; Urbach, J. S.; Roepe, P. D. *Biochemistry* **2006**, *45*, 12400.

112. Burgess, S. J.; Kelly, J. X.; Shomloo, S.; Wittlin, S.; Brun, R.; Liebmann, K.; Peyton, D. H. *J. Med. Chem.* **2010**, *53*, 6477.
113. Combrinck, J. M.; Mabotha, T. E.; Ncokazi, K. K.; Ambele, M. A.; Taylor, D.; Smith, P. J.; Hoppe, H. C.; Egan, T. J. *ACS Chem. Biol.* **2013**, *8*, 133.
114. Fitch, C. D. *Proc. Natl. Acad. Sci. USA.* **1969**, *64*, 1181.
115. Fitch, C. D. *Antimicrob. Agents Chemother.* **1973**, *3*, 545.
116. Ceithaml, J.; Evans, E. A. *Arch. Biochem. Biophys.* **1946**, *10*, 397.
117. Marshall, E. K.; Dearborn, E. H. *J. Pharmacol.* **1946**, *88*, 142.
118. Colowick, S. P.; Womack, F. C. *J. Biol. Chem.* **1969**, *244*, 774.
119. Yayon, A.; Cabantchik, Z. I.; Ginsburg, H. *Proc. Natl. Acad. Sci. USA.* **1985**, *82*, 2784.
120. Bray, P. G.; Howells, R. E.; Ward, S. A. *Biochem. Pharmacol.* **1992**, *43*, 1219.
121. MacIntyre, A. C.; Cutler, D. J. *J. Pharm. Sci.* **1993**, *82*, 592.
122. Warhurst, D. C.; Craig, J. C.; Adagu, I. S.; Meyer, D. J.; Lee, S. Y. *Malar. J.* **2003**, *2*, 14.
123. Egan, T. J. *Mini Rev. Med. Chem.* **2001**, *1*, 113.
124. Geary, T. G.; Jensen, J. B.; Ginsburg, H. *Biochem. Pharmacol.* **1986**, *35*, 3805.
125. Bray, P. G.; Janneh, O.; Raynes, K. J.; Mungthin, M.; Ginsburg, H.; Ward, a. S. A. *J. Cell Biol.* **1999**, *145*, 363.
126. Scheibel, L. W.; Sherman, I. W., Malaria: Principles and Practice of Malariology, Wernsdorfer, W. H.; McGregor, I., Eds., *Churchill Livingstone*, Edinburgh, U.K., **1988**, pp. 219-252.
127. Janjua, R. M.; Schultka, R.; Goebbel, L.; Pait, T. G.; C.B., S. *Neurosurgery* **2010**, *66*, 758.
128. Virchow, R. *Arch. Pathol. Anatomie. Physiol. Klin. Med.* **1849**, *2*, 587.
129. Brown, W. H. *J. Exp. Med.* **1911**, *13*, 290.
130. Sullivan, D. J., Hemozoin: a Biocrystal Synthesized during the Degradation of Hemoglobin, *Biopolymers Online,* **2005**, 9, http://www.wiley-vch.de/books/biopoly/pdf_v09/vol09_13.pdf (March 2015).
131. Ashong, J. O.; Blench, I. P.; Warhurst, D. C. *Trans. R. Soc. Trop. Med. Hyg.* **1989**, *83*, 167.
132. Fitch, C. D.; Kanjananggulpan, P. *J. Biol. Chem.* **1987**, *262*, 15552.
133. Slater, A. F. G.; Swiggard, W. J.; Orton, B. R.; Flitter, W. D.; Goldberg, D. E.; Cerami, A.; Henderson, G. B. *Proc. Natl. Acad. Sci. USA.* **1991**, *88*, 325.
134. Bohle, D. S.; Debrunner, P.; Jordan, P. A.; Madsen, S. K.; Schulz, C. E. *J. Am. Chem. Soc.* **1998**, *120*, 8255.
135. Bohle, D. S.; Dinnebier, R. E.; Madsen, S. K.; Stephens, P. W. *J. Biol. Chem.* **1997**, *272*, 713.
136. Pagola, S.; Stephens, P. W.; Bohle, D. S.; Kosar, A. D.; Madsen, S. K. *Nature* **2000**, *404*, 307.
137. Sullivan, D. J.; Gluzman, I. Y.; Russell, D. G.; Goldberg, D. E. *Proc. Natl. Acad. Sci. USA.* **1996**, *93*, 11865.
138. Chugh, M.; Sundararaman, V.; Kumar, S.; Reddy, V. S.; Siddiqui, W. A.; Stuart, K. D.; Malhotra, P. *Proc. Natl. Acad. Sci. USA.* **2013**, *110*, 5392.
139. Egan, T. J.; Ross, D. C.; Adams, P. A. *FEBS Lett.* **1994**, *352*, 54.
140. Dorn, A.; Stoffel, R.; Matile, H.; Bubendorf, A.; Ridley, R. *Nature* **1995**, *374*, 269.
141. Dorn, A.; Vippagunta, S.; Matile, H.; Bubendorf, A.; Vennerstrom, J.; Ridley, R. *Biochem. Pharmacol.* **1998**, *55*, 737.
142. Bendrat, K.; Berger, B. J.; Cerami., A. *Nature* **1995**, *378*, 138.
143. Fitch, C. D.; Cai, G.-Z.; Chen, Y. F.; Shoemaker, J. D. *Biochim. Biophys. Acta.* **1999**, *1454*, 31.
144. Olafsona, K. N.; Ketchuma, M. A.; Rimera, J. D.; Vekilova, P. G. *Proc. Natl. Acad. Sci.* **2015**, *In Press*, www.pnas.org/cgi/doi/10.1073/pnas.1501023112.
145. Pisciotta, J. M.; Coppens, I.; Tripathi, A. K.; Scholl, P. F.; Shuman, J.; Bajad, S.; Shulaev, V.; Sullivan, D. J. *Biochem. J.* **2007**, *402*, 197.
146. Ambele, M. A.; Sewell, B. T.; Cummings, F. R.; Smith, P. J.; Egan, T. J. *Cryst. Growth Des.* **2013**, *13*, 4442.
147. Hoang, A. N.; Ncokazi, K. K.; de Villiers, K. A.; Wright, D. W.; Egan, T. J. *Dalton Trans.* **2010**, *39*, 1235.
148. Hoang, A. N.; Sandlin, R. D.; Omar, A.; Egan, T. J.; Wright, D. W. *Biochem.* **2010**, *49*, 10107.
149. Ambele, M. A.; Egan, T. J. *Malar. J.* **2012**, *11*, 337.
150. Kuter, D.; Benjamin, S. J.; Egan, T. J. *J. Inorg. Biochem.* **2014**, *133*, 40.
151. Sullivan, D. J.; Gluzman, I. Y.; Goldberg, D. E. *Science* **1996**, *271*, 219.
152. Hoang, A. N.; Sandlin, R. D.; Omar, A.; Egan, T. J.; Wright, D. W. *Biochemistry* **2010**, *49*, 10107.
153. Ridley, R. G. *Trends Microbiol.* **1996**, *4*, 253.

154. Buller, R.; Peterson, M. L.; Almarsson, O.; Leiserowitz, L. *Cryst. Growth Des.* **2002**, *2*, 553.
155. Gildenhuys, J.; Roex, T. l.; Egan, T. J.; de Villiers, K. A. *J. Am. Chem. Soc.* **2013**, *135*, 1037.
156. Leed, A.; DuBay, K.; Ursos, L. M. B.; Sears, D.; de Dios, A. C.; Roepe, P. D. *Biochemistry* **2002**, *41*, 10245.
157. de Villiers, K. A.; Marques, H. M.; Egan, T. J. *J. Inorg. Biochem.* **2008**, *102*, 1660.
158. de Villiers, K. A.; Gildenhuys, J.; Roex, T. l. *ACS Chem. Biol.* **2012**, *7*, 666.
159. Kurosawa, Y.; Satoh, T.; Dorn, A.; Matile, H.; Kitsuji-Shirane, M.; Hofheinz, W.; Kansy, M.; Ridley, R. G. *Antimicrob. Agents Chemother.* **2000**, *44*, 2638.
160. Carter, M. D.; Phelan, V. V.; Sandlin, R. D.; Bachmann, B. O.; Wright, D. W. *Comb. Chem. High T. Scr.* **2010**, *3*, 285.
161. Sandlin, R. D.; Carter, M. D.; Lee, P. J.; Auschwitz, J. M.; Leed, S. E.; Johnson, J. D.; Wright, D. W. *Antimicrob. Agents Chemother.* **2011**, *55*, 3363.
162. Rosen, M. J.; Kunjappu, J. T., Surfactants and Interfacial Phenomena, *John Wiley & Sons.*, New Jersey, **2012**, Chapter 1.
163. Domínguez, A.; Fernández, A.; González, N.; Iglesias, E.; Montenegro, L. *J. Chem. Educ.* **1997**, *74*, 1227.
164. Huy, N. T.; Uyen, D. T.; Maeda, A.; Trang, D. T. X.; Oida, T.; Harada, S.; Kamei, K. *Antimicrob. Agents Chemother.* **2007**, *51*, 350.
165. Bohle, D. S.; Helms, J. B. *Biochem. Biophys. Res. Commun.* **1993**, *193*, 504.
166. Egan, T. J.; Hempelmann, E.; Mavuso, W. W. *J. Inorg. Biochem.* **1999**, *73*, 101.
167. Akoyunoglou, J.-H. A.; Olcott, H. S.; Brown, W. D. *Biochemistry* **1963**, *2*, 1033.
168. Brown, S. B.; Dean, T. C.; Jones, P. *Biochem. J.* **1970**, *117*, 733.
169. Munro, O. Q.; Marques, H. M. *Inorg. Chem.* **1996**, *35*, 3768.
170. Marques, H. M.; Munro, O. Q.; Crawcour, M. L. *Inorg. Chim. Acta* **1992**, *196*, 221.
171. Gamo, F.-J.; Sanz, L. M.; Vidal, J.; de Cozar, C.; Alvarez, E.; Lavandera, J.-L.; Vanderwall, D. E.; Green, D. V. S.; Kumar, V.; Hasan, S.; Brown, J. R.; Peishoff, C. E.; Cardon, L. R.; Garcia-Bustos, J. F. *Nature* **2010**, *465*, 305.
172. Guiguemde, W. A.; Shelat, A. A.; Bouck, D.; Duffy, S.; Crowther, G. J.; Davis, P. H.; Smithson, D. C.; Connelly, M.; Clark, J.; Zhu, F.; Jimenez-Diaz, M. B.; Martinez, M. S.; Wilson, E. B.; Tripathi, A. K.; Gut, J.; Sharlow, E. R.; Bathurst, I.; Mazouni, F. E.; Fowble, J. W.; Forquer, I.; McGinley, P. L.; Castro, S.; Angulo-Barturen, I.; Ferrer, S.; Rosenthal, P. J.; DeRisi, J. L.; Sullivan Jr, D. J.; Lazo, J. S.; Roos, D. S.; Riscoe, M. K.; Phillips, M. A.; Rathod, P. K.; Van Voorhis, W. C.; Avery, V. M.; Guy, R. K. *Nature* **2010**, *465*, 311.
173. Munos, B., Can open-source R&D reinvigorate drug research?, *Nat. Rev. Drug Discov.,* Indiana, **2006**, www.nature.com/reviews/drugdisc (February 2015).
174. World Health Organisation. Antimalarial Drug Resistance, **2014**, Switzerland, http://www.who.int/malaria/areas/drug_resistance/overview/en/ (October 2014).
175. Zhang, J.-H.; Chung, T. D. Y.; Oldenburg, K. R. *J. Biomol. Screen.* **1999**, *4*, 67.
176. Swinney, D. C.; Anthony, J. *Nat. Rev. Drug Discov.* **2011**, *10*, 507.
177. Kotz, J. *SciBX* **2012**, *5*, 1.
178. Guiguemde, W. A.; Shelat, A. A.; Garcia-Bustos, J. F.; Diagana, T. T.; Gamo, F.-J.; Guy, R. K. *Chem. Biol.* **2012**, *19*, 116.
179. Cameron, A.; Read, J.; Tranter, R.; Winter, V. J.; Sessions, R. B.; Brady, R. L.; Vivas, L.; Easton, A.; Kendrick, H.; Croft, S. L.; Barros, D.; Lavandera, J. L.; Martin, J. J.; Risco, F.; García-Ochoa, S.; Gamo, F. J.; Sanz, L.; Leon, L.; Ruiz, J. R.; Gabarró, R.; Mallo, A.; Gómez de las Heras, F. *J. Biol. Chem.* **2004**, *279*, 31429.
180. Rush, M. A.; Baniecki, M. L.; Mazitschek, R.; Cortese, J. F.; Wiegand, R.; Clardy, J.; Wirth, D. F. *Antimicrob. Agents Chemother.* **2009**, *53*, 2564.
181. Plouffe, D.; Brinker, A.; McNamara, C.; Henson, K.; Kato, N.; Kuhen, K.; Nagle, A.; Adrian, F.; Matzen, J. T.; Anderson, P.; Nam, T.-g.; Gray, N. S.; Chatterjee, A.; Janes, J.; Yan, S. F.; Trager, R.; Caldwell, J. S.; Schultz, P. G.; Zhou, Y.; Winzeler, E. A. *Proc. Natl. Acad. Sci. USA.* **2008**, *105*, 9059.
182. Egan, T. J.; Hunter, R.; Kaschula, C. H.; Marques, H. M.; Misplon, A.; Walden, J. *J. Med. Chem.* **2000**, *43*, 283.
183. Zhu, C.; Cook, S. P. *J. Am. Chem. Soc.* **2012**, *134*, 13577.
184. Bryan, M., Silas Cook '99 Finds New Way to Synthesize Antimalarial Drug, *Reed,* Reed College Oregon, **2013**, 92, 1, http://www.reed.edu/reed_magazine/march2013/articles/features/cook.html (January 2014).

185. Graham, I. A.; Besser, K.; Blumer, S.; Branigan, C. A.; Czechowski, T.; Elias, L.; Guterman, I.; Harvey, D.; Isaac, P. G.; Khan, A. M.; Larson, T. R.; Li, Y.; Pawson, T.; Penfield, T.; Rae, A. M.; Rathbone, D. A.; Reid, S.; Ross, J.; Smallwood, M. F.; Segura, V.; Townsend, T.; Vyas, D.; Winzer, T.; Bowles, D. *Science* **2010**, *327*, 328.
186. Cohen, J.; Patel, S.; Singh, I.; Cutler, M.; Ansboro, R.; Grove, N.; Nguyen, T.; Laverty, M.; Yadav, P.; Bathurst, I.; Jagoe, G.; Wells, T.; Gabriel, S.; Pilloy, J.; Binoche, N.; Farabolini, P.; Farret, H.; Marchal, F.; Bosman, A.; Bowles, D.; Clayton, D.; Graham, I., Artemisinin Scenario Analysis, **2009**, *The Boston Consulting Group*, pp. 1-14, http://www.mmv.org/sites/default/files/uploads/docs/artemisinin/03b_Artemisinin_Scenario_Analysis_BCG.pdf (February 2015).
187. Corsello, M. A.; Garg, N. K. *Nat. Prod. Rep.* **2015**, *32*, 359.
188. Venture, M. f. M., Optimizing artemisinin production, **2015**, Medicines for Malaria Venture, http://www.mmv.org/artemisinin (January 2015).
189. Charman, S. A.; Arbe-Barnes, S.; Bathurst, I. C.; Brun, R.; Campbell, M.; Charman, W.; Chiu, F.; Chollet, J.; Craft, J.; Creek, D.; Dong, Y.; Matile, H.; Maurer, M.; Morizzi, J.; Nguyen, T.; Papastogiannidis, P.; Scheurer, C.; Shackleford, D.; Sriraghavan, K.; Stingelin, L.; Tang, Y.; Urwyler, H.; Wang, X.; White, K.; Wittlin, S.; Zhou, L.; Vennerstrom, J. L. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4400.
190. Bayes, T. *Philos. Trans. R. Soc. London* **1763**, *53*, 370.
191. Balfer, J.; Bajorath, J. *J. Chem. Inf. Model.* **2014**, *54*, 2451.
192. Xia, X.; Maliski, E. G.; Gallant, P.; Rogers, D. *J. Med. Chem.* **2004**, *47*, 4463.
193. Klon, A. E.; Glick, M.; Thoma, M.; Acklin, P.; Davies, J. W. *J. Med. Chem.* **2004**, *47*, 2743.
194. Diller, D. J.; Hobbs, D. W. *J. Med. Chem.* **2004**, *47*, 6373.
195. Bender, A.; Mussa, H. Y.; Glen, R. C. *J. Biomol. Screen.* **2005**, *10*, 658.
196. Renault, N.; Laurent, X.; Farce, A.; Bakali, J. E.; Mansouri, R.; Gervois, P.; Millet, R.; Desreumaux, P.; Furman, C.; Chavatte, P. *Chem. Biol. Drug Des.* **2013**, *81*, 442.
197. van Hoorn, W. P.; Bell, A. S. *J. Chem. Inf. Model.* **2009**, *49*, 2211.
198. Crisman, T. J.; Bender, A.; Milik, M.; Jenkins, J. L.; Scheiber, J.; Sukuru, S. C. K.; Fejzo, J.; Hommel, U.; Davies, J. W.; Glick, M. *J. Med. Chem.* **2008**, *51*, 2481.
199. Fang, J.; Yang, R.; Gao, L.; Zhou, D.; Yan, S.; Liu, A.-L.; Du, G. *J. Chem. Inf. Model.* **2013**, *53*, 3009.
200. Ekins, S.; Reynolds, R. C.; Kim, H.; Koo, M.-S.; Ekonomidis, M.; Talaue, M.; Paget, S. D.; Woolhiser, L. K.; Lenaerts, A. J.; Bunin, B. A.; Connell, N.; Freundlich, J. S. *Chem. Biol.* **2013**, *20*, 370.
201. Ekins, S.; Freundlich, J. S.; Hobrath, J. V.; White, E. L.; Reynolds, R. C. *Pharm. Res.* **2014**, *31*, 414.
202. Lundholt, B. K.; Scudder, K. M.; Pagliaro, L. *J. Biomol. Screen.* **2003**, *8*, 566.
203. Sittampalam, G. S.; Coussens, N. P.; Nelson, H., Assay Guidance Manual, **2004**, Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences, http://www.ncbi.nlm.nih.gov/books/NBK53196/ (December 2014).
204. Sandlin, R. D.; Fong, K. Y.; Wicht, K. J.; Carrell, H. M.; Egan, T. J.; Wright, D. W. *Int. J. Parasitol.* **2014**, *4*, 316.
205. Klein, K.; Koch, O.; Kriege, N.; Mutzel, P.; Schäfer, T. *Molecular Informatics* **2013**, *32*, 964.
206. Ertl, P.; Jelfs, S.; Muhlbacher, J.; Schuffenhauer, A.; Selzer, P. *J. Med. Chem.* **2006**, *49*, 4568.
207. Egan, T. J.; Mavuso, W. W.; Ncokazi, K. K. *Biochem.* **2001**, *40*, 204.
208. Joshi, M. C.; Wicht, K. J.; Taylor, D.; Hunter, R.; Smith, P. J.; Egan, T. J. *Eur. J. Med. Chem.* **2013**, *69*, 338.
209. Worachartcheewan, A.; Nantasenamat, C.; Isarankura-Na-Ayudhy, C.; Prachayasittikul, V. *Chem. Pap.* **2013**, *67*, 1462.
210. Parenti, M. D.; Pacchioni, S.; Ferrari, A. M.; Rastelli, G. *J. Med. Chem.* **2004**, *47*, 4258.
211. Shah, P.; Siddiqi, M. I. *SAR QSAR Environ. Res.* **2010**, *21*, 527.
212. Raynes, K.; Galatis, D.; Cowman, A. F.; Tilley, L.; Deady, L. W. *J. Med. Chem.* **1995**, *38*, 204.
213. Wu, T.; Nagle, A.; Sakata, T.; Henson, K.; Borboa, R.; Chen, Z.; Kuhen, K.; Plouffe, D.; Winzeler, E.; Adrian, F.; Tuntland, T.; Chang, J.; Simerson, S.; Howard, S.; Ek, J.; Isbell, J.; Deng, X.; Gray, N. S.; Tully, D. C.; Chatterjee, A. K. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6970.
214. Jung, M.; Park, W.-H.; Jung, J.-C.; Lim, E.; Lee, Y.; Oh, S.; Hyung-InMoon *Chem. Biol. Drug Des.* **2009**, *73*, 346.

215. MoSS - Molecular Substructure Miner, v6.10, **2013**, *European Centre for Soft Computing*, Spain
216. Patrick, G. L., An Introduction to Medicinal Chemistry, *Oxford University Press*, United Kingdom, **2013**.
217. Smith, G. F., Designing drugs to avoid toxicity, *Elsevier*, Boston, **2011**, 50.
218. Purohit, V.; Basu, A. K. *Chem. Res. Toxicol.* **2000**, *13*, 673.
219. Padda, R. S.; Wang, C.; Hughes, J. B.; Kutty, R.; Bennett, G. N. *Environ. Toxicol. Chem.* **2003**, *22*, 2293.
220. Cruciani, G.; Carosati, E.; Boeck, B. D.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. *J. Med. Chem.* **2005**, *48*, 6970.
221. MetaSite, v2.0, **2012**, *Molecular Discovery*, Italy, http://www.moldiscovery.com/.
222. VolSurf+, v1.0.7.1, **2012**, *Molecular Discovery*, Italy, http://www.moldiscovery.com/software/vsplus/.
223. Graczyk, P. P.; Khan, A.; Bhatia, G. S.; Palmer, V.; Medland, D.; Numata, H. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 4666.
224. Black J.W.; Durant G.J.; Emmett J.C.; R.C., G. *Nature* **1974**, *248*, 65.
225. Uaucu R.; Karaburun N.G.; I., I. *Farmaco* **2001**, *56*, 285.
226. Mjalli, A.; Sarshar, S., 1,2,4,5-tetra substituted imidazoles as modulators of multi-drug resistance, **1997**, Patent US 5700826, pp 19.
227. Antolini, M.; Bozzoli, A.; Ghiron, C.; Kennedy, G.; Rossi, T.; Ursini, A. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 1023.
228. Khan, M. S.; Siddiqui, S. A.; Siddiqui, M. S. R. A.; Goswami, U.; Srinivasan, K. V.; Khan, M. I. *Chem. Biol. Drug. Des.* **2008**, *72*, 197.
229. Rathore, D.; Jani, D.; Nagarkatti, R., HDP (heme detoxification protein) involved in hemozoin formation in Plasmodium and Theileria as an anti-protozoal target, and high-throughput screening for antimalarial HDP inhibitors, **2007**, Patent US 2007/0148185 Al.
230. Pink, R.; Hudson, A.; Mouries, M. A.; Bendig, M. *Nat. Rev. Drug Discov.* **2005**, *4*, 727.
231. Bréthous, L.; Garcia-Delgado, N.; Schwartz, J.; Bertrand, S.; Bertrand, D.; Reymond, J.-L. J. *J. Med. Chem.* **2012**, *55*, 4605.
232. Mao, W.; Ning, M.; Liu, Z.; Zhu, Q.; Leng, Y.; Zhang, A. *Bioorg. Med. Chem.* **2012**, *20*, 2982.
233. Mallakpour, S.; Zarei, M. *High Perform. Polym.* **2013**, *25*, 245.
234. Il'ina, I. G.; Mikhalev, O. V. *Russ. J. Org. Chem.* **2010**, *47*, 1488.
235. Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61*, 10827.
236. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J. *J. Comb. Chem.* **1999**, *1*, 55.
237. Naredla, R. R.; Klumpp, D. A. *Tetrahedron Lett.* **2012**, *53*, 4779.
238. Wu, X.-F.; Neumann, H.; Beller, M. *Chem. Asian J.* **2010**, *5*, 2168.
239. Fang, W.; Deng, Q.; Xu, M.; Tu, T. *Org. Lett.* **2013**, *15*, 3678.
240. Bjerglund, K.; Lindhardt, A. T.; Skrydstrup, T. *J. Org. Chem.* **2012**, *77*, 3793.
241. Luca, L. D.; Giacomelli, G.; Porcheddu, A. *J. Org. Chem.* **2002**, *67*, 6272.
242. Donaruma, L. G.; Heldt, W. Z. *Org. React.* **1960**, *11*, 1.
243. Beckmann, E. *Berichte der Deutschen Chemischen Gesellschaft* **1886**, *19*, 988.
244. Jursic, B. S.; Zdravkovski, Z. *Synth. Commun.* **1993**, *23*, 2761.
245. Valeur, E.; Bradley, M. *Chem. Soc. Rev.* **2009**, *38*, 606.
246. Pizey, J. S., Synthetic Reagents, *Wiley*, New York, **1974**, Vol. 1, pp. 321–357.
247. Chu, W.; Tu, Z.; McElveen, E.; Xu, J.; Taylor, M.; Luedtke, R. R.; Mach, R. H. *Bioorg. Med. Chem.* **2005**, *13*, 77.
248. Adams, R.; Ulrich, L. H. *J. Am. Chem. Soc.* **1920**, *42*, 599.
249. Knapp, S.; Gibson, F. S., In Organic Syntheses, *Wiley*, New York, **1998**, Collect. Vol. IX, pp. 516–521.
250. Pearson, A. J.; Roush, W. R., Handbook of Reagents for Organic Synthesis: Activating Agents and Protecting Groups, *Wiley*, New York, **1999**, pp. 333-338.
251. Klosa, J. J. *Prakt. Chem. Leipzig.* **1962**, *19*, 45.
252. Clayden, J.; Greeves, N.; Warren, S.; Wothers, P., Organic Chemistry, *Oxford University Press*, New York, **2007**, Vol. 1.
253. Venkataraman, K.; Wagle, D. R. *Tetrahedron Lett.* **1972**, *32*, 3037.
254. Sanchez-Sancho, F.; Mann, E.; Herradon, B. *Synlett* **2000**, *4*, 509.
255. Ragnarsson, U.; Grehn, L. *Acc. Chem. Res.* **1998**, *31*, 494.
256. Topliss, J. G. *J. Med. Chem.* **1972**, *15*, 1006.

257. Shaterian, H. R.; Ranjbar, M. *J. Mol. Liq.* **2011**, *160*, 40.
258. Van den Bossche, H.; Willemsens, G.; Cools, W.; Marichal, P.; Lauwers, W. *Biochem. Soc. Trans.* **1983**, *11*, 665.
259. Azizi, N.; Dado, N.; Amiri, A. K. *Can. J. Chem.* **2012**, *90*, 195.
260. Wan, L.; Wang, X.; Wang, S.; Li, S.; Li, Q.; Tian, R.; Li, M. *J. Phys. Org. Chem.* **2008**, *22*, 331.
261. Khosropour, A. R. *Can. J. Chem.* **2008**, *86*, 264.
262. Bahrami, K.; Nejati, A. *Monatsh. Chem.* **2011**, *142*, 159.
263. Chauveau, E.; Marestin, C.; Schiets, F.; Mercier, R. *Green Chem.* **2010**, *12*, 1018.
264. Chary, M. V.; Keerthysri, N. C.; Vupallapati, S. V. N.; Lingaiah, N.; Kantevari, S. *Cat. Comm.* **2008**, *9*, 2013.
265. Kakuta, H.; Zheng, X.; Oda, H.; Harada, S.; Sugimoto, Y.; Sasaki, K.; Tai, A. *J. Med. Chem.* **2008**, *51*, 2400.
266. Wang, H.; Cao, X.; Xiao, F.; Liu, S.; Deng, G.-J. *Org. Lett.* **2013**, *15*, 4900.
267. Moore, W. J., Physical Chemistry, *Prentice-Hall*, London, **1962**, pp. 140-142.
268. "Azeotropic data for ninary mixtures", in CRC Handbook of Chemistry and Physics, **2005**, CRC Press, Florida, http://www.hbcpnetbase.com (January 2015).
269. MarvinSketch v6.3.0, **2014**, *ChemAxon Ltd.*, http://www.chemaxon.com.
270. Lagnoux, D.; Delort, E.; Douat-Casassus, C.; Esposito, A.; Reymond, J.-L. *Chem. Eur. J.* **2004**, *10*, 1215.
271. Jentoft, N.; Dearborn, D. G. *J. Biol. Chem.* **1979**, *254*, 4359.
272. Lane, C. F. *Synthesis* **1975**, *3*, 135.
273. Engel, T.; Reid, P., Physical Chemistry, *Pearson Education Benjamin Cummings*, San Fransisco, **2006**, Vol. 1.
274. Safari, J.; Khalili, S. D.; Banitaba, S. H. *Syn. Org. Chem.* **2011**, *41*, 2359.
275. Silversmith, E. F., 2,4,5-Triarylimidazoles, *E I du Pont de Nemours & Co.*, **1965**, Patent FR 1395112, 17 pp.
276. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512.
277. Joo, C.; Kanga, S.; Kim, S. M.; Han, H.; Yang, J. W. *Tetrahedron Lett.* **2010**, *51*, 6006.
278. Sklute, G.; Oizerowich, R.; Shulman, H.; Keinan, E. *Chem.-Eur. J.* **2004**, *10*, 2159.
279. Reusch, W., Supplemental NMR Topics, **2013**, International Organization for Chemical Sciences, Michigan, USA, https://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/Spectrpy/nmr/nmr2.htm (March 2015).
280. Agrawal, A.; Tratnyek, P. G. *Environ. Sci. Technol.* **1995**, *30*, 153.
281. Hung, H.-M.; Ling, F. H.; Hoffmann, M. R. *Environ. Sci. Technol.* **2000**, *34*, 1758.
282. Shaban, E.; Wicht, K. J.; Wang, N.; Mei, Z.-W.; Hayashi, I.; Aleem, A. A.; Gokha, E.; Kaiser, M.; Tantawy, I. E.; Sayed, E.; Egan, T. J.; Inokuchi, T. *Heterocycles* **2014**, *89*, 1055.
283. Patrick, G. L., An Introduction to Medicinal Chemistry, *Oxford University Press*, Oxford, **2001**, pp. 276.
284. Katritzky, A. R.; Pozharski, A. F., Handbook of Heterocyclic Chemistry, *Pergamon/Elsevier*, United Kingdom, **2000**, Vol. 1.
285. Egan, T. J. *J. Inorg. Biochem.* **2006**, *100*, 916.
286. Kelly, J. X.; Winter, R.; Riscoe, M.; Peyton, D. H. *J. Inorg. Biochem.* **2001**, *86*, 617.
287. Gaussian Software, v3.0, **2004**, *Gaussian, Inc.*, Wallingford CT, http://www.gaussian.com/.
288. Chou, A. C.; Fitch, C. D. *Biochem. Biophys. Res. Comm.* **1993**, *195*, 422.
289. Materials Studio, v4.0, **2013**, *Accelrys Software Inc.*, San Diego, www.accelrys.com/products/materials-studio/.
290. Molecular Modelling Pro Plus, v7.0.2, **2013**, *ChemSw Inc.*, United States, http://www.norgwyn.com/mmpplus.html.
291. Brook, I. *Clin Infect Dis.* **1989**, *11*, 361.
292. Gluzman, I. Y.; Schlesinger, P. H.; Krogstad, D. J. *Antimicrob. Agents Chemother.* **1987**, *31*, 32.
293. O'Neill, P. M.; Willock, D. J.; Hawley, S. R.; Bray, P. G.; Storr, R. C.; Ward, S. A.; Park, B. K. *J. Med. Chem.* **1997**, *40*, 437.
294. Walden, J. C., Smith, P. J., Egan, T. J., Characterisation of mefloquine accumulation in Plasmodium falciparum, **2003**, Department of Medicine, *University of Cape Town*, Doctor of Philosophy, http://open.uct.ac.za/handle/11427/8601 (February 2015).
295. Yayon, A.; Timberg, R.; Friedman, S.; Ginsburg, H. *J. Protozool.* **1984**, *31*, 367.
296. Chou, A. C.; Chevli, R.; Fitch, C. D. *Biochemistry* **1980**, *19*, 1543.
297. Fitch, C. D., Malaria and the Red Cell 2, *Alan R. Liss*, New York, **1989**, pp. 45–52.

298. Biagini, G. A.; Richier, E.; Bray, P. G.; Calas, M.; Vial, H.; Ward, S. A. *Antimicrob. Agents Chemother.* **2003**, *47*, 2584.
299. Omodeo-Sale, F.; Cortelezzi, L.; Basilico, N.; Casagrande, M.; Sparatore, A.; Taramelli, D. *Antimicrob. Agents Chemother.* **2009**, *53*, 4339.
300. Kuter, D.; Chibale, K.; Egan, T. J. *J. Inorg. Biochem.* **2011**, *105*, 684.
301. Martis, E. A.; Radhakrishnan, R.; Badve, R. R. *J. Appl. Pharm. Sci.* **2011**, *1*, 2.
302. Bender, A. *Springer Sci.* **2011**, *672*, 175.
303. Wicht, K. J.; Combrinck, J. M.; Smith, P. J.; Egan, T. J. *Bioorg. Med. Chem. Lett.* **2015**, *In Press.*
304. Gupta, S. K. *Int. J. App. Basic Med. Res.* **2012**, *2*, 3.
305. Discovery Studio Modeling Environment, v4.0, **2013**, *Accelrys Software Inc.*, San Diego, www.accelrys.com/products/discovery-studio/.
306. Spitzmüller, A.; Mestres, J. *PLoS Comput. Biol.* **2013**, *9*, 1.
307. Nzila, A. *J. Antimicrob. Chemother.* **2006**, *57*, 1043.
308. Vhquez-Vivar, J.; Augusto, O. *J. Biol. Chem.* **1992**, *267*, 6848.
309. Dahl, E. L.; Shock, J. L.; Shenai, B. R.; Gut, J.; DeRisi, J. L.; Rosenthal, P. J. *Antimicrob. Agents Chemother.* **2006**, *50*, 3124.
310. Baggish, A. L.; Hill, D. R. *Antimicrob. Agents Chemother.* **2002**, *46*, 1163.
311. Alumasa, J. N.; Gorka, A. P.; Casabianca, L. B.; Comstock, E.; de Dios, A. C.; Roepe, P. D. *J. Inorg. Biochem.* **2011**, *105*, 467.
312. World Health Organisation, Q&A on artemisinin resistance, **2015**, World Health Organisation, Switzerland, http://www.who.int/malaria/media/artemisinin_resistance_qa/en/ (March 2015).
313. Microsoft Office Professional Plus, **2010**, *Microsoft Corporation*, United States, http://products.office.com/en-us/excel.
314. GraphPad Prism, v5.0, **2007**, San Diego California USA, www.graphpad.com.
315. SoftMax Pro, **2013**, *Molecular Devices*, USA, http://www.moleculardevices.com/systems/microplate-readers/softmax-pro-data-acquisition-and-analysis-software.
316. Egan, T. J.; Mavuso, W. W.; Ross, D. C.; Marquesc, H. M. *J. Inorg. Biochem.* **1997**, *68*, 137.
317. Bemis, G. W.; Murcko, M. A. *J. Med. Chem.* **1996**, *39*, 2887.
318. Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. *Nucleic Acids Res.* **2006**, *34*, 901.
319. Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. *Nucleic Acids Res.* **2008**, *36*, 901.
320. Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. *J. Chem. Inf. Model.* **2012**, *52*, 1757.
321. Raj, R.; Mehra, V.; Gut, J.; Rosenthal, P. J.; Wicht, K. J.; Egan, T. J.; Hopper, M.; Wrischnik, L. A.; Land, K. M.; Kumar, V. *Euro. J. Med. Chem.* **2014**, *84*, 425.
322. Ahn, K. H.; Chung, Y.; Jun, M.; Kim, D. J. *J. Org. Chem.* **2009**, *74*, 4849.
323. George, S.; Shanmugapandiyan, P. *Int. J. Pharm. Pharm. Sci.* **2011**, *4*, 104.
324. Miller, T. M.; Neenan, T. X. *Chem. Mater.* **1990**, *2*, 346.
325. Odago, M. O.; Hoffman, A. E.; Carpenter, R. L.; Tse, D. C. T.; Sun, S.-S.; Lees, A. J. *Inorg. Chim. Acta.* **2011**, *374*, 558.
326. Qin, Z.; Jennings, M. C.; Puddephatt, R. J. *Inorg. Chem.* **2003**, *42*, 1956.
327. Lukin, O.; Kubota, T.; Okamoto, Y.; Schelhase, F.; Yoneva, A.; Müller, W. M.; Müller, U.; Vögtle, F. *Angew. Chem. Int. Ed.* **2003**, *42*, 4542.
328. Horino, H.; Sakaba, H.; Arai, M. *Synthesis* **1989**, *9*, 715.
329. Palmans, A. R. A.; Vekemans, J. A. J. M.; Fischer, H.; Hikmet, R. A.; Meijer, E. W. *Chem. Eur. J.* **1997**, *3*, 300.
330. Puratchikody, A.; Gopalakrishnan, S.; Nallu, M. *Ind. J. Pharm. Sci.* **2005**, *67*, 725.