Anti-malarial Agents Targeting Apical Membrane Antigen 1

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

by

Cael Debono
B. Pharm. Sci. (Hons) (Monash University)

Monash Institute of Pharmaceutical Sciences
Faculty of Pharmacy and Pharmaceutical Sciences
Medicinal Chemistry Theme
381 Royal Parade, Parkville, Victoria, 3052, Australia
Monash University
June 2018

Declaration

To the best of the authors knowledge and belief, this thesis contains no material which has been accepted for the award of any other degree or diploma in any university or other institution, and contains no material previously published or written, except where due reference is made. This work disclosed is based on my own work except where due reference is made in the text of the thesis.

Cael Debono

June 2018

Contents

Acknowledgements

I originally began my scientific career completing a bachelor degree in science at the University of Melbourne. It wasn't until my second year at university that I finally committed to studying chemistry, in particular medicinal chemistry, since we had to major in one branch of science on offer for the course. It was at this point, my passion for chemistry had developed into more than a fleeting interest. After transferring to Monash University to complete my honours, I realized I enjoyed research and working in the lab so the decision to complete my doctorate was an easy one.

The completion of a PhD is not a solo effort and many people have helped me along the way. First and foremost the three individuals that deserve the utmost acknowledgement for their boundless support are my three supervisors Professor Peter Scammells, Professor Ray Norton and Dr. Shane Devine, thanks for all your assistance, your patience and most importantly your abilities and interest within the project which has led to our work being published which was a personal goal I wanted to achieve through my research. I would also like to extend my gratitude to Drs. Damien Drew, Taryn Guinan, Chris MacRaild and Indu Chandrashekaran. Thank you for being involved in the project and with all your help with the biology and testing associated with the project. To Dr. Jason Dang, thanks for all your help with chemical analysis, for your help with NMR, HPLC, HRMS and LCMS. Your understanding of these analytical instruments was very helpful for myself and I am sure for many of the researchers at Monash. An extended praise goes to the members of lab 123, lab 108 and lab 111, in particular Leigh Ford, Bankala Krishnarjuna and Mansura Akter, you guys were always available to help me out with any of my queries. Also I would like to extend my gratitude to all the fellow students I have shared my PhD experience with Matthew Wakeham, Cassandra Yong, Anthony La, Jeffrey Seow. Thanks for all the great memories, drunken social functions and the friendship support. The accomplishment of a PhD is not solely dependent on the support of your work colleagues. The most integral support came from my parents Mary Rose and Anthony and siblings Aaron and Ashton. Thanks so much for caring; being supportive and understanding of stressful times I endured. Also for your financial support to make a goal of mine a reality.

Abbreviations

AMA1	Apical membrane antigen 1
°C	Degrees Celsius
^{13}C-NMR	Carbon-13 NMR
^{1}H-NMR	Proton NMR
Boc	*tert*-Butyloxycarbonyl
Calcd.	Calculated
Compd	Compound
CPMG	Carr-Purcell-Meiboom-Gill
D	Doublet
DCM	Dichloromethane
DIPEA	*N, N*-Diisopropylethylamine (Hünig's base)
DMSO	Dimethylsulfoxide
EtOAc	Ethyl acetate
EDC	*N*-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide
FBDD	Fragment-Based Drug Discovery
equiv	Equivalent
g	Grams
IC_{50}	concentration of compound required to produce 50% inhibition
h	hours
H-bond	Hydrogen bond
Hz	Hertz
HSQC	Heteronuclear single quantum coherence
LCMS	Liquid chromatography mass spectrometry
J	J-coupling

LID	Ligand-induced dissociation
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
μL	micromolar
mL	millimolar
M	Molar
MJ	Moving Junction
m	Multiplet
m/z	Mass to charge ration
min	Minutes
MW	Molecular weight
NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide
NMR	Nuclear magnetic resonance
Pf	*Plasmodium falciparum*
q	Quartet
RON	Rhoptry neck protein
rt	Room temperature
SAR	Structure activity relationship
STD	Saturation transfer difference
T	triplet
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
tlc	Thin layer chromatography
TOF	Time-of-flight

Publications/Presentations

Work conducted during this PhD has resulted in a number of publications and presentations to date and these are listed below:

Papers: (The complete paper may be found either throughout thesis or in Appendix A):

1) Krishnarjuna, B.; Lim, S. S.; Devine, S. M.; **Debono, C. O**.; Lam, R.; Chandrashekaran, I. R.; Jaipuria, G.; Yagi, H.; Atreya, H. S.; Scanlon, M. J.; MacRaild, C. A.; Scammells, P. J.; Norton, R. S. Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials. *Journal of Molecular Recognition* **2016**, *29*, 281–291.
2) Devine, S. M.; Mulcair, M. D.; **Debono, C. O**.; Leung, E. W.; Nissink, J. W. M.; Lim, S. S.; Chandrashekaran, I. R.; Vazirani, M.; Mohanty, B.; Simpson, J. S.; Baell, J. B.; Scammells, P. J.; Norton, R. S.; Scanlon, M. J. Promiscuous 2-aminothiazoles (PrATs): a frequent hitting scaffold. *Journal of Medicinal Chemistry* **2015**, *58*, 1205-1214.
3) Ge, X.; MacRaild, C. A.; Devine, S. M.; **Debono, C. O**.; Wang, G.; Scammells, P. J.; Scanlon, M. J.; Anders, R. F.; Foley, M.; Norton, R. S. Ligand-Induced Conformational Change of *Plasmodium falciparum* AMA1 detected Using ^{19}F NMR. *Journal of Medicinal Chemistry* **2014**, *57*, 6419-6427.
4) Lim, S. S.; Yang, W.; Krishnarjuna, B.; Kannan Sivaraman, K.; Chandrashekaran, I. R.; Kass, I.; MacRaild, C. A.; Devine, S. M.; **Debono, C. O**.; Anders, R. F.; Scanlon, M. J.; Scammells, P. J.; Norton, R. S.; McGowan, S. Structure and Dynamics of Apical Membrane Antigen 1 form *Plasmodium falciparum* FVO. *Biochemistry* **2014**, *53*, 7310-7320.
5) Devine, S. M; Lim, S. S.; Chandrashekaran, I. R.; MacRaild, C. A.; Drew, D. R.; **Debono, C. O**.; Lam, R.; Anders, R. F.; Beeson, J. G.; Scanlon, M. J.; Scammells, P. J.; Norton, R. S. A critical evaluation of pyrrolo[2,3-d]pyrimidine-4-amines as *Plasmodium falciparum* apical membrane antigen 1 (AMA1) inhibitors. *MedChemComm* **2014**, *5*, 1500-1506.
6) Lim, S. S.; **Debono, C. O**.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening. *Australian Journal of Chemistry* **2013**, *66*, 1530–1536.

Oral presentation:

Fragment-based Drug Design Down Under at Monash University (Parkville, July 2016). **Debono, C. O**.; Lim, S. S.; MacRaild, C. A.; Chandrashekaran, I. R.; Devine, S. M.; Krishnarjuna, B.; Anders, R. F.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Anti-malarial Agents Targeting Apical Membrane Antigen 1.

Poster presentation:

EFMC-ISMC XXIV EFMC International Symposium on Medicinal Chemistry (Manchester, August-September 2016). **Debono, C. O**.; Lim, S. S.; MacRaild, C. A.; Chandrashekaran, I. R.; Devine, S. M.; Krishnarjuna, B.; Anders, R. F.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Anti-malarial Agents Targeting Apical Membrane Antigen 1.

Abstract

AMA1 is a type-1 integral membrane protein essential for invasion of host cells by zoites of all apicomplexan parasites, in particular *Plasmodium* merozoites and sporozoites. This protein is considered an essential component of the formation of the moving junction (MJ) that allows for the invasion of human erythrocytes and hepatocytes during malarial infection. AMA1 is parasite-derived and suggested to facilitate a conserved invasion mechanism across the diverse host cells targeted by apicomplexan parasites. Thus, AMA1 presents itself as a novel and promising antimalarial target. AMA1 has been studied extensively as a potential component of a vaccine against malaria in addition to being investigated as a viable therapeutic target by small molecule inhibitors.

Fragment-based NMR screening conducted at MIPS had identified fragments that bound AMA1 in competition with the peptide R1, a high-affinity peptide for 3D7 *Pf*AMA1. R1 is known to bind the entire length of the highly conserved hydrophobic cleft, a crucial binding region for the formation of the MJ. This provided a platform for an FBDD investigation of select fragment scaffolds in search of higher affinity strain-transcending binders of AMA1. The work discussed involved the design, synthesis and screening of multiple series of fragment analogues, derived from two separate scaffolds, all targeting the hydrophobic cleft of AMA1. This work focused on the investigation of the 2-aminothiazole and 2-substituted aryl benzimidazole scaffolds, opting to optimize each core in search of SAR using both NMR and SPR techniques to validate analogue binding. Furthermore, the evaluation of SAR for each series was conducted. However, a lack of reliable data resulting from increasingly poor analogue physicochemical properties suggested more accurate affinity measurements were required to determine whether the ongoing development of fragment hits would lead to suitable compounds.

Towards the identification of SAR, a flow-cytometry-based growth assay for inhibitor screening was investigated to screen the 2-substituted aryl benzimidazole scaffold. Investigation of this benzimidazole scaffold was able to identify numerous benzimidazole analogues with nanomolar potency against transgenic W2Mef *P. falciparum* parasites. This activity did not correlate distinctively with AMA1 inhibition, suggesting another inhibitory mechanism was plausible.

An alternative strategy for fragment development was the employment of reactive functional groups to produce selective covalent probes of AMA1 in search of structural information. The synthesis of covalently modifying molecules allowed for the selective targeting of AMA1 binding

regions previously identified by NMR to permit analogue binding. A range of covalent probes was designed based upon selected 2-substituted aryl benzimidazole pharmacophores in search of higher affinity and selectivity. Probes were screened using MALDI-TOF MS leading to the development of several covalent adducts demonstrating a 1:1 binding interaction, which is the first reported use of covalently modifying molecules against AMA1.

Collectively, our ability to identify binders of the highly conserved hydrophobic cleft via NMR together with the ability to selectively target particular regions of AMA1, provide valuable insights that should contribute to the design of higher affinity strain-transcending agents targeting *P. falciparum* AMA1.

Paper Declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis. This thesis includes 1 original paper published in peer-reviewed journals. The core theme of the thesis is the development and screening of small molecule inhibitors targeted against AMA1, an essential component involved in the invasion of host cells during malarial infection. The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the Doctor of Philosophy, majoring in Medicinal Chemistry under the supervision of Professor Peter Scammells, Professor Raymond Norton and Doctor Shane Devine.

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: [redacted] **Date:** 13/03/2018

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor signature: [redacted] **Date:** 13/03/2018

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research. In the case of chapter 2 my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
2	*Promiscuous 2-aminothiazoles (PrATs): A Frequent Hitting Scaffold*	*Accepted*	*5%. SPR testing and input into manuscript*	1) Shane M. Devine, *input into manuscript 30%* *2) Martin J. Scanlon, corresponding author 20%* *3) Mark Mulcair, SPR testing 10%* *2) J. Willem N. Nissink, input into manuscript 10%* *9) Jonathan B. Baell, input into manuscript 10%* *3) Elenor W. W. Leung, NMR testing 5%* *4) San Sui Lim, NMR testing 5%* *5) Indu R. Chandrashekaran, input into manuscript* *6) Mansha Vazirani, NMR testing* *7) Biswaranjan Mohanty, NMR testing 5%* *8) Jamie S. Simpson, supervised testing* *10) Peter J. Scammells, supervised testing and input into manuscript* *11) Raymond S. Norton, supervised testing and input into manuscript*	*No* *No* *No* *No* *No* *No* *Yes* *No* *Yes* *No* *No* *No* *No*

Chapter 1 - Introduction

1.1 Malaria

Malaria is an infectious disease caused by parasitic protozoans (unicellular eukaryotic organisms) belonging to the genus *Plasmodium*.[1] Ronald Ross in 1896 was the first to demonstrate that transmission of malaria in humans was facilitated by a mosquito vector.[2] Transmission occurs via the proboscis of an infected *Anopheles* mosquito.[3] Five *Plasmodium* species are known to cause malaria in humans, with *Plasmodium falciparum* (*Pf*) being responsible for most cases of mortality and morbidity amongst humans.[4] The *Plasmodium* species, *P. vivax* also causes malaria across parts of the world although its percentage contribution towards malaria-related illness is not yet fully understood.[5] Currently, malaria is regarded as one of the world's most prevalent diseases, resulting in serious global health problems and imposing substantial hardship on the world's most vulnerable societies. *P. falciparum* has the potential to affect half of the world's human population through the transmission of malaria. The World Health Organization (WHO) world malaria report states the global tally of malaria in 2016 was 216 million cases leading to 445,000 deaths with a high proportion of mortalities being attributed to young children and pregnant women.[1] In 2016, malaria financing totaled US$ 2.7 billion highlighting the annual financial burden attributed to the treatment and control of the global malaria threat.[1] The emergence of resistant strains to first-line chemotherapies has only worsened the global impact of malaria on governments and countries affected, warranting an urgent response from the global scientific community to develop new drugs with novel mechanisms of action that could minimize the emergence and impact of resistant strains.[6,7]

1.2 *Plasmodium* Life Cycle

The *Plasmodium* parasite belongs to the phylum Apicomplexa, which is characterized by the presence of an apical prominence that drives parasitic invasion of host cells.[8] *Plasmodium* species have a complex life cycle, moving between female *Anopheles* mosquitoes (vector) and vertebrate (host), and requiring the formation of unique invasive forms for the invasion of particular host cell types (Figure 1.1).[9] Once invasive sporozoites enter the host, they travel to the liver and infect liver hepatocytes. The sporozoites mature in the hepatocytes and, once matured, become merozoites, which are then released from the hepatocytes to begin the asexual cycle within the blood. These merozoites invade the host erythrocytes within the blood. The blood stage begins once the release of mature merozoites from hepatocytes into the blood stream is initiated, thus leading to the asexual

reproduction of merozoites upon erythrocyte invasion. The blood stage cycle is responsible for the clinical symptoms associated with malarial infection such as fever, fatigue and eventually death.

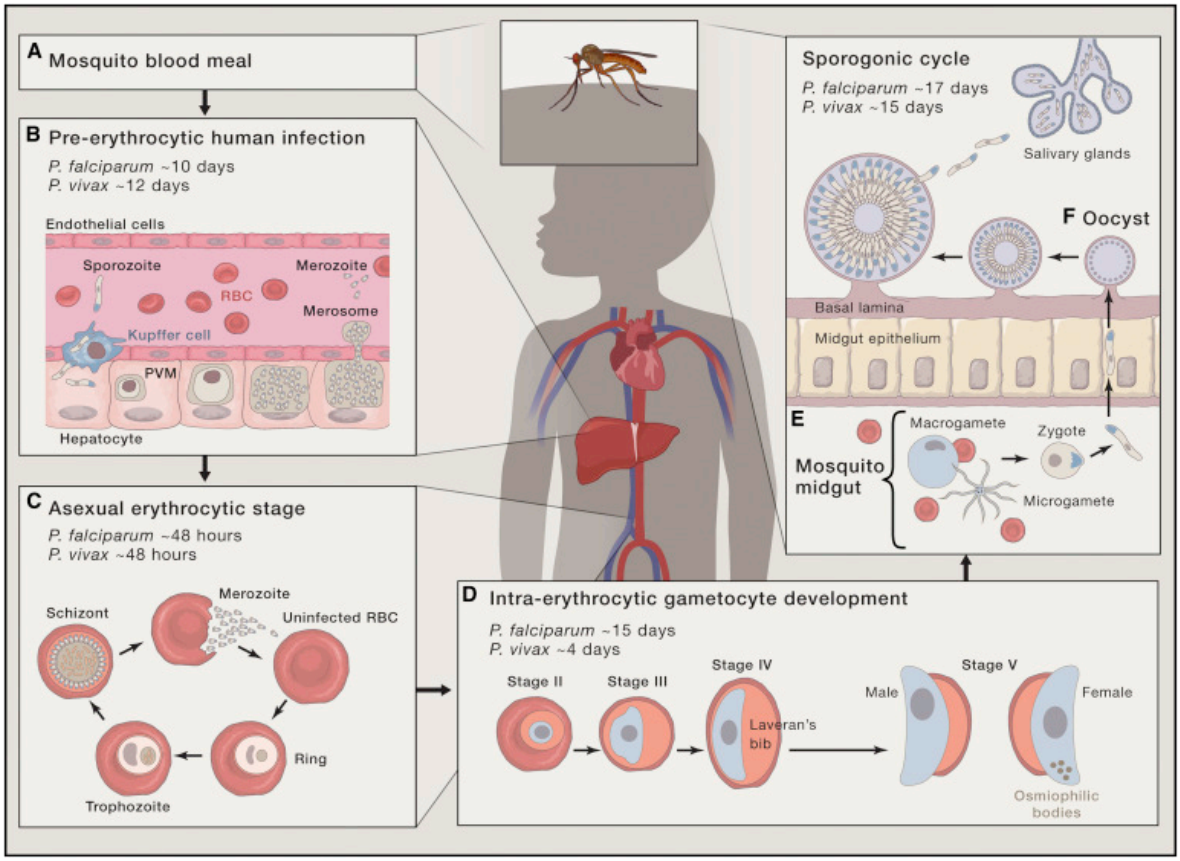


Figure 1.1 Life cycle of *P. falciparum*.[9] (A) Malaria infection is initiated with the injection of sporozoites. (B) The liver stage begins once the sporozoites are transported to the liver, entering hepatocytes and subsequently undergoing schizogony until tens of thousands of daughter merozoites are released in merosomes into the vasculature. (C) When released, erythrocytes are invaded beginning the chronic cycle of asexual schizogony in the bloodstream. (D) Some asexually producing merozoites are reprogrammed to undergo gametocytogenesis (sexual cycle). (E) Within an 11-15 day period, gametocytes develop within the bone marrow and, once mature, enter the peripheral circulation to be taken up by a mosquito during a blood meal. (F) The fusion of micro- and macrogamets forms a zygote that transforms over 24 h into an ookinete, which moves through the mosquito midgut epithelium and encysts to become an oocyst where asexual sporogenic replication occurs. Motile sporozoites are released upon oocyst rupture and pass into salivary glands ready to be injected.

1.2.1 Liver Stage

Plasmodium sporozoites are injected into the host dermis via the mosquito proboscis during a blood feed. The sporozoites rely on gliding motility, which is a random process to move a large proportion of the sporozoites towards the blood vessel and inadvertently enter the blood stream (**Figure 1.1**).[9] At this point the sporozoites are deemed to be in a migratory mode as they make their way towards the liver.[10] Those sporozoites that do not successfully arrive in the bloodstream are destroyed and drained by the host lymphatic system, resulting in a host immune response.[9] Once in the bloodstream, sporozoites quickly make their way towards the liver via a process known as traversal, which involves crossing the sinusoidal barrier made up of fenestrated endothelial cells and macrophage-like Kupffer cells.[11] A number of proteins facilitate this process, such as SPECT (sporozoite microneme protein) (also known as perforin-like protein 1, cell traversal protein, phospholipase and gamete egress and sporozoite traversal protein (GEST)).[12-14] Although the exact role of each of these proteins is still to be confirmed, SPECT2 is understood to contain a membrane attack complex/perforin-like domain, suggesting that it plays a critical role in punching holes within membranes to allow the parasite to enter.[13]

Once the sporozoites have interacted with hepatocytes in the liver, they turn from migratory mode to invasive mode, which is a direct result of the sporozoite binding to a range of proteins found on the hepatocyte surface, such as tetraspanin CD81 and scavenger receptor B1 (SC-B1).[10] The actual invasion of the hepatocyte requires the binding of circumsporozoite protein to highly sulfated proteoglycans found on the surface.[15] The subsequent steps involve a variety of proteins, including apical membrane antigen 1 (AMA1), which is released from the apical organelles of the sporozoite.[16] Over the next 2-10 days following sporozoite invasion, the sporozoite transforms into its exo-erythrocyte form, eventually culminating in the development of an invasive merozoites that are released into the blood stream to begin the asexual blood stage cycle of malaria.[17] For each hepatocyte infected, approximately 40,000 merozoites are released via a parasite-filled vesicle known as a merosome.[17]

1.2.2 Erythrocyte Invasion

The invasion of erythrocytes in the bloodstream is a fast, dynamic and multi-step process that is split up into three main stages known as pre-invasion, active invasion and echinocytosis, which is complete within 2 min (**Figure 1.2**).[18] The pre-invasion process involves the initial interaction between the merozoite and erythrocyte, notably through merozoite surface protein 1 (MSP1), which acts as a platform for at least 3 large complexes with different extrinsic proteins that bind

erythrocytes.[19,20] Various studies working with merozoites lacking MSP1 have demonstrated the ability to invade erythrocytes, suggesting that the protein might not be necessary for successful invasion.[21] The interaction of the merozoite and erythrocyte results in parasite actomyosin motor-driven deformation of the host cell brought about by two ligand families of type 1 membrane proteins in *Pf*.[22] The erythrocyte binding-like proteins (EBLs) such as EBA-175 which trigger the release of proteins from the rhoptries, and *Pf* reticulocyte-binding protein homologues (*Pf*Rhs) play an important role in signaling the activation of subsequent steps following pre-invasion.[23] Another key protein involved in the attachment of the merozoite is calcineurin, which helps to stabilize the dimerization of EBL and *Pf*Rh proteins, a critical process for host-receptor ligation and signal transduction for subsequent events during the invasion process.[24]

After initial contact and recognition of the host cell surface receptors, the merozoite is re-oriented such that the apical end is in direct contact with the erythrocyte membrane.[25] The merozoite moves into the erythrocyte via the formation of a tight junction, known as the moving junction (MJ), formed between primarily AMA1 and the rhoptry neck protein (RON) complex.[26,27] The RON complex is released from the merozoite and deposited on the erythrocyte surface to bind AMA1, lipid-rich rhoptry contents are also released to help form the parasitophorous vacuole membrane, which forms a protective seal around the merozoite as the parasite moves inside the cell via an actomyosin motor.[28] The vacuole is sealed at the posterior end of the merozoite once completely inside, protecting the merozoite from the cytoplasm of the erythrocyte.[9] The process of echinocytosis follows once the parasite is inside the erythrocyte, resulting in the host cell shrinking and forming spiky protrusions to withstand harsher conditions caused by the infection.[18] The merozoite encased inside the vacuole undergoes cell division (schizogony) over the next 48 h eventually leading to the destruction of the host cell and resulting in the release of between 16 and 32 merozoites into the bloodstream.[29] The release of merozoites is tightly regulated and involves a number of protein kinases such as cyclic guanosine monophosphate-dependent protein kinase.[30]

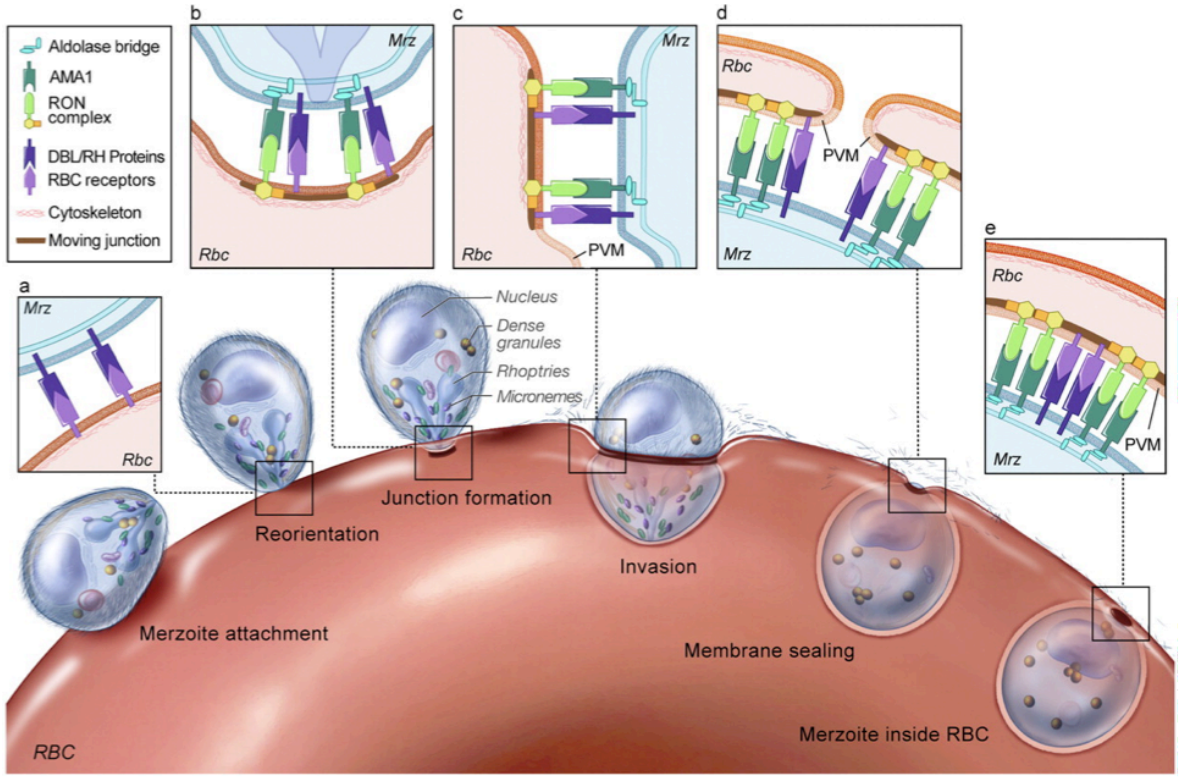


Figure 1.2 Schematic model of the steps involved in *Pf* merozoite invasion.[31] This figure outlines the 5-step process entitled a) to e) of the invasive mechanism, beginning with a) the reorientation of the merozoite on the erythrocyte surface, b) MJ formation, c) the bulk of the invasion mechanism, d) sealing stage of invasive parasite via formation of parasitophorous vacuole and e) merozoite inside erythrocyte.

1.2.3 Sexual Stage

As merozoites divide in the blood stream, a developmental switch occurs for some parasites, initiating commitment to sexual development to form male and female gametocytes. The transmission of malaria from the human host to the mosquito vector is dependent on development of the sexual stages, although molecular events surrounding this developmental switch are not fully characterized yet. This process is understood to occur at some point in the previous schizogony cycle, where daughter merozoites from a single schizont-infected cell are committed to developing into either gametocytes or asexual schizonts.[9] In particular, *Pf* gametocyte maturation is an extended process compared to other species. In total, it takes around 11 days once the commitment to this process has been initiated for mature gametocytes that are infectious to the vector to develop, during which period these gametocytes are located in the human bone marrow to avoid splenic clearance.[32]

1.3 Current front-line antimalarial therapies

The fourth decade of the 17th century marked the first specific remedy for periodic fevers later recognized to be the first antimalarial remedy when Alphonse Laveran identified the intra-erythrocytic protozoan parasite that caused malaria in 1880.[33] The Cinchona alkaloids (quinine, quinidine and cinchonine) were considered the earliest frontline therapies arresting the intra-erythrocytic life cycle of malaria.[34] However the products of a ubiquitous shrub known as *Artemesia annua* (sweet wormwood), most notably artemisinin (qinghaosu in Chinese) slowly took charge eventually becoming apart of what is considered the current frontline antimalarial therapy (**Figure 1.3**).[35] Before the discovery of artemisinin, the efficacy of previous front-line therapies such as chloroquine and sulfadoxine–pyrimethamine, both of differing pharmacological profiles, had diminished owing to the emergence of resistant strains. Chinese scientists first announced Artemisinin to the rest of the world in 1979, eight years after its initial discovery in 1971.[36] Artemisinin is understood to attribute its antimalarial activity to its peroxide bridge however the exact mode of action remains uncertain. Owing to the development of a full chemical synthesis of artemisinin four years later, the synthesis of dihydroartemisinin (DHA) and its derivatives artemether and artesunate eventually replaced the parent drug, artemisinin, owing to their greater antimalarial activity. Although artemisinin derivatives had become the leading antimalarial therapies on the market, when given alone, adherence to a 7-day regiment was required to maximize cure rates however artemisinin regiments were poorly adhered too. The failure to adhere to required treatment programs eventually led to the emergence of resistance due to prolonged exposure of parasites to insufficient dosing. In order to improve the efficacy and treatment time using artemisinin, WHO recommended that artemisinins be used as a combination therapy with various slowly eliminating antimalarial drugs such mefloquine, sulfadoxine-pyrimethamine or aodiaquine as the slower acting partner drug. This therapy was subsequently known as artemisinin-based combination therapies (ACTs).[37] Through using ACTs, WHO recommended aiming for a cure rate of 95%, assessed at 28 days following initial drug treatment to determine whether initial aim had been achieved. If a 95% cure rate was not achieved, WHO also recommended changing antimalarial treatment policy. Considering numerous strains are believed to cause malaria such as *P. vivax, P. malariae* and *P. ovale*, the advantages of using an ACT is that the ACT is highly effective against all infective strains of malaria.

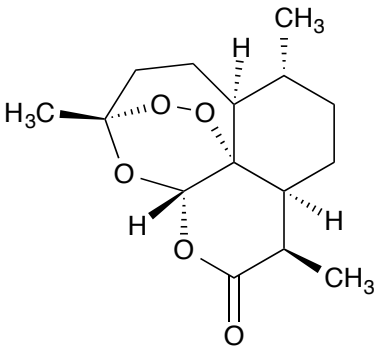


Figure 1.3 The structure of artemisinin (Qinghaosu in Chinese).[35]

1.3.1 Growth of resistance

Because the current antimalarial landscape is dominated by the use of artemisinin-derived therapies, the continual use of ACTs has led to considerable concern that resistance would emerge to this class as it has for previous frontline treatments. Observational studies near the Thai–Cambodian border in the early 2000s have not eased concerns, alluding to an alarming increase in ACT treatment failure due to delayed parasite clearance time.[38] This region is notorious considering its recognized as the epicenter from which chloroquine-resistant and later multidrug-resistant *P. falciparum* spread 50 years ago.[38] Treatment failure rates after artesunate-mefloquine and artemether-lumefantrine in western Cambodia often exceed 10%, consistently higher than anywhere else in the world.[38] A report from western Cambodia suggested the use of sub-standard and/or sub-optimal drugs to be one of the factors contributing to the increasing incidence of resistance, eventually resulting in artemisinin monotherapy cessation after 30 years of clinical use in the region.[39,40] Used as a monotherapy, artemisinin has a very short half-life in plasma, typically in the order of 1-3 h and usually takes 7 days or more of repeated administration to treat malarial infection.[41] The short half-life of a monotherapy treatment was another factor which stressed the need for ACTs in order to ensure parasites were repeatedly exposed to high drug levels in the plasma. When used as a combination therapeutic approach, the treatment duration could be slashed by half and provide a much more effective cure as opposed to a monotherapy treatment program.[42] Another factor that has contributed to the development of malarial resistance is the very low rate of transmission in some affected areas causing insufficient immunity towards parasites that may have survived drug treatment in the area.[43] With limited alternative chemotherapeutics available and the increasing prevalence of ACT resistance, it is apparent that the fight against malaria requires the development of novel therapeutics.

1.4 Apical Membrane Antigen 1 (AMA1)

1.4.1 AMA1 as an essential component of the invasion process

The molecular events that occur upon the merozoites invading host erythrocytes have been extensively studied. Parasite molecules contained within the apical complex of merozoites are considered to play a critical role in the invasion process. The major organelles that make up the apical complex of the apicomplexa phylum are known as the rhoptries, which are membrane-bound electron-dense structures that appear to discharge contents during invasion of erythrocytes. In 1989, a study by Peterson *et al.*,[44] investigated the novel cDNA clone Ag352 structure and its gene; this study led to the identification of a polypeptide termed AMA1. Ag352 was identified by screening a library of cDNA prepared from *P. falciparum* isolate NF7 with human antimalarial antibodies.[45] Indirect Immunofluorescence microscopy with affinity-purified human antibodies was used to identify the location and movement of antigens located in the rhoptries.[46] This study was able to identify AMA1 as the first antigen to be located at the apical complex and subsequently exported to the merozoite surface upon schizont rapture. AMA1 was also revealed to be lost during invasion due to the inability to detect immunofluorescence; this result suggested AMA1 might be involved in the invasion process.

Another study in 2005 centered around *Toxoplasma gondii* identified a complex of RON proteins that are released from rhoptries localize to the MJ during invasion of tachyzoites into mammalian host cells.[47,48] This was relevant because the *T. gondii* orthologue of AMA1 was able to co-localize and co-purify with the RON complex similarly to previously identified *P. falciparum* homologues of the RON proteins associating with AMA1.[29,49,50] In addition, studies of *T. gondii* and *P. knowlesi* have also shown rhoptries to contain both RON proteins[51] and lipids[52] which upon exposure to the host cell cytoplasm have been linked to forming part of the membrane of the parasitophorous vacuole. Owing to orthologous interactions known to be present in other *Plasmodium* species, it is plausible that AMA1 is able to recognize a parasite-derived receptor for host-cell invasion that would be released as part of the rhoptry secretion. These orthologous interactions that are known to be present across the apicomplexan genera may also suggest a conserved invasion mechanism involving AMA1 and RON complex proteins.[53] The highly conserved nature of the AMA1 sequence across the apicomplexan species is another feature that suggests AMA1 possesses an essential function within the invasion process. To this point, AMA1 has been targeted using anti-bodies against AMA1 which have shown successful growth inhibition of *P. falciparum* cultures[54] as well as the inhibition of erythrocyte[55] and hepatocyte invasion.[47] The

inhibition studies of AMA1 and the RON complex will be discussed in more detail throughout this chapter.

1.4.2 Structure

Mature AMA1 is a 66-kDa type 1 integral membrane protein that is made up of three domains (I, II, III), containing eight intramolecular disulfide bridges (**Figure 1.4 A**).[56,57] AMA1 was first identified in *P. knowlesi*;[58] a comparison of the AMA1 sequence with GenBank and National Biochemical Research Foundation databases failed to identify any similar proteins. When compared with other blood stage antigens, AMA1 is considered a unique protein in that AMA1 lacks repetitive sequences.[44] Initial crystal structures of the AMA1 ectodomain were of a construct of *P. vivax* AMA1 comprising all three domains (**Figure 1.4 B**),[59] and of a *Pf*AMA1 construct lacking domain III.[60] Domains I and II are strongly conserved across the apicomplexan phylum, suggesting a combined similar role for these domains, whereas domain III is divergent and thought to be less involved in parasitic invasion.[20] Domains I and II share a common core topology, folding together to form a so-called PAN motif or apple fold, this fold consists of a central five-strand β-sheet flanked by a single α-helix on one face and a smaller three-stranded sheet on the other face.[58,61] This particular fold for domains I and II forms a large hydrophobic cleft (**Figure 1.4 C**) that is central in the formation of the AMA1/RON2 invasion complex.[62,63] The base of this cleft is formed by 12 conserved and structurally well-defined residues largely contributed by several loops that make up part of domains I and II.[64] Although several sections of AMA1, notably the loops which make up the sides of the cleft, cannot be resolved in the electron density maps for some AMA1 crystal structures, it is clear that the hydrophobic cleft is made up of both conserved and polymorphic regions.[64]

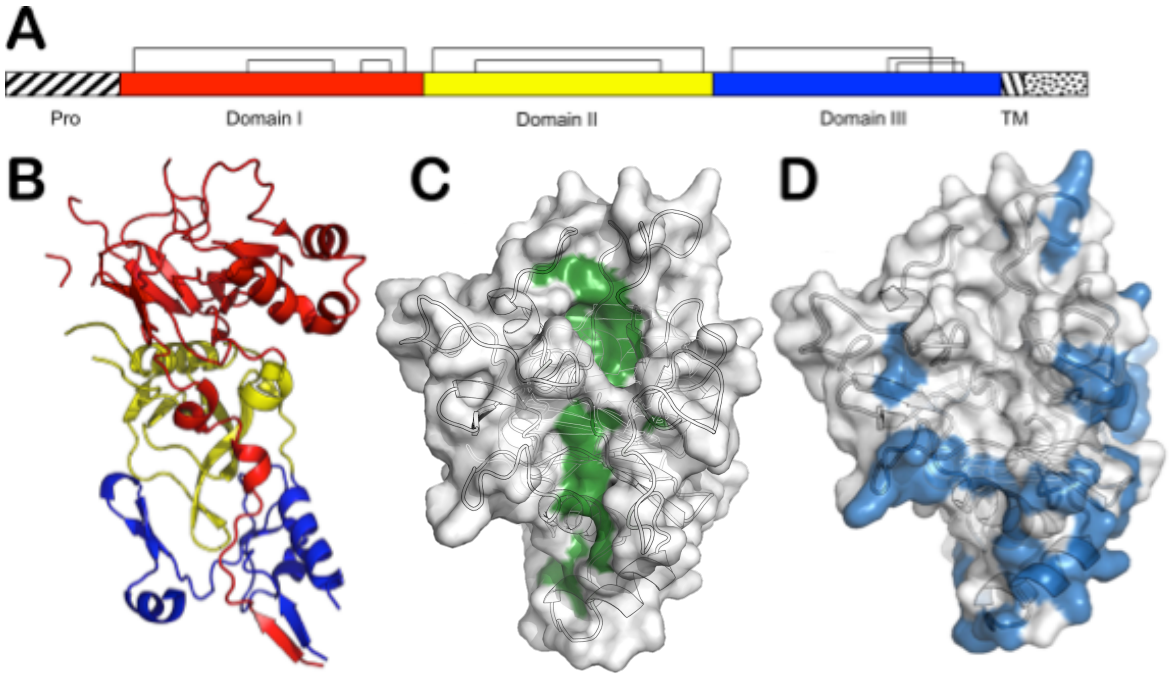


Figure 1.4 Structure of AMA1.[64] (**A**) The primary structure of AMA1, depicting disulphide bridge positioning. (**B**) Structure of the *P. vivax* AMA1 ectodomain (1W8K), showing the organization of domains I, II and III as indicated by A.[65] (**C**) Top view of surface of *Pf*AMA1 (IZ40) showing the conserved hydrophobic cleft (green).[59] (**D**) Top view of surface of AMA1 with the polymorphic residues in *P. falciparum* shaded (blue).

Within *Plasmodium* species, polymorphisms are concentrated in domain I, particularly loops Id and Ie, which surround one end of the cleft (**Figure 1.4 D**). This observed structural feature of the conserved hydrophobic cleft surrounded by variable loops is strongly suggestive of a functionally important site under stringent immune surveillance.[60] The apparent spatial positioning of each loop surrounding the cleft is another striking feature that may allow for the conformational masking of both ends of the cleft owing to the combination of flexible loops at one end of the cleft and the more polymorphic but less flexible loops at the other end of the cleft.[64] Precedence does exist for immune evasion by sequence or conformational variability as suggested for AMA1, the antigen Rv1977 in *Mycobacterium tuberculosis* is one example involved in diversifying selection to evade host immunity.[66] This conserved region of AMA1 has been studied extensively using multiple classes of AMA1-binding molecules targeting the conserved region of AMA1 with the hope to inhibit merozoite invasion of erythrocytes *in vitro*. These probe molecules will be discussed further throughout this chapter.

1.4.3 Inhibitory interactions

The conserved nature of the hydrophobic cleft has long been considered an essential target for inhibition of host cell invasion. Evidence supporting this hypothesis has manifested itself in the form of several *in vitro* studies completed using three different classes of AMA1-binding molecules known as monoclonal antibodies (mAbs),[61,67] shark antibodies (IgNARs)[68] and small peptides identified via phage-display libraries,[69] in particular the 20-residue peptide known as R1.[70] Each class has been shown to inhibit the formation of the AMA1/RON complex through different binding modes in and around the hydrophobic cleft of AMA1.

The monoclonal antibody, known as 1F9, which binds specifically to 3D7 AMA1 via phage display, reacts with a 57-residue epitope apart of domain I of AMA1 (**Figure 1.5**). The binding of monoclonal antibodies, such as 1F9, has aided in the identification of regions of AMA1 associated with inhibition of invasion, albeit in a strain-specific manner. It has been shown that the interaction of three of the domain I loops (Ic, Id, Ie) make up most of the total buried surface area (~90%) which is attributed to 1F9 heavy chain binding of the 1F9-AMA1 complex (**Figure 1.5**). This resultant inhibition of merozoite invasion may occur from several mechanisms, one example is the reported blocking of AMA1 processing events by inhibitory antibodies.[71] The strain-specificity of 1F9 is attributed to four residues in the Id loop that make the strongest interactions with 1F9. Theses residues, known as E197, H200, F201, and D204 are understood to make the largest interactions (each contributing ~100 Å in total buried surface area) with the 1F9 antibody. Each of the four residues identified are considered polymorphic as confirmed via mutagenesis studies although the extent of polymorphism can vary, particular residue 204, which is strictly dimorphic. These four residues largely contribute to the strain-specific binding nature of 1F9 as determined via mutagenesis studies.[61] The subsequent determination of the AMA1-1F9 crystal complex was able to confirm that many of the major contacts formed within the complex were situated at one end of the hydrophobic cleft.

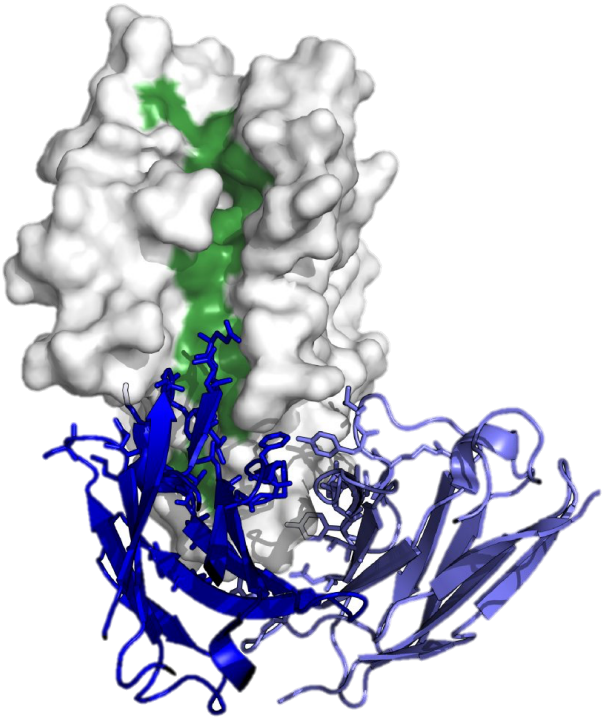


Figure 1.5 AMA1-1F9 crystal structure.[64] The top surface of AMA1 (hydrophobic cleft in green) in complex with mAb 1F9 (light chain in light blue, heavy chain in dark blue).

Another class of inhibitory antibody investigated against AMA1 is the single domain shark antibodies (IgNARs) (**Figure 1.6**). IgNARs are made up of a disulphide-bonded dimer of two protein chains, each containing a single variable and five constant domains. Owing to individual variable (V_{NAR}) domains binding antigens independently, a library of V_{NAR}s extensively variable in the complementarity-determining region (particularly CDR1 and CDR3) loops was developed and subsequently screened to explore the dynamics of V_{NAR} antigen binding, in this case against AMA1.[68] The extended CDR3 loop of these IgNARs has been shown to penetrate deep into the AMA1 hydrophobic cleft similarly to the binding mode of previously discussed mAb, 1F9.[72] In addition to IgNARs burying a large area of AMA1 surface, early generations of IgNAR also exhibited similar strain-specificity for the 3D7 *P. falciparum* strain, as determined for 1F9.[72] Subsequent targeted mutagenesis of IgNARs resulted in broader binding specificity and enhanced affinity due to the CDR3 loop extending further into the hydrophobic cleft and increasing contact with more conserved hydrophobic residues by the CDR3 loop.[72] The improved broader strain specificity identified for mutated IgNARs was the result of mutated antibodies interacting with the more conserved residues of AMA1. Although polymorphic loops contributed the largest to the resultant epitope of IgNAR binding, the opportunity to develop a strain-transcending inhibitory

molecule remained conceivable by targeting the more conserved region of hydrophobic cleft of AMA1.

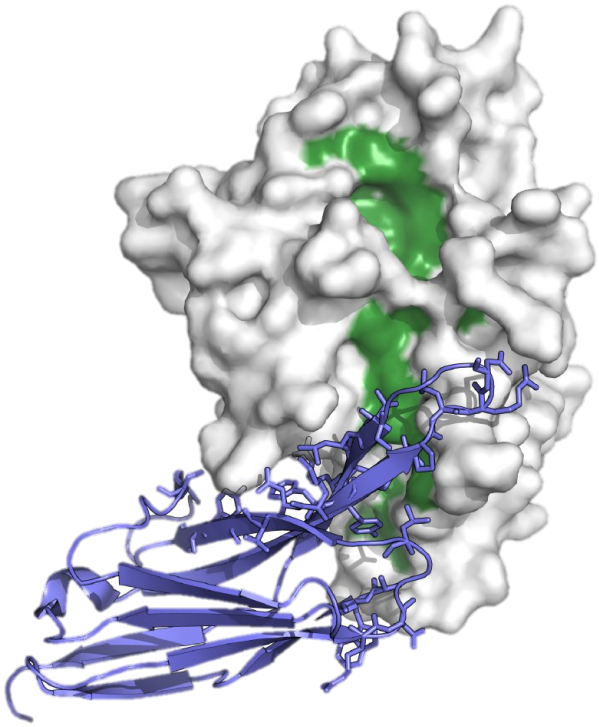


Figure 1.6 AMA1-14I-1 crystal structure.[64] The top surface representation of AMA1 (hydrophobic cleft in green) showing the complex between V_{NAR} domain 14I-1 (light blue).

Using a random peptide library displayed on the surface of filamentous phage panned on recombinant AMA1, a 20-residue peptide R1 was identified to bind *Pf*AMA1. Binding of R1 was dependent on AMA1 having the proper conformation and is strain-specific for the 3D7 and D10 *Pf*AMA1, whom share >98% identity.[70] The measured affinity of R1 binding is 100 nM.[29,70] In competition experiments, R1 was shown to partially block the binding of mAb 1F9 and IgNARs suggesting R1 also interacts near the polymorphic domain I loops.[73] R1 has also been shown to compete with a second inhibitory anti-AMA1 mAb, 4G2, known to bind to an epitope located at the base of the domain II loop, towards the more conserved end of the hydrophobic cleft.[73] These results, along with crystallography studies, recognize R1 as capable of binding the entire length of the hydrophobic cleft in a similar fashion to the endogenous ligand RON2.[73] By binding the entire length of the hydrophobic cleft, R1 is able to block the interaction between *Pf*AMA1 and the RON complex thereby inhibit the formation of the MJ.[29] Unexpectedly, the crystallization of the *Pf*AMA1-R1 complex revealed two molecules of the R1 peptide bound to *Pf*AMA1, denoted as R1

major, making the most contact with the cleft, and R1-minor, which lies slightly above R1-major (**Figure 1.7**). In total it is calculated that 3025 Å of R1's molecular surface is buried within the cleft, with R1-major making up 75% of this buried surface.[74] The solution structure of R1 identifies R1 as an elongated structure with a slight turn towards the center of the peptide due to non-polar residues residing in that region of the peptide.[70]

Although crystallography displayed the presence of R1-minor, a study by Wang *et al.*[75] upon analysis of a comparison of the experimental and predicted NMR data of the R1-AMA1 complex, deemed the R1-minor conformation most likely an artifact due to the high concentration of R1 peptide used during crystallographic study. Stoichiometric binding studies of the AMA1-R1 complex via both surface plasmon resonance (SPR) and isothermal titration calorimetry (ITC) also contrasted with the unexpected 2:1 crystallography binding, defining the R1 binding as a 1:1 affinity.[74] One key reason for the strong binding capacity of R1 can be attributed to the occupation of a deep pocket, stabilized around an arginine knob-in-hole interaction containing 6 hydrogen bonds, which resides at the more conserved end of the cleft. This preformed pocket of *Pf*AMA1 is considered to act as an anchor point for the endogenous ligand RON2, occupied by the residue Arg2041 with its guanidyl group fitting snugly inside.[74] Although each class of inhibitory molecule was able to block merozoite invasion, each class of AMA1-binding molecule investigated was considered inadequate owing to strain-specificity, particularly for *Pf*AMA1, thereby as therapeutic agents each class was limited against malaria due to the existence of multiple strains of AMA1. The development of inhibitory molecules of broad-strain activity can be overcome, to some extent, by targeting key interactions in the more conserved region of the hydrophobic cleft.

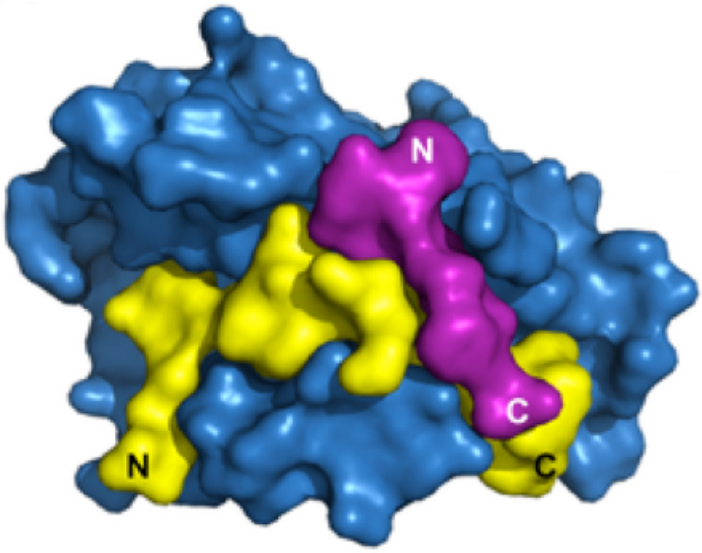


Figure 1.7 Structure of 3D7 *Pf*AMA1 crystallized with R1 peptide (3SRJ).[74] The co-crystal structure of *Pf*AMA1 (blue surface) with R1 reveals two bound peptides, R1 major (yellow) and R1 minor (purple).

1.4.4 AMA1 as a drug target

The hydrophobic cleft of AMA1 plays an important role in facilitating the formation of the AMA1/RON complex involved in the invasion of host organism cells. The conserved nature of the hydrophobic cleft characterizes the hydrophobic cleft as a useful target of action for inhibitory agents against AMA1. Knock-out studies have also shown AMA1 as a crucial antigen for parasitic survival across a variety of apicomplexan parasites.[76,77] One particular study by Yap *et al.*,[78] placed the *Pfama1* gene under conditional control using dimerizable Cre recombinase (DiCre). By controlling the expression of the gene, the expression of AMA1 was decreased by approximately 80% during one intraerythrocytic growth cycle eventually leading to a 40% reduction in growth. The reduction in growth resulted from a decreased invasion efficiency characterized by a post-invasion defect in sealing of the parasitophorous vacuole.[78] Although, the importance of AMA1 has been questioned in the past, as publicized in a study by Bargieri *et al.*,[16] which argued AMA1, rather than being essential for invasion, is a dispensable adhesion of apicomplexan zoites. This study generated *Plasmodium* sporozoites and merozoites along with *T. tachyzoites* lacking AMA1 to examine whether normal host cell penetration would be maintained. The results indicated all zoites maintaining normal host cell penetration through the junction. Despite this study, investigations into mAbs, IgNARs and inhibitory peptides as previously discussed have shown both the effective inhibition of parasitic invasion and the ability to target multiple strains of AMA1 via interactions with more conserved sections of the hydrophobic cleft.

AMA1 is essential for the invasion of all host cells by all apicomplexan zoites; thereby allowing for the possibility to target both the liver and blood stage of the disease acting against both sporozoites and merozoites respectively. A significant advantage of targeting the AMA1-RON interaction is that the site of action would be located within the bloodstream of the host organism thereby avoiding any need to cross cell membrane to target intracellular stages of the parasite. This is a significant factor in identifying AMA1 as an ideal drug target considering the implication of transporter proteins mediating resistance to existing anti-malarials such as artemisinin.[79] Although the involvement of AMA1 in both the liver and blood stage is considered an advantage, it is a significant challenge to maintain effective drug levels in the bloodstream throughout the entire time the individual is at risk. The blood stage is preferred for prophylactic intervention due to the invasive phase of the ~48-72 hour bloody stage cycle allowing for either the administration of a single dose with a sufficient plasma half-life or repeat dosing of a inhibitor with a short-lived plasma half-life.[64] The invasive machinery, AMA1 and the RON proteins, are also unique to the apicomplexa phylum and no homologous proteins exist in the human host avoiding any non-specific activity against host cells.[64] As opposed to more commonly investigated protein kinase inhibitors for malaria that require a high degree of selectivity, targeting AMA1 would avoid such a scenario, since both proteins are of parasitic origin.[80]

The development of new inhibitory agents will need to address the issue of strain-specificity that has limited previously investigated inhibitory agents. To date it has been shown that strain-transcending inhibitory agents could be developed. An example of this is the small modifications to the peptide R1 helping to improved R1 binding for multiple strains of *Pf*AMA1, such as for strains HB3 and W2mef *Pf*AMA1, which were previously unaffected.[81] Similarly, selective mutations of the AMA1-inhibitory IgNARs enhanced interactions with conserved residues of the hydrophobic cleft which led to the improved inhibition of multiple strains previously uneffected.[68,72] By avoiding the polymorphic loops concentrated towards one end of the cleft through mutagenesis, these studies highlighted targeting the conserved residues of the cleft as crucial to overcoming strain-specificity for inhibitory agents. By interacting at the more conserved end of the cleft, inhibitory agents could possess strain-transcending activity considering the cleft is conserved across the apicomplexan phylum. A deep cavity away from the polymorphic loops and underneath the domain II loop of the hydrophobic cleft was one region considered attractive, particularly for small inhibitory molecule design. AMA1 has also been considered as a candidate component of subunit vaccines being developed for malaria,[82] though the extensive polymorphisms in AMA1 are proving a major problem for vaccine development. A drug targeting the conserved hydrophobic cleft will not suffer

from this problem. Although some progress towards AMA1 small molecule inhibitors has been made, particularly by Srinivasan *et al.*[83] who investigated the pyrrol pyrimidine scaffold against AMA1, there is still much to do and learn. To date no small molecule inhibitors have been shown to bind the hydrophobic cleft of AMA1 however the fragment-based drug discovery (FBDD) strategy is one strategy being looked at to help identify binders of the hydrophobic cleft. This strategy will be discussed in further detail throughout.

1.5 Fragment-based drug discovery (FBDD)

Over the past 15 years, FBDD has become a mainstream alternative to traditional methods for generating starting points for drug discovery.[84-86] The essence of FBDD involves the screening of low molecular weight compounds (M_r < 250 Da), traditionally smaller than the drug-like lead structures often screened for high-throughput screening (HTS) campaigns.[87] These low molecular weight compounds are usually defined as having less than 20 non-hydrogen (or 'heavy') atoms as well as a lower affinity when compared to molecules usually tested in a HTS study.[87] Owing to recent practical applications of FBDD exploration over the last decade,[88,89] many reviews and analysis of the technique can be sourced to providing expert commentary and knowledge on this powerful tool designed to discover drug leads.[84,90,91]

1.5.1 History

Although the concept of FBDD has only recently become a mainstream contributor to drug discovery, the FBDD concept has been in development for over three decades. The late William Jencks is accredited to have first postulated the idea of FBDD, explaining the change in the Gibbs free energy when a molecule bound to a protein, A-B, as a summation of its component parts, A and B, labelled as the "intrinsic binding energies" of A and B.[92] Using this equation, Jencks was able to define the binding affinities of large molecules as being derived from "component parts" or fragments of the parent compound.[92] The energies of these component parts, A and B, would be represented as ΔG_A^i and ΔG_B^i respectively, and what Jencks referred to as "connection Gibbs energy", shown as ΔG^S which was derived largely from changes in translation and rotational entropy, would combine to give the total binding of the larger molecule, A-B (**Equation 1.8**).[92]

$$\textbf{Equation 1.8 - } \Delta G_{AB}^0 = \Delta G_A^i + \Delta G_B^i + \Delta G^S \qquad \textbf{(1.8)}$$

In the years that followed Jencks explanation of the change in Gibbs free energy, several studies by Nakamura *et al.*[93] and Andrews *et al.*[94] analyzed and identified which particular functional groups of large molecules contributed mostly to binding affinity. These studies further developed

the understanding of the role of functional groups in binding. Jencks explanation also helped to examine and explain an earlier study that focused on biotin, which was used as an example to identify the impact biotin's "component parts" had on binding of the larger molecule when constructed. The deconstruction of biotin into its representative fragments identified each "part" (2 in total) to bind albeit weakly, to streptavidin when compared to biotin.[95] Once it had been established that the binding affinity of a molecule was based on its components parts, the practical implementation of Jencks' total binding equation (**Equation 1.8**) posed two major challenges identified as: (1) identifying what the "components parts" are that make up the larger molecule and (2) how to subsequently link them to form the viable ligand.[92] It was this realization that held back the adoption of FBDD as a mainstay for modern drug design. Not until 1996 did the FBDD landscape begin to change, coinciding around the published work of Abbott laboratory members Shuker, Hajduk, Meadows, and Fesik that described the first practical implementation of FBDD called 'SAR by NMR'.[96] In implementing an FBDD approach, the group firstly screened fragments to identify those that bound to the protein target in question using NMR to measure binding via proton chemical shift changes. Once the group had identified a lead from screening efforts, a series of analogues based on the initial lead fragment was purchased/synthesized to optimize binding for a site identified from the protein target. The same process was completed for another site proximal to the initial binding site with the hope that two optimal fragments bound to these sites could be linked to generate a higher-affinity ligand. This was the first case where Jencks FBDD concept was implemented to develop a high-affinity ligand based on its 'component parts'.

Although several "ligand-detected" NMR techniques had already existed, such as STD NMR which is used to detect changes in NMR properties of bound fragments, it was only after 'SAR by NMR' in the mid-1990s that NMR experimental techniques became sufficiently sensitive and rapid enough for the concept of FBDD to become more practical.[97] Through 'SAR by NMR', groups were now able to provide structural information on the binding site of lead compounds, information not accessible via the use of techniques such as STD.[84] To date a number of additional methods to screen fragments have evolved, including several NMR techniques such as CPMG and HSQC, X-ray crystallography, SPR, mass spectrometry, thermal shift and ITC.[98-100] These techniques are all considered effective screening methods of fragments, however only a few can provide the necessary additional structural information considered pivotal for progression towards a successful FBDD campaign. In particular, NMR and X-ray crystallography can provide structural information while methods such as SPR, ITC and mass spectrometry can provide affinity data. One of the most commonly used methods for FBDD is still NMR owing to the ability to screen fragments under

both a protein-observed or ligand-observed context, allowing for the determination of structural information and to a smaller degree affinity via titration experiments for certain experiments.[101] In practice, a typical FBDD campaign employs a number of techniques in search of lead compounds, beginning with a primary screen, which is accomplished using mixtures that provide a coarse read-out to identify binders and non-binders. Subsequently, a secondary screen is also performed, individually assessing the binding fragments of the primary screen in an effort to verify binding affinity for the target as well as identifying possible SAR, helping to guide the design and elaboration of hit(s).[84,102]

A number of FBDD based projects have been published and commercialized over the years, each overcoming common challenges described as 1) identifying the right fragment to elaborate, and 2) elaborating the chemical structure to generate a useful lead. The development of vemurafinib (Zelboraf, PLX4032), considered the first fragment-derived drug to reach the market, is exemplary of current FBDD projects (**Figure 1.9**).[88] The development of vemurafinib began with a search for known protein kinase scaffolds, eventually leading to a select library of 20,000 compounds (ranging in molecular mass between 150 and 350 Daltons). This library was screened against a panel of multiple divergent but structurally characterized kinases.[88] The positive hits identified through the screening campaign were followed up via co-crystallography with Pim1 eventually leading to the discovery of a single ATP-binding mode for 3-anilino-7-azaindole (**Figure 1.9**). The identification of this lead led to the elaboration of the lead to probe the surrounding chemical space identified from co-crystallography in search of a high-affinity ligand. The development of this drug, which was approved for treatment of late stage melanoma in mid-2011, is considered a benchmark in the field of FBDD in which structural information aided in the conversion of a fragment into a higher affinity marketable drug.

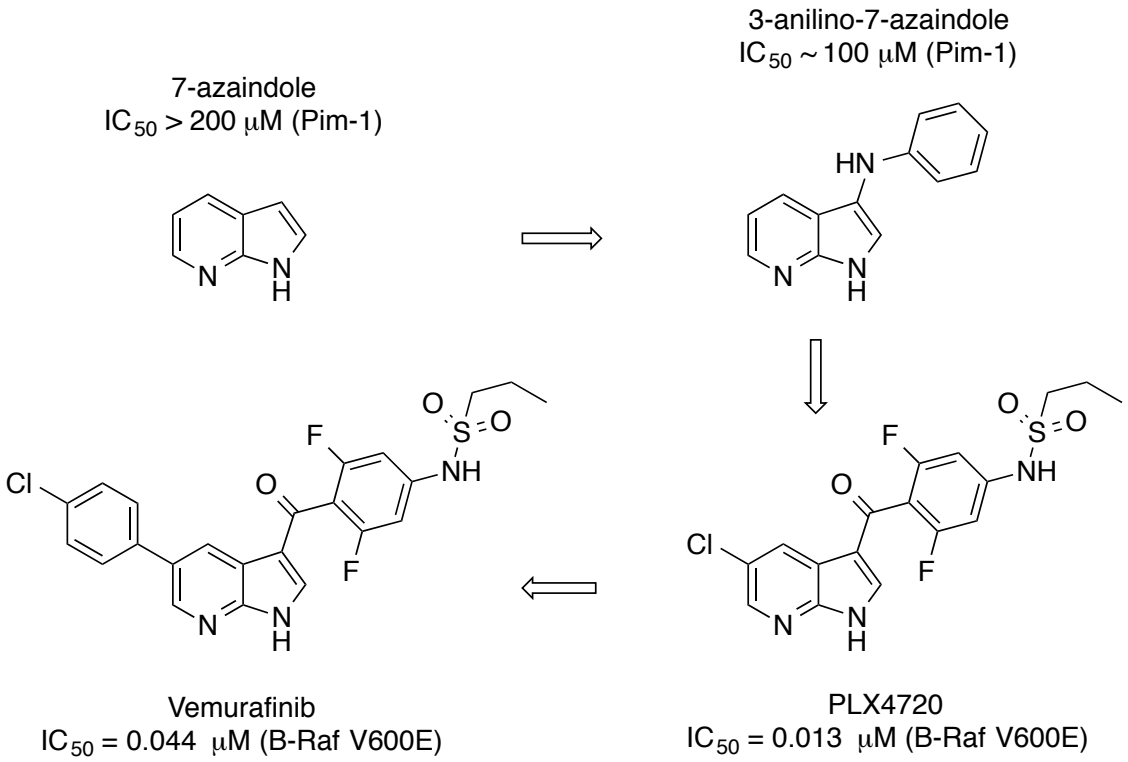


Figure 1.9 Stages of the discovery of vemurafinib. IC_{50} for each molecule is shown.[88]

The drug Venetoclax (ABT-199) was the second FBDD success story to be approved by the FDA.[103] In April 2016, Venetoclax gained the FDA's 'breakthrough therapy' designation in the treatment of certain patients with chronic lymphocytic leukaemia. Venetoclax is a Bcl-2 inhibitor; the overexpression of proteins of the anti-apoptotic Bcl-2 family is a trademark of many cancers. The first challenge faced during the development of these anti-apoptotic protein inhibitors was the need to inhibit protein-protein interactions (PPIs). Researchers at Abbott were able to screen a library of fragments using protein-detected NMR resulting in the identification of two fragments that bound two adjacent regions of Bcl-xL.[104] Owing to the presence of two adjacent sites, a suitable acylsulfonamide linker was identified and subsequent parallel synthesis led to the development of compound 1 (**Figure 1.10**), which was a potent inhibitor of Bcl-xL (K_D = 36 nM). Compound 1 also bound tightly to human serum albumin (HSA), the most abundant protein found in human blood plasma. The determination of the structural information of a close analogue of compound 1 in complex with domain III of HSA led to the development of ABT-737, later identified as a potent inhibitor of several anti-apoptotic Bcl-2 family members. The availability of co-structures with many inhibitors developed through the optimization of ABT-737 helped to design the inhibitor ABT-199 (Venetoclax), which is selective for Bcl-2 over Bcl-xL.

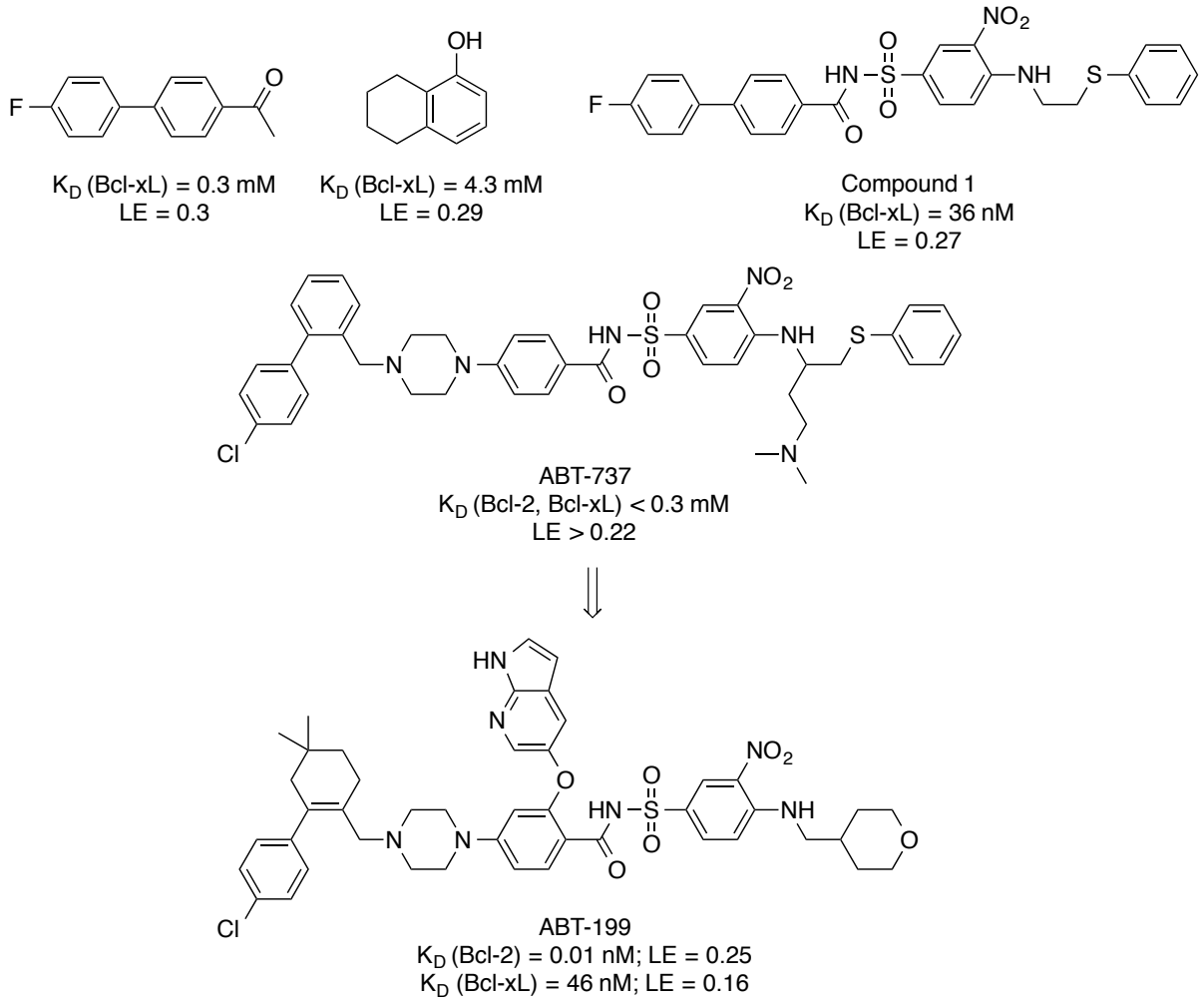


Figure 1.10 Discovery of Venetoclax through FBDD. K_D and LE are shown for each molecule.[103,104]

Although the concept of FBDD continues to be overshadowed by other hit-finding techniques (such as HTS), the continual development of FBDD remains a key priority in the field of drug discovery. FBDD still poses numerous challenges such as the content of fragment libraries, technology available for monitoring fragments and the identification of ways to optimize leads; the capability to target difficult PPIs were other techniques have failed remains a key reason behind FBDD's acceptance and utility within the industry and academia.[87]

1.5.2 Defining a fragment

FBDD is predicated on the notion that a small fragment can be identified and then either grown, merged, or linked with another fragment to improve potency. Several criteria describing the metrics of fragments have been established, more notably the 'rule of 3' proposed by Astex

Therapeutics.[105] The 'rule of 3' uses guidelines of $M_r \leq 300$, ≤ 3 H-bond donors, ≤ 3 H-bond acceptors and a CLogP (calculated partition coefficient of a compound) ≤ 3; derived from Chris Lipinski's 'rule of 5' which is used to describe the physicochemical properties of most orally available drugs.[106] Contrary to Lipinski's entitled 'rule of 5' theory, as opposed to stringent rules the 'rule of 5' should be regarded as guidelines to follow when assessing low molecular weight fragment hits. Although over the years, owing to differences in opinion, many different compounds have been considered fragments due to possessing low complexity and low molecular weight contributing to a low affinity even after satisfying the 'rule of 3' criteria.

1.5.3 Advantages of FBDD

One of the distinctive successes of FBDD is that targets that show poor druggability owing to difficult PPIs being investigated are more amenable to FBDD approaches when other methods of drug discovery/lead generation such as HTS have failed.[87,107] Methods such as HTS have failed to identify inhibitors of PPIs owing to their binding sites often being fairly flat and the initial hit compounds binding weakly.[87] Even though methods such as HTS and combinatorial chemistry are able to screen larger sets of compounds, FBDD is able to explore chemical space more efficiently since the estimated drug-like chemical space coverage of FBDD is 10^{60} compounds in comparison to even the largest HTS library, estimated at 10^6-10^7 molecules.[108,109] The cost associated with the screening and maintenance of fragment libraries is another factor which has allowed FBDD to become more readily accessible considering the costs associated with utilizing large HTS libraries can be significant in comparison.[84] In addition, HTS-based lead identification for novel targets is nearly twice as likely to fail as for known targets.[110]

1.5.4 Disadvantages of FBDD

FBDD does possess some disadvantages, in particular, the need for sensitive and robust screening methods to identify low affinity fragment binders. This is particularly the case in functional assays in which fragments are too weak to show activity.[84] This is particularly important in the projects that would like to link two fragments from adjacent binding sites on the one target, few methods are available to help researchers distinguish between the two separate binding sites.[111] Despite a reliance on biophysical assays to screen low affinity fragments, the two most widely used and most informative methods, NMR and X-ray crystallography, are continually being developed to improve data output for FBDD campaigns.

Another disadvantage of FBDD is the requirement of optimization once a lead has been identified to produce higher affinity compounds. This factor quite often requires multiple rounds of medicinal chemistry and testing which would demand more funding whilst also taking a longer period of time to be successful. The success of an adequate follow-up of lead compounds is also heavily dependent on the availability of structural data for the protein target, this could hamper an FBDD campaign if the binding site of a compound of interest is not clearly defined. The presence of non-specific binding when targeting a large protein surface using fragments is another factor that can affect SAR determination and thus negatively impact the optimization of the lead compound. This potential lack of selectivity from small compounds can be overcome however it can also hinder the subsequent optimization of leads. Despite these issues, the concept of FBDD has become a mainstay in the medicinal chemistry field and is widely deployed in both academia and commercial laboratories.

1.5.5 Evolution of fragments

The elaboration of fragments poses a number of challenges namely the selection of an ideal optimization strategy (e.g. growth, merge, link *etc.*), the monitoring of physicochemical properties throughout the optimization process and synthetic feasibility.[84] As previously discussed, structural information has become vital for defining the optimal vectors and strategies for fragment evolution to avoid negative or flat SAR, which can be difficult to interpret. Identifying the ideal combination of these factors is crucial to the evolution of fragments since they also help chemists to rank and prioritize leads to ensure the success of the FBDD campaign. The most common strategy for improving binding affinity is known as the traditional growth method involving the initial identification of a positive hit followed by multiple rounds of design, synthesis and screening. As like all FBDD campaigns, each interesting fragment is identified through a tailored screening campaign using multiple methods such as NMR and X-ray crystallography. It is these methods that help to determine compound binding/activity and identify structural information all whilst monitoring acceptable physicochemical properties such as ligand efficiency (LE). This method relies heavily on the quality of structural data associated with fragments of interest. The development of vemurafinib is a classic example of this strategy, with X-ray crystallography being the driving force behind development of vemurafinib.[88]

Another technique employed is the merging technique utilized when two compounds occupying a similar binding pocket are shown to extend into different chemical regions. One example of this strategy is the formation of novel p38α MAP kinase inhibitors by Gill *et al.*[89] This work discussed

the structure-guided optimization of the molecular fragments 2-amino-3-benzyl-oxypyridine (IC_{50} 1.3 mM) and 3-(2-(4-pyridyl)ethyl)indole (IC_{50} 35 μM) identified using X-ray crystallographic screening of p38α MAP kinase. By reviewing the literature, the group was able to identify a non-selective kinase inhibitor, which when crystallized with their protein target, displayed overlap with the two fragments identified in their crystallography screen. The subsequent merging of the literature inhibitor with their hits gave two high-affinity analogues with improved selectivity towards p38α MAP kinase. This method seems simple enough however one aspect that must be considered is the binding modes of each individual fragment used to make up the merged final compound, which could negatively impact overall binding affinity if not maintained as per individual pharmacophores.

The least favoured strategy for the elaboration of fragments is fragment linking. At face value, the linking strategy is an attractive prospect owing to the hypothesized summation of affinities derived from the two binders of adjacent sites[111] but the choice and length of the linker plays a huge role in defining final compound affinity. This strategy is considered the most difficult method to improve fragment affinity. The length and type of the linker must take into account the binding mode of each adjacent fragment and the available chemical space between the two adjacent fragments.[112] The linker method is also heavily reliant on good structural data to aid in the selection of a linker type and length suitable for the protein target in question.[113] A recent example of this strategy was work targeted against lactate dehydrogenase A (LDHA) that followed a traditional path of initial growth via an SAR study using X-ray crystallography, followed by the utilization of a linking step to give the final lead compound.[114] The presence of a secondary binding site in the crystal structure meant that a specifically designed library would be employed to screen this secondary site for possible hits considering the additional site required inhibitory activity to be effective. By targeting this additional site the opportunity to link binders of both sites was applicable. The selection of which FBDD strategy to employ is made on a case-by-case basis which is usually influenced by the level of data available for the FBDD campaign, particularly structural data which plays a critical role in determining how useful each method can be.

1.6 Previous work and basis for current work

An FBDD study of AMA1 was favoured over the investigation of existing peptides or antibodies considering the lower production costs and potential oral bioavailability associated with using small molecule inhibitors as a potential malarial treatment.[115] Our group has ongoing investigations of AMA1 and the development of inhibitors in collaboration with the Anders group at La Trobe University. This FBDD investigation began with the completion of a screening campaign of the Monash Institute of Pharmaceutical Sciences (MIPS) fragment library (~1200 compounds) against AMA1. The primary screen of the MIPS library employed STD NMR; fragments were screened in cocktails of 6 compounds helping to efficiently identify weak binders of AMA1 for further analysis. The STD NMR technique will be explained in more detail separately. Considering STD NMR provided no structural information on the binding sites of positive hits; R1, a known high-affinity binder of the 3D7 hydrophobic cleft and to a lesser extent FVO (unpublished data), was utilized to compete off bound compounds to identify fragments interacting in or around the hydrophobic cleft.[116] The introduction of another NMR technique, known as CPMG NMR also helped to identify compounds that bound the hydrophobic cleft. This technique will also be discussed in detail separately. Higher concentrations of fragments and 3D7 *Pf*AMA1 (500 and 10mM, respectively, compared with 300 and 5mM, respectively, in the initial cocktail screens) were employed to increase the signal intensity and make any competitive effect easier to detect. Only fragment hits competed off by R1 when examined using both STD and CPMG NMR were characterized further via SPR, in order to estimate the binding affinity of each hit for AMA1. SPR was considered the secondary screen for the FBDD campaign against AMA1. The characterization of the protein target, AMA1 was also investigated in parallel to this fragment screen in collaboration with other members of the group. The strains 3D7 and FVO *Pf*AMA1, which share a sequence homology of 94%,[117] were employed throughout this FBDD study for mapping studies owing to the continual efforts within the group to assign all observable backbone resonances (NH, ^{15}N, $^{13}C^{\alpha}$, $^{13}C'$) for both 3D7 and FVO *Pf*AMA1.

1.6.1 Saturation transfer difference NMR

STD NMR spectroscopy has emerged as one of the most popular ligand-based NMR techniques for the study of protein-ligand interactions.[118] STD spectroscopy relies on the selective saturation of protons of the macromolecular receptor, which then allows saturation to quickly propagate across the entire receptor due to effective spin diffusion. STD is an ideal technique when assessing small molecules that generally display weak binding affinity owing to the critical need of STD for fast exchange between free and bound ligands to work. If the small molecule ligand binds the receptor, saturation will then spread onto the ligand. The resultant 1D spectrum will show the intensity of the ligand signal attenuated. Subtraction of the resulting spectrum from the reference spectrum without saturation yields the STD spectrum, which contains only signals of the saturated protons of the ligand (**Figure 1.11**).[119] The success of this technique is a consequence of its robustness requiring minimal amounts of unlabeled protein for testing. For our purpose, a competition experiment was conducted using STD NMR in the presence and absence of the competing R1 peptide. The R1 peptide was added to compete off any positive hits interacting at the hydrophobic cleft which would be observed as a decreased saturation transfer from the protein to fragments and hence display a reduction in STD signal intensity if competed off by R1 (**Figure 1.11**).

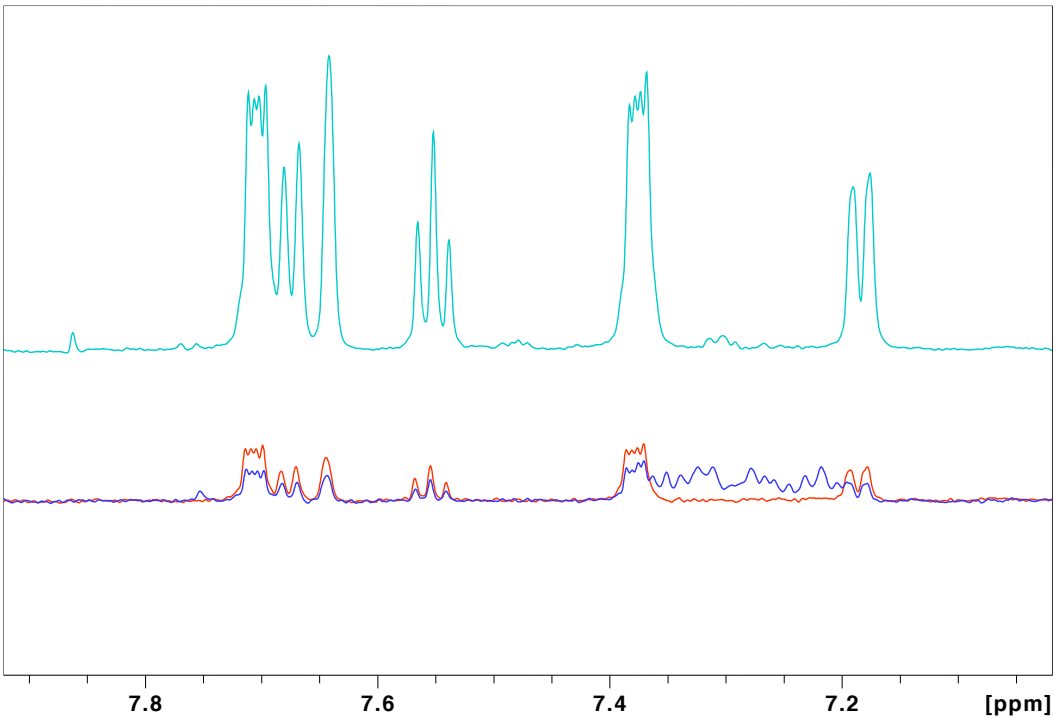


Figure 1.11 Example STD NMR data for a compound displaying the compound's ^{1}H reference spectrum (cyan) along with the R1 absent STD spectrum (red) and R1 present STD spectrum (blue).

1.6.2 Carr-Purcell-Meiboom-Gill NMR

The CPMG pulse train is a fundamental component of pulse sequences used for the measurement of dynamic processes by NMR spectroscopy. CPMG spectroscopy is a technique widely employed in FBDD that measures the transverse or spin-spin T_2 relaxation times of any nucleus. Once a fragment interacts with a protein target, its relaxation rate is affected. CPMG NMR was used to facilitate a competition experiment in both the presence and absence of the competing R1 peptide. The peptide was expected to compete with and displace any weakly bound fragments interacting at the hydrophobic cleft. If a fragment is competed off by R1, this would be indicated by an increase in fragment signal when compared to the CPMG spectra of the R1 absent sample (**Figure 1.12**). An increase in fragment signal would result from a reduction in the transverse relaxation rate for the fragment owing to the free fragment being in fast exchange and no longer bound to the protein.[116]

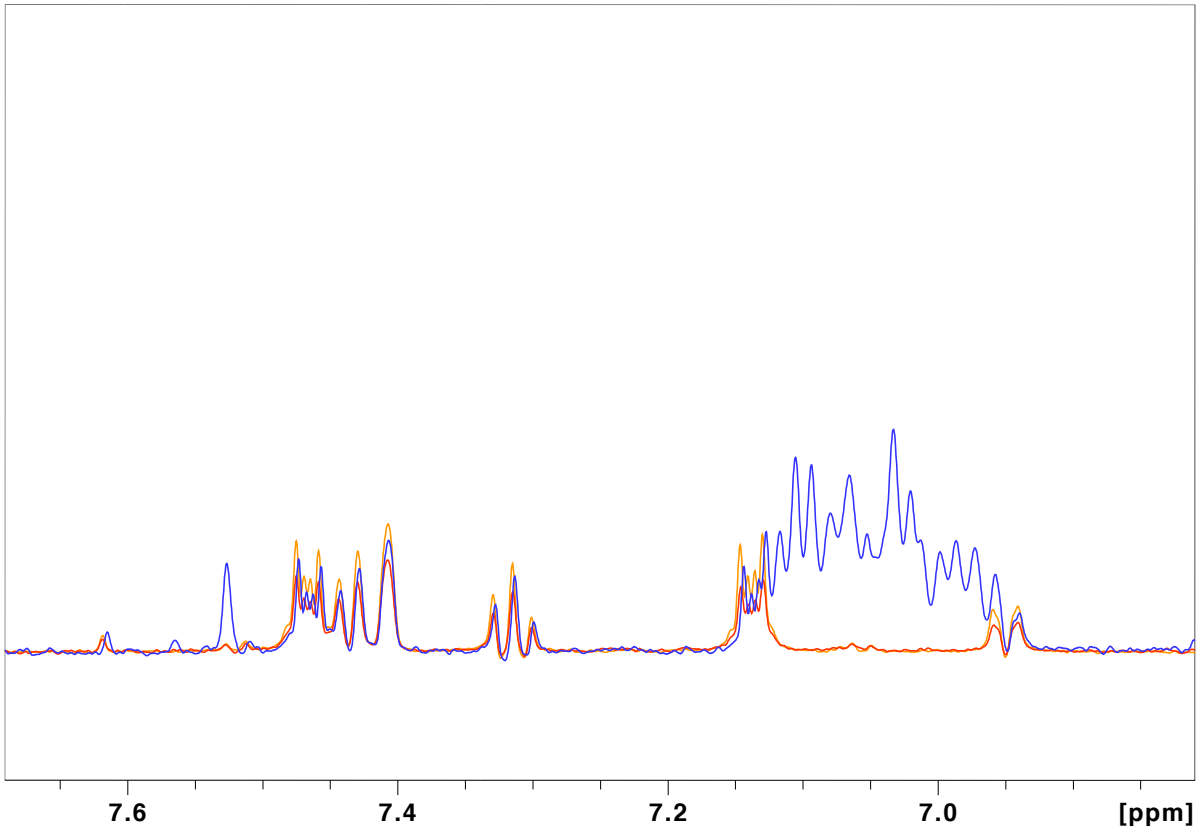


Figure 1.12 Example CPMG NMR data for a compound displaying the compound's ^{1}H NMR reference spectrum (orange) along with the R1 absent CPMG spectrum (red) and the R1 present CPMG spectrum (blue).

1.6.3 AMA1 screening

A fragment-based screening approach was employed using the MIPS fragment library of low-molecular weight molecules against the hydrophobic cleft of AMA1. A screen involving 1140 fragments (initially screened as cocktails of 6 fragments each at 300 μM, 5 μM AMA1) was completed on the 3D7 *Pf*AMA1$_{[104-442]}$ corresponding to the domain I and II of the AMA1 ectodomain (**Figure 1.13**).[116] Each fragment was tested via a STD NMR cocktail screen, eventually affording 208 hits at a hit rate of 18 %. Owing to the lack of structural information on the binding sites of the 208 hits, the introduction of both an STD and CPMG NMR competition experiment to identify hits interacting with the hydrophobic cleft. By this means, 90 fragments were shown to be false positives and 57 were competed by R1 in both STD and CPMG experiments. Overall our fragment screening approach using the two NMR techniques described previously was able to identify 57 fragments from the fragment library interacting with the hydrophobic cleft of *Pf*AMA1, at a hit rate of 5 %.

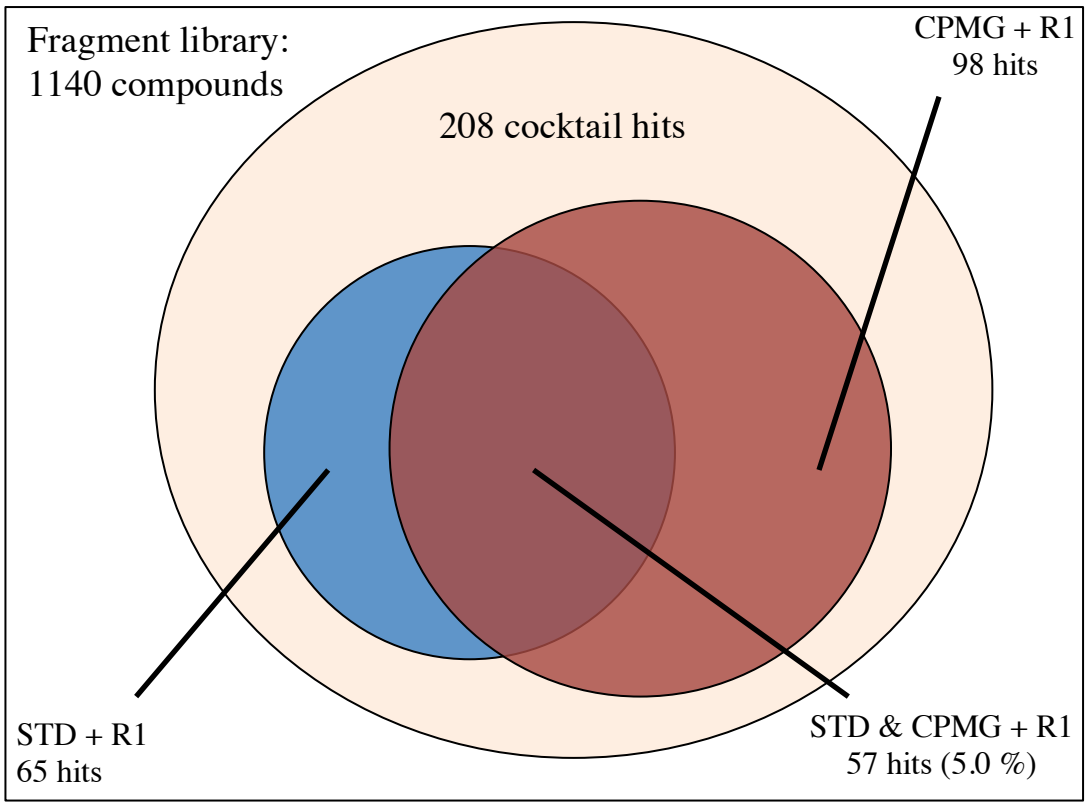


Figure 1.13 Venn diagram showing the numbers of hits identified in the primary cocktail screen and R1 competition experiments.

The identified R1-competing hits varied structurally with fragments of similar structure clustered into different classes either described as 2-aminothiazoles, 2-arylfurans, benzimidazoles, among others (**Figure 1.14**). An analysis of the 57 hits showed an obvious trend in the partition coefficient (log*P*) values between R1-competing hits and the non-competing fragments, with the hits appearing to be more hydrophobic (higher log*P*) on average than the rest of the library compounds. As the cleft is largely hydrophobic, a preference for the overall more non-polar fragments was expected. The R1-competing hits also have slightly higher molecular weights, with the increase in size being attributable to increasing numbers of rings, such as the thiazole and benzimidazole series, rather than additional rotatable bonds.

Each competing hit of 3D7 *Pf*AMA1 was subsequently tested by SPR against immobilized 3D7 *Pf*AMA1 at 50, 100 and 200 mM to estimate their binding affinities. Of the evaluated 57 fragments, 46 compounds showed binding by SPR. It is believed that the remaining 11 compounds simply showed binding affinities that were too weak to detect via SPR. For all fragments tested, binding to AMA1 failed to reach saturation over the elected concentration range consistent with the relatively weak affinity expected for primary fragment hits. All except two hits showed binding affinities weaker than 1 mM, with the strongest hits displaying K_D values of approximately 600 μM. Most of the 46 hits identified via SPR (67.4 %), bound AMA1 with a LE of between 0.2 and 0.3 kcal mol^{-1} heavy $atom^{-1}$. The LE is a measure of the binding energy per atom of a ligand to its binding partner. A measure of the LE assists in describing the optimal combinations of physicochemical properties and pharmacological properties in search of lead compounds.[86] For all chemical classes identified, no X-ray crystal structures were solved either owing to low-resolution crystal structures or missing fragment electron densities hence any SAR investigation of a certain chemical class identified from the library would use competition NMR experiments coupled with SPR affinity determination as the primary assay to track the progress of elaborated analogues.

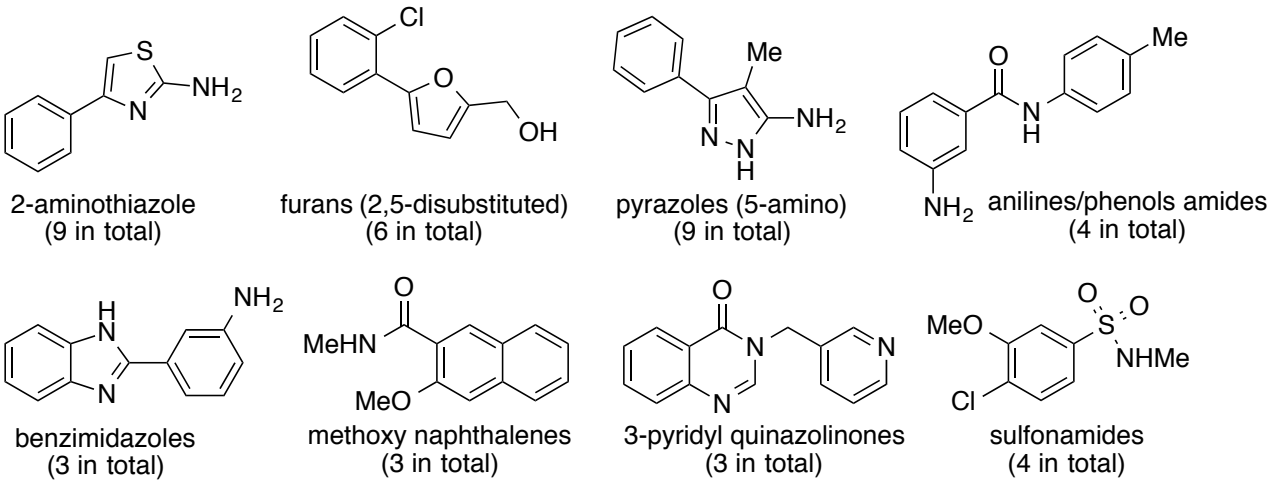


Figure 1.14 Different chemical classes identified during screening campaign. The number of fragments per class is also shown. A miscellaneous group was introduced for the remaining 5 fragments identified as hits for AMA1.

Following on from the initial fragment screen, the focus of this work was to identify small molecule inhibitors of AMA1 utilizing conventional FBDD methods of testing and synthesis in the process. The outcomes of the fragment screen previously mentioned led to the elaboration of the 2-aminothiazole series and eventually followed by an investigation into the benzimidazole series. A relative ranking based on the calculated LE for each fragment hit helped identify the 2-aminothiazole and benzimidazole classes as the most desirable starting points in search of higher affinity strain-transcending inhibitors of AMA1. The concept of LE defines an 'optimal' drug as having a LE of 0.29 kcal mol^{-1} heavy atom,[120,121] thereby a comparison of the LE for all positive hits led to the selection of these two scaffolds. In FBDD, LE is considered an important property for initially ranking and tracking the progress of fragment optimization. The synthetic feasibility of each binding scaffold also played a role during the selection process. Each investigation into a select chemical class intended to develop weakly but specifically binding fragment hits into more potent lead compounds. Conventional FBDD techniques such as SPR, STD and CPMG NMR and too a lesser extent X-ray crystallography were all employed to help elucidate binding affinity and the binding mode of selected scaffolds. The affinities of each of the library hits are not shown considering the weak affinity of hits can make accurate measurements of affinity challenging hence it was anticipated that as affinity was gained, these measurements would become more reliable. In order to improve the accuracy of affinity measurement, all analogues developed for a certain binding scaffold were tested as a 5-point dose response as opposed to the 3-point analysis of library fragments. All relative outcomes will be discussed throughout including additional studies

completed in parallel to the elaboration of the 2-aminothiazole and benzimidazole scaffold that influenced the direction of the overall study.

1.7 Bibliography

(1) World Health Organization World Malaria Report 2017. **2017**, 1–196.
(2) Ross, R. Observations on Malaria Parasites made in Secunderabad, Deccan. *British Medical Journal* **1896**, *1*, 260–261.
(3) Guerra, C. A.; Tatem, A. J.; Noor, A. M.; Snow, R. W. The global distribution and population at risk of malaria: past, present, and future. *The Lancet Infectious Diseases* **2004**, *4*, 327–36.
(4) Breman, J. G.; Alilio, M. S.; Mills, A. Conquering the intolerable burden of malaria: what's new, what's needed: a summary. *The American Journal of Tropical Medicine and Hygiene* **2004**, *71*, 1–15.
(5) Naing, C.; Whittaker, M. A.; Nyunt Wai, V.; Mak, J. W. Is *Plasmodium vivax* malaria a severe malaria?: a systematic review and *meta*-analysis. *PLoS Neglected Tropical Diseases* **2014**, *8*, e3071.
(6) World Health Organization World malaria report 2015 - without profiles. **2015**, 1–157.
(7) Fidock, D. A.; Eastman, R. T.; Ward, S. A.; Meshnick, S. R. Recent highlights in antimalarial drug resistance and chemotherapy research. *Trends in Parasitology* **2008**, *24*, 537–544.
(8) Martinsen, E. S.; Perkins, S. L.; Schall, J. J. A three-genome phylogeny of malaria parasites (*Plasmodium* and closely related genera): Evolution of life-history traits and host switches. *Molecular Phylogenetics and Evolution* **2008**, *47*, 261–273.
(9) Cowman, A. F.; Healer, J.; Marapana, D.; Marsh, K. Malaria: Biology and Disease. *Cell* **2016**, *167*, 610–624.
(10) Coppi, A.; Tewari, R.; Bishop, J. R.; Bennett, B. L.; Lawrence, R.; Esko, J. D.; Billker, O.; Sinnis, P. Heparan Sulfate Proteoglycans Provide a Signal to *Plasmodium* Sporozoites to Stop Migrating and Productively Invade Host Cells. *Cell Host & Microbe* **2007**, *2*, 316–327.
(11) Tavares, J.; Formaglio, P.; Thiberge, S.; Mordelet, E.; Van Rooijen, N.; Medvinsky, A.; Ménard, R.; Amino, R. Role of host cell traversal by the malaria sporozoite during liver infection. *Journal of Experimental Medicine* **2013**, *210*, 905–915.
(12) Ishino, T.; Yano, K.; Chinzei, Y.; Yuda, M. Cell-Passage Activity Is Required for the Malarial Parasite to Cross the Liver Sinusoidal Cell Layer. *PLoS Biology* **2004**, *2*, 77-82.
(13) Risco-Castillo, V.; Topçu, S.; Marinach, C.; Manzoni, G.; Bigorgne, A. E.; Briquet, S.; Baudin, X.; Lebrun, M.; Dubremetz, J-F.; Silvie, O. Malaria Sporozoites Traverse Host Cells within Transient Vacuoles. *Cell Host & Microbe* **2015**, *18*, 593–603.
(14) Bhanot, P.; Schauer, K.; Coppens, I.; Nussenzweig, V. A Surface Phospholipase Is Involved in the Migration of *Plasmodium* Sporozoites through Cells. *Journal of Biological Chemistry* **2005**, *280*, 6752–6760.
(15) Herrera, R.; Anderson, C.; Kumar, K.; Molina-Cruz, A.; Nguyen, V.; Burkhardt, M.; Reiter, K.; Shimp, R., Jr.; Howard, R. F.; Srinivasan, P.; Nold, M. J.; Ragheb, D.; Shi, L.; DeCotiis, M.; Aebig, J.; Lambert, L.; Rausch, K. M.; Muratova, O.; Jin, A.; Reed, S. G.; Sinnis, P.; Barillas-Mury, C.; Duffy, P. E.; MacDonald, N. J.; Narum, D. L. Reversible Conformational Change in the *Plasmodium falciparum* Circumsporozoite Protein Masks Its Adhesion Domains. *Infection and Immunity* **2015**, *83*, 3771–3780.
(16) Bargieri, D. Y.; Andenmatten, N.; Lagal, V.; Thiberge, S.; Whitelaw, J. A.; Tardieux, I.; Meissner, M.; nard, R. M. E. Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion. *Nature Communications* **2014**, *4*, 1–13.
(17) Sturm, A.; Amino, R.; Van de Sand, C.; Regen, T. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **2006**, 313, 1287-1290.
(18) Weiss, G. E.; Gilson, P. R.; Taechalertpaisarn, T.; Tham, W-H.; de Jong, N. W. M.; Harvey, K. L.; Fowkes, F. J. I.; Barlow, P. N.; Rayner, J. C.; Wright, G. J.; Cowman, A. F.; Crabb, B. S. Revealing the Sequence and Resulting Cellular Morphology of Receptor-Ligand Interactions during *Plasmodium falciparum* Invasion of Erythrocytes. *PLoS Pathogens* **2015**, 1–25.

(19) Lin, C. S.; Uboldi, A. D.; Epp, C.; Bujard, H.; Tsuboi, T.; Czabotar, P. E.; Cowman, A. F. Multiple *Plasmodium falciparum* Merozoite Surface Protein 1 Complexes Mediate Merozoite Binding to Human Erythrocytes. *Journal of Biological Chemistry* **2016**, *291*, 7703–7715.
(20) Bannister, L. H.; Mitchell, G. H.; Butcher, G. A.; Dennis, E. D.; Cohen, S. Structure and development of the surface coat of erythrocytic merozoites of *Plasmodium knowlesi*. *Cell Tissue Research* **1986**, *245*, 281–290.
(21) Das, S.; Hertrich, N.; Perrin, A. J.; Withers-Martinez, C.; Collins, C. R.; Jones, M. L.; Watermeyer, J. M.; Fobes, E. T.; Martin, S. R.; Saibil, H. R.; Wright, G. J.; Treeck, M.; Epp, C.; Blackman, M. J. Processing of *Plasmodium falciparum* Merozoite Surface Protein MSP1 Activates a Spectrin-Binding Function Enabling Parasite Egress from RBCs. *Cell Host & Microbe* **2015**, *18*, 433–444.
(22) Tham, W.-H.; Healer, J.; Cowman, A. F. Erythrocyte and reticulocyte binding-like proteins of *Plasmodium falciparum*. *Trends in Parasitology* **2012**, *28*, 23–30.
(23) Singh, S.; Alam, M. M.; Pal-Bhowmick, I.; Brzostowski, J. A.; Chitnis, C. E. Distinct External Signals Trigger Sequential Release of Apical Organelles during Erythrocyte Invasion by Malaria Parasites. *PLoS Pathogens* **2010**, *6*, e1000746.
(24) Paul, A. S.; Saha, S.; Engelberg, K.; Jiang, R. H. Y.; Coleman, B. I.; Kosber, A. L.; Chen, C.-T.; Ganter, M.; Espy, N.; Gilberger, T. W.; Gubbels, M.-J.; Duraisingh, M. T. Parasite Calcineurin Regulates Host Cell Recognition and Attachment by Apicomplexans. *Cell Host & Microbe* **2015**, *18*, 49–60.
(25) Cowman, A. F.; Crabb, B. S. Invasion of Red Blood Cells by Malaria Parasites. *Cell* **2006**, *124*, 755–766.
(26) Keeley, A.; Soldati, D. The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. *Trends in Cell Biology* **2004**, *14*, 528–532.
(27) Blackman, M. J.; Bannister, L. H. Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation. *Molecular & Biochemical Parasitology* **2001**, *117*, 11-25.
(28) Riglar, D. T.; Richard, D.; Wilson, D. W.; Boyle, M. J.; Dekiwadia, C.; Turnbull, L.; Angrisano, F.; Marapana, D. S.; Rogers, K. L.; Whitchurch, C. B.; Beeson, J. G.; Cowman, A. F.; Ralph, S. A.; Baum, J. Super-Resolution Dissection of Coordinated Events during Malaria Parasite Invasion of the Human Erythrocyte. *Cell Host & Microbe* **2011**, *9*, 9–20.
(29) Richard, D.; MacRaild, C. A.; Riglar, D. T.; Chan, J. A.; Foley, M.; Baum, J.; Ralph, S. A.; Norton, R. S.; Cowman, A. F. Interaction between *Plasmodium falciparum* Apical Membrane Antigen 1 and the Rhoptry Neck Protein Complex Defines a Key Step in the Erythrocyte Invasion Process of Malaria Parasites. *Journal of Biological Chemistry* **2010**, *285*, 14815–14822.
(30) Collins, C. R.; Hackett, F.; Strath, M.; Penzo, M.; Withers-Martinez, C.; Baker, D. A.; Blackman, M. J. Malaria Parasite cGMP-dependent Protein Kinase Regulates Blood Stage Merozoite Secretory Organelle Discharge and Egress. *PLoS Pathogens* **2013**, *9*, e1003344.
(31) Srinivasan, P.; Beatty, W. L.; Diouf, A.; Herrera, R.; Ambroggio, X.; Moch, J. K.; Tyler, J. S.; Narum, D. L.; Pierce, S. K.; Boothroyd, J. C.; Haynes, J. D.; Miller, L. H. Binding of Plasmodium merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 13275-13280.
(32) Joice, R.; Nilsson, S. K.; Montgomery, J.; Dankwa, S.; Egan, E.; Morahan, B.; Seydel, K. B.; Bertuccini, L.; Alano, P.; Williamson, K. C.; Duraisingh, M. T.; Taylor, T. E.; Milner, D. A.; Marti, M. *Plasmodium falciparum* transmission stages accumulate in the human bone marrow. *Science Translational Medicine* **2014**, *6*, 244re5.
(33) Morrissette, N. S.; Ajioka, J. W. The early years of Toxoplasma research: What's past is prologue. *International Journal for Parasitology* **2009**, *39*, 865–869.
(34) Butler, A. R.; Khan, S.; Ferguson, E. A brief history of malaria chemotherapy. *Journal of the Royal College of Physicians of Edinburgh* **2010**, *40*, 172–177.

(35) Meshnick, S. R.; Taylor, T. E.; Kamchonwongpaisan, S. Artemisinin and the antimalarial endoperoxides: from herbal remedy to targeted chemotherapy. *Microbiological Reviews.* **1996**, *60*, 301–315.
(36) White, N. J. Qinghaosu (artemisinin): the price of success. *Science* **2008**, *320*, 330-334.
(37) World Health Organization. Global report on antimalarial drug efficacy and drug resistance: 2000-2010. *World Health Organization* **2010**, 1-115.
(38) Denis, M. B.; Tsuyuoka, R.; Poravuth, Y.; Narann, T. S.; Seila, S.; Lim, C.; Incardona, S.; Lim, P.; Sem, R.; Socheat, D.; Christophel, E. M.; Ringwald, P. Surveillance of the efficacy of artesunate and mefloquine combination for the treatment of uncomplicated falciparum malaria in Cambodia. *Tropical Medicine & International Health* **2006**, *11*, 1360–1366.
(39) Newton, P. N.; Fernández, F. M.; Plançon, A.; Mildenhall, D. C.; Green, M. D.; Ziyong, L.; Christophel, E. M.; Phanouvong, S.; Howells, S.; McIntosh, E.; Laurin, P.; Blum, N.; Hampton, C. Y.; Faure, K.; Nyadong, L.; Soong, C. W. R.; Santoso, B.; Zhiguang, W.; Newton, J.; Palmer, K. A collaborative epidemiological investigation into the criminal fake artesunate trade in South East Asia. *PLoS Medicine* **2008**, *5*, e32.
(40) Dondorp, A. M.; Yeung, S.; White, L.; Nguon, C.; Day, N. P. J.; Socheat, D.; Seidlein, von, L. Artemisinin resistance: current status and scenarios for containment. *Nature Reviews Drug Discovery* **2010**, *8*, 272–280.
(41) Stepniewska, K.; Ashley, E.; Lee, S. J.; Anstey, N.; Barnes, K. I.; Binh, T. Q.; D'Alessandro, U.; Day, N. P. J.; de Vries, P. J.; Dorsey, G.; Guthmann, J. P.; Mayxay, M.; Newton, P. N.; Olliaro, P.; Osorio, L.; Price, R. N.; Rowland, M.; Smithuis, F.; Taylor, W. R. J.; Nosten, F.; White, N. J. In Vivo Parasitological Measures of Artemisinin Susceptibility. *Journal of Infectious Diseases* **2010**, *201*, 570–579.
(42) Hastings, I. How artemisinin-containing combination therapies slow the spread of antimalarial drug resistance. *Trends in Parasitology* **2011**, *27*, 67–72.
(43) White, N. J.; Pongtavornpinyo, W. The de novo selection of drug-resistant malaria parasites. *Proceedings of the Royal Society B: Biological Sciences* **2003**, *270*, 545–554.
(44) Peterson, M. G.; Marshall, V. M.; Smythe, J. A.; Crewther, P. E.; Lew, A.; Silva, A.; Anders, R. F.; Kemp, D. J. Integral membrane protein located in the apical complex of *Plasmodium falciparum*. *Molecular and Cellular Biology.* **1989**, *9*, 3151–3154.
(45) Stahl, H. D.; Bianco, A. E.; Crewther, P. E.; Burkot, T.; Coppel, R. L.; Brown, G. V.; Anders, R. F.; Kemp, D. J. An asparagine-rich protein from blood stages of *Plasmodium falciparum* shares determinants with sporozoites. *Nucleic Acids Research* **1986**, *14*, 3089–3102.
(46) Holder, A. A.; Freeman, R. R.; Uni, S.; Aikawa, M. Isolation of a *Plasmodium falciparum* rhoptry protein. *Molecular and Biochemical Parasitology* **1985**, *14*, 292-303.
(47) Silvie, O.; Franetich, J.-F. O.; Charrin, S. P.; Mueller, M. S.; Siau, A.; Bodescot, M.; Rubinstein, E.; Hannoun, L.; Charoenvit, Y.; Kocken, C. H.; Thomas, A. W.; van Gemert, G.-J.; Sauerwein, R. W.; Blackman, M. J.; Anders, R. F.; Pluschke, G.; Mazier, D. A Role for Apical Membrane Antigen 1 during Invasion of Hepatocytes by *Plasmodium falciparum* Sporozoites. *Cellular Microbiology* **2004**, *279*, 1823–1833.
(48) Alexander, D. L.; Mital, J.; Ward, G. E.; Bradley, P.; Boothroyd, J. C. Identification of the Moving Junction Complex of Toxoplasma gondii: A Collaboration between Distinct Secretory Organelles. *PLoS Pathogens* **2005**, *1*, e17.
(49) Alexander, D. L.; Arastu-Kapur, S.; Dubremetz, J. F.; Boothroyd, J. C. *Plasmodium falciparum* AMA1 Binds a Rhoptry Neck Protein Homologous to TgRON4, a Component of the Moving Junction in Toxoplasma gondii. *Eukaryotic Cell* **2006**, *5*, 1169–1173.
(50) Cao, J.; Kaneko, O.; Thongkukiatkul, A.; Tachibana, M.; Otsuki, H.; Gao, Q.; Tsuboi, T.; Torii, M. Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in *Plasmodium falciparum* merozoites. *Parasitology International* **2009**, *58*, 29–35.

(51) Besteiro, S. B.; Michelin, A.; Poncet, J. L.; Dubremetz, J.-F. O.; Lebrun, M. Export of a Toxoplasma gondii Rhoptry Neck Protein Complex at the Host Cell Membrane to Form the Moving Junction during Invasion. *PLoS Pathogens* **2009**, *5*, e1000309.
(52) Bannister, L. H.; Butcher, G. A.; Dennis, E. D.; Mitchell, G. H. Structure and invasive behaviour of Plasmodium *knowlesi* merozoites in vitro. *Parasitology* **2007**, 482–491.
(53) Proellocks, N. I.; Coppel, R. L.; Waller, K. L. Dissecting the apicomplexan rhoptry neck proteins. *Trends in Parasitology* **2010**, *26*, 297–304.
(54) Hodder, A. N.; Crewther, P. E.; Anders, R. F. Specificity of the protective antibody response to apical membrane antigen 1. *Infection and Immunity* **2001**, *69*, 3286-3294.
(55) Boyle, M. J.; Wilson, D. W.; Richards, J. S.; Riglar, D. T.; Tetteh, K. K. A.; Conway, D. J.; Ralph, S. A.; Baum, J.; Beeson, J. G. Isolation of viable *Plasmodium falciparum* merozoites to define erythrocyte invasion events and advance vaccine and drug development. Proceedings of the National Academy of Sciences of the United States of America **2010**, *107*, 14378–14383.
(56) Narum, D. L.; Thomas, A. W. Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites. *Molecular & Biochemical Parasitology* **1994**, *67*, 59–68.
(57) Hodder, A. N.; Crewther, P. E.; Matthew, M. L.; Reid, G. E.; Moritz, R. L.; Simpson, R. J.; Anders, R. F. The disulfide bond structure of Plasmodium apical membrane antigen-1. *Journal of Biological Chemistry* **1996**, *271*, 29446–29452.
(58) Remarque, E. J.; Faber, B. W.; Kocken, C. H. M.; Thomas, A. W. Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends in Parasitology* **2008**, *24*, 74–84.
(59) Pizarro, J. C.; Normand, B. V-L.; Chesne-Seck, M. -L.; Christine R. Collins, C. R.; Withers-Martinez, C.; Hackett, F.; Blackman, M. J.; Faber, B. W.; Remarque, E. J.; Kocken, C. H. M.; Thomas, A. W.; Bentley, G. A. Crystal structure of the malaria vaccine candidate apical membrane antigen 1. *Science* **2005**, *308*, 408-411.
(60) Bai, T.; Becker, M.; Gupta, A.; Strike, P.; Murphy, V. J.; Anders, R. F.; Batchelor, A. H. Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proceedings of the National Academy of Sciences* **2005**, *102*, 12736–12741.
(61) Coley, A. M.; Gupta, A.; Murphy, V. J.; Bai, T.; Kim, H.; Anders, R. F.; Foley, M.; Batchelor, A. H. Structure of the Malaria Antigen AMA1 in Complex with a Growth-Inhibitory Antibody. *PLoS Pathogens* **2007**, *3*, e138.
(62) Aikawa, M.; Miller, L. H.; Johnson, J.; Rabbege, J. Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *Journal of Cellular Biology* **1978**, *77*, 72–82.
(63) Baum, J.; Richard, D.; Healer, J.; Rug, M.; Krnajski, Z.; Gilberger, T. W.; Green, J. L.; Holder, A. A.; Cowman, A. F. A Conserved Molecular Motor Drives Cell Invasion and Gliding Motility across Malaria Life Cycle Stages and Other Apicomplexan Parasites. *Journal of Biological Chemistry* **2006**, *281*, 5197–5208.
(64) MacRaild, C. A.; Anders, R. F.; Foley, M.; Norton, R. S. Apical membrane antigen 1 as an anti-malarial drug target. *Current Topics in Medicinal Chemistry* **2011**, *11*, 2039–2047.
(65) Polley, S. D.; Chokejindachai, W.; Conway, D. J. Allele frequency-based analyses robustly map sequence sites under balancing selection in a malaria vaccine candidate antigen. *Genetics* **2003**, *165*, 555–561.
(66) Jiang, Y.; Liu, H.; Wang, X.; Li, G.; Qiu, Y.; Dou, X.; Wan, K. Conserved hypothetical protein Rv1977 in Mycobacterium tuberculosis strains contains sequence polymorphisms and might be involved in ongoing immune evasion. *International Journal of Clinical and Experimental Pathology* **2015**, *8*, 6891–6900.

(67) Coley, A. M.; Parisi, K.; Masciantonio, R.; Hoeck, J.; Casey, J. L.; Murphy, V. J.; Harris, K. S.; Batchelor, A. H.; Anders, R. F.; Foley, M. The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infection and Immunity* **2006**, *74*, 2628–2636.
(68) Nuttall, S. D.; Humberstone, K. S.; Krishnan, U. V.; Carmichael, J. A.; Doughty, L.; Hattarki, M.; Coley, A. M.; Casey, J. L.; Anders, R. F.; Foley, M.; Irving, R. A.; Hudson, P. J. Selection and affinity maturation of IgNAR variable domains targeting *Plasmodium falciparum* AMA1. *Proteins* **2004**, *55*, 187–197.
(69) Li, F.; Dluzewski, A.; Coley, A. M.; Thomas, A.; Tilley, L.; Anders, R. F.; Foley, M. Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *Journal of Biological Chemistry* **2002**, *277*, 50303–50310.
(70) Harris, K. S.; Casey, J. L.; Coley, A. M.; Masciantonio, R.; Sabo, J. K.; Keizer, D. W.; Lee, E. F.; McMahon, A.; Norton, R. S.; Anders, R. F.; Foley, M. Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infection and Immunity* **2005**, *73*, 6981–6989.
(71) Dutta, S.; Haynes, J. D.; Barbosa, A.; Ware, L. A. Mode of action of invasion-inhibitory antibodies directed against apical membrane antigen 1 of *Plasmodium falciparum*. *Infection and Immunity* **2005**, *73*, 2116-2122.
(72) Henderson, K. A.; Streltsov, V. A.; Coley, A. M.; Dolezal, O.; Hudson, P. J.; Batchelor, A. H.; Gupta, A.; Bai, T.; Murphy, V. J.; Anders, R. F.; Foley, M.; Nuttall, S. D. Structure of an IgNAR-AMA1 Complex: Targeting a Conserved Hydrophobic Cleft Broadens Malarial Strain Recognition. *Structure* **2007**, *15*, 1452–1466.
(73) Collins, C. R.; Withers-Martinez, C.; Bentley, G. A.; Batchelor, A. H.; Thomas, A. W.; Blackman, M. J. Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *Journal of Biological Chemistry* **2007**, *282*, 7431–7441.
(74) Vulliez-Le Normand, B.; Tonkin, M. L.; Lamarque, M. H.; Langer, S.; Hoos, S.; Roques, M.; Saul, F. A.; Faber, B. W.; Bentley, G. A.; Boulanger, M. J.; Lebrun, M. Structural and Functional Insights into the Malaria Parasite Moving Junction Complex. *PLoS Pathogens* **2012**, *8*, e1002755.
(75) Wang, G.; MacRaild, C. A.; Mohanty, B.; Mobli, M.; Cowieson, N. P.; Anders, R. F.; Simpson, J. S.; McGowan, S.; Norton, R. S.; Scanlon, M. J. Molecular Insights into the Interaction between *Plasmodium falciparum* Apical Membrane Antigen 1 and an Invasion-Inhibitory Peptide. *PLoS One* **2014**, *9*, e109674.
(76) Triglia, T.; Healer, J.; Caruana, S. R.; Hodder, A. N.; Anders, R. F.; Crabb, B. S.; Cowman, A. F. Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species. *Molecular Microbiology* **2000**, *38*, 706–718.
(77) Hehl, A. B.; Lekutis, C.; Grigg, M. E.; Bradley, P. J.; Dubremetz, J. F.; Ortega-Barria, E.; Boothroyd, J. C. Toxoplasma gondii homologue of plasmodium apical membrane antigen 1 is involved in invasion of host cells. *Infection and Immunity* **2000**, *68*, 7078–7086.
(78) Yap, A.; Azevedo, M. F.; Gilson, P. R.; Weiss, G. E.; O'Neill, M. T.; Wilson, D. W.; Crabb, B. S.; Cowman, A. F. Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cellular Microbiology* **2014**, *16*, 642–656.
(79) Sanchez, C. P.; Dave, A.; Stein, W. D.; Lanzer, M. Transporters as mediators of drug resistance in *Plasmodium falciparum*. *International Journal for Parasitology* **2010**, *40*, 1109–1118.
(80) Tyler, J. S.; Boothroyd, J. C. The C-Terminus of Toxoplasma RON2 Provides the Crucial Link between AMA1 and the Host-Associated Invasion Complex. *PLoS Pathogens* **2011**, *7*, e1001282.
(81) Harris, K. S.; Casey, J. L.; Coley, A. M.; Karas, J. A. Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning. *Journal of Biological Chemistry* **2009**, *284*, 9361-9371.

(82) Anders, R. F.; Adda, C. G.; Foley, M.; Norton, R. S. Recombinant protein vaccines against the asexual blood-stages of *Plasmodium falciparum*. *Human Vaccines* **2014**, *6*, 39–53.
(83) Yasgar, A.; Luci, D. K.; Beatty, W. L.; Hu, X.; Andersen, J.; Narum, D. L.; Moch, J. K.; Sun, H.; Haynes, J. D.; Maloney, D. J.; Jadhav, A.; Simeonov, A.; Miller, L. H.; Srinivasan, P. Disrupting malaria parasite AMA1-RON2 interaction with a small molecule prevents erythrocyte invasion. *Nature Communications* **2013**, *4*, 1–9.
(84) Erlanson, D. A. Introduction to fragment-based drug discovery. *Topics in Current Chemistry* **2012**, *317*, 1–32.
(85) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nature Reviews Drug Discovery* **2007**, *6*, 211–219.
(86) Bembenek, S. D.; Tounge, B. A.; Reynolds, C. H. Ligand efficiency and fragment-based drug discovery. *Drug Discovery Today* **2009**, *14*, 278–283.
(87) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nature Reviews Drug Discovery* **2016**, 1–15.
(88) Tsai, J.; Lee, J. T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N. K.; Sproesser, K.; Li, L.; Smalley, K. S. M.; Fong, D.; Zhu, Y.-L.; Marimuthu, A.; Nguyen, H.; Lam, B.; Liu, J.; Cheung, I.; Rice, J.; Suzuki, Y.; Luu, C.; Settachatgul, C.; Shellooe, R.; Cantwell, J.; Kim, S.-H.; Schlessinger, J.; Zhang, K. Y. J.; West, B. L.; Powell, B.; Habets, G.; Zhang, C.; Ibrahim, P. N.; Hirth, P.; Artis, D. R.; Herlyn, M.; Bollag, G. Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity. *Proceedings of the National Academy of Sciences* **2008**, *105*, 3041–3046.
(89) Gill, A. L.; Frederickson, M.; Cleasby, A.; Woodhead, S. J.; Carr, M. G.; Woodhead, A. J.; Walker, M. T.; Congreve, M. S.; Devine, L. A.; Tisi, D.; O'Reilly, M.; Seavers, L. C. A.; Davis, D. J.; Curry, J.; Anthony, R.; Padova, A.; Murray, C. W.; Carr, R. A. E.; Jhoti, H. Identification of Novel p38α MAP Kinase Inhibitors Using Fragment-Based Lead Generation. *Journal of Medicinal Chemistry* **2005**, *48*, 414–426.
(90) Doak, B. C.; Norton, R. S.; Scanlon, M. J. The ways and means of fragment-based drug design. *Pharmacology and Therapeutics* **2016**, *167*, 28-37.
(91) Murray, C. W.; Verdonk, M. L.; Rees, D. C. Experiences in fragment-based drug discovery. *Trends in Pharmacological Sciences* **2012**, *33*, 224–232.
(92) Jencks, W. P. On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences* **1981**, *78*, 4046–4050.
(93) Nakamura, C. E.; Abeles, R. H. Mode of interaction of beta-hydroxy-beta-methylglutaryl coenzyme A reductase with strong binding inhibitors: compactin and related compounds. *Biochemistry* **1985**, *24*, 1364–1376.
(94) Andrews, P. R.; Craik, D. J.; Martin, J. L. Functional group contributions to drug-receptor interactions. *Journal of Medicinal Chemistry* **1984**, *27*, 1648–1657.
(95) Green, N. M. Avidin. *Advances in Protein Chemistry* **1975**, 29, 85–133.
(96) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.
(97) Becker, E. D. A brief history of Nuclear Magnetic Resonance. *Analytical Chemistry* **1993**, 65, 295A-302A
(98) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-Based Drug Discovery. *Journal of Medicinal Chemistry* **2004**, *47*, 3463–3482.
(99) Patel, D.; Bauman, J. D.; Arnold, E. Advantages of crystallographic fragment screening: Functional and mechanistic insights from a powerful platform for efficient drug discovery. *Progress in Biophysics and Molecular Biology* **2014**, *116*, 92–100.
(100) Labbé, C. M.; Laconde, G.; Kuenemann, M. A.; Villoutreix, B. O.; Sperandio, O. iPPI-DB: a manually curated and interactive database of small non-peptide inhibitors of protein–protein interactions. *Drug Discovery Today* **2013**, *18*, 958–968.

(101) Lepre, C. A. 9 - Practical Aspects of NMR-Based Fragment Screening. *Fragment Based Drug Design Tools, Practical Approaches, and Examples* **2011**, *493*, 219–239.
(102) Hubbard, R. E.; Murray, J. B. 20 - Experiences in Fragment-Based Lead Discovery. *Fragment Based Drug Design Tools, Practical Approaches, and Examples* **2011**, *493*, 509–531.
(103) Leverson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; Huang, D. C. S.; Hymowitz, S. G.; Jin, S.; Khaw, S. L.; Kovar, P. J.; Lam, L. T.; Lee, J.; Maecker, H. L.; Marsh, K. C.; Mason, K. D.; Mitten, M. J.; Nimmer, P. M.; Oleksijew, A.; Park, C. H.; Park, C.-M.; Phillips, D. C.; Roberts, A. W.; Sampath, D.; Seymour, J. F.; Smith, M. L.; Sullivan, G. M.; Tahir, S. K.; Tse, C.; Wendt, M. D.; Xiao, Y.; Xue, J. C.; Zhang, H.; Humerickhouse, R. A.; Rosenberg, S. H.; Elmore, S. W.; Souers, A. J. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nature Medicine* **2013**, *19*, 202–208.
(104) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **2005**, *435*, 677–681.
(105) Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. A "Rule of Three" for fragment-based lead discovery?. *Drug Discovery Today* **2003**, *8*, 876–877.
(106) Lipinski, C. A.; Lombardo, F.; Dominy, B. W. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **2012**, *23*, 3-25.
(107) Keserü, G. M.; Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews Drug Discovery* **2009**, *8*, 203–212.
(108) Reymond, J. -L.; van Deursen, R.; Blum, L. C.; Ruddigkeit, L. Chemical space as a source for new drugs. *MedChemComm* **2010**, *1*, 30.
(109) Fink, T.; Bruggesser, H.; Reymond, J.-L. Virtual Exploration of the Small-Molecule Chemical Universe below 160 Daltons. *Angewandte Chemie International Edition English* **2005**, *44*, 1504–1508.
(110) Macarron, R. Critical review of the role of HTS in drug discovery. *Drug Discovery Today* **2006**, *11*, 277–279.
(111) Ichihara, O.; Barker, J.; Law, R. J.; Whittaker, M. Compound Design by Fragment-Linking. *Molecular Informatics* **2011**, *30*, 298–306.
(112) Chung, S.; Parker, J. B.; Bianchet, M.; Amzel, L. M.; Stivers, J. T. Impact of linker strain and flexibility in the design of a fragment-based inhibitor. *Nature Chemical Biology* **2009**, *5*, 407–413.
(113) Röhrig, C. H.; Loch, C.; Guan, J.-Y.; Siegal, G.; Overhand, M. Fragment-Based Synthesis and SAR of Modified FKBP Ligands: Influence of Different Linking on Binding Affinity. *ChemMedChem* **2007**, *2*, 1054–1070.
(114) Ward, R. A.; Brassington, C.; Breeze, A. L.; Caputo, A.; Critchlow, S.; Davies, G.; Goodwin, L.; Hassall, G.; Greenwood, R.; Holdgate, G. A.; Mrosek, M.; Norman, R. A.; Pearson, S.; Tart, J.; Tucker, J. A.; Vogtherr, M.; Whittaker, D.; Wingfield, J.; Winter, J.; Hudson, K. Design and Synthesis of Novel Lactate Dehydrogenase A Inhibitors by Fragment-Based Lead Generation. *Journal of Medicinal Chemistry* **2012**, *55*, 3285–3306.
(115) Kar, S.; Kar, S. Control of malaria. *Nature Reviews Drug Discovery* **2010**, *9*, 511–512.
(116) Lim, S. S.; Debono, C. O.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening. *Australian Journal of Chemistry* **2013**, *66*, 1530–1536.

(117) Lim, S. S.; Yang, W.; Krishnarjuna, B.; Kannan Sivaraman, K.; Chandrashekaran, I. R.; Kass, I.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Anders, R. F.; Scanlon, M. J.; Scammells, P. J.; Norton, R. S.; McGowan, S. Structure and Dynamics of Apical Membrane Antigen 1 from *Plasmodium falciparum* FVO. *Biochemistry* **2014**, *53*, 7310–7320.
(118) Meyer, B.; Peters, T. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angewandte Chemie International Edition English* **2003**, *42*, 864–890.
(119) Viegas, A.; Manso, J.; Nobrega, F. L.; Cabrita, E. J. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *Journal of Chemical Education* **2011**, *88*, 990–994.
(120) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430–431.
(121) Verdonk, M. L.; Rees, D. C. Group Efficiency: A Guideline for Hits-to-Leads Chemistry. *ChemMedChem* **2008**, *3*, 1179–1180.

Chapter 2. Development of 2-aminothiazoles

2.1 Background

The aminothiazole core is considered an active pharmacophore that has been widely used in medicinal chemistry for drug discovery research. The aminothiazole core is made up of an amino group linked to a five membered heterocyclic unit containing sulfur and nitrogen located at the 1- and 3-positions, respectively. Over the years, the aminothiazole scaffold has been explored as a viable pharmacophore in many important therapeutic settings, exhibiting antimicrobial,[1] antiviral,[1] anticancer,[2] antitumor[3] and anticonvulsant[4] activity. Over the last decade, numerous drugs containing a 2-aminothiazole core have successfully been developed as clinically relevant pharmaceuticals. Famotidine is an example of an 2-aminothiazole containing drug clinically used for the treatment of peptic ulcers and to help control gastroesophageal reflux.[5] Abafungin is another example of an aminothiazole-containing drug used as an antimicrobial agent for the treatment of dermatomycoses.[6] Recently the aminothiazole core has also been able to show antiplasmodial activity.[7,8] An analysis of the 46 positive hits identified from the MIPS library revealed three fragments which bound AMA1 displaying a high LE of $\geq$ 0.3 kcal mol^{-1} heavy $atom^{-1}$; these three fragments were identified as belonging to the 2-aminothiazole chemical class.[9] As described by Hajduk,[10] fragments with a high LE (LE $\geq$ 0.3 kcal mol^{-1} heavy $atom^{-1}$) are regarded as efficient binders and have a higher probability of maintaining efficient binding throughout an elaboration campaign. Thereby, the high LE displayed by these three 2-aminothiazole-derived fragments (**Figure 2.1**, compounds **1**, **2** and **3**) provided an attractive starting point for the development of a preliminary structure-activity relationship (SAR) profile based on the 2-aminothiazole scaffold. The thiazole scaffold was identified as the largest chemical class, making up 15% of the total positive hits after SPR follow-up. Of the nine thiazole-based hits identified from the MIPS library (**Figure 2.1**), two different aminothiazole sub-types, recognized as the 4-aryl substituted series (demonstrated by compounds **1, 2** and **3**) and the 4-methyl-5-aryl substituted series (demonstrated by compounds **4** and **5**), were considered ideal starting points for this study targeting AMA1.

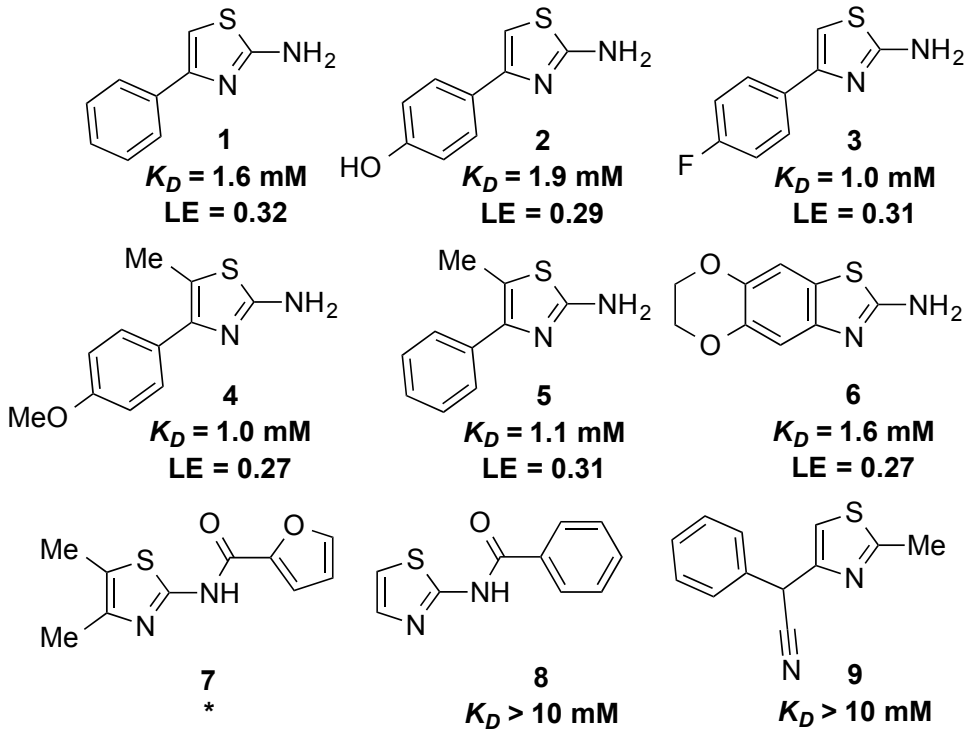


Figure 2.1 2-Aminothiazole fragment hits identified through fragment-based screening. *The binding affinity of **7** could not be determined resulting in an SPR response of zero for each tested sample.[9]

The six 2-aminothiazole-containing hits (**1**-**6**) displayed binding affinities between 1.0 and 1.9 mM in comparison to the 2-amidothiazole compounds (**7** and **8**), which were considered non-binders. The replacement of the amino group with a methyl group at the 2-position of the thiazole core resulted in **9,** which was also considered a non-binder. This suggested the importance of the 2-amino moiety with regard to maintaining affinity for AMA1. Unfortunately attempts to crystallize 2-aminothiazole fragments within AMA1 had failed to yield any success and instead competition STD and CPMG NMR would be the most suitable assay techniques for determining fragment binding within the hydrophobic cleft, which is a significant binding pocket of interest in this body of work (**Figure 2.2**). The inability to crystallized hits with AMA1 when compounds were either co-crystallized or soaked with 3D7 *Pf*AMA1 could be the consequence of weak affinity, which is typically encountered with fragment compounds with a clear lack of specific binding. Since structural information revealing the binding mode of 2-aminothiazole analogues was insufficient at this stage of our work, a rational structure-based drug discovery (SBDD) approach for the development of chemical leads could not be instigated. Instead STD and CPMG NMR were utilized in order to identify compounds competed off by the highly specific binding peptide, R1. It was

hypothesized that any positive hits via STD and CPMG NMR could be further rationally improved through chemical elaboration with the overall aim being the improvement in affinity for resultant hits and subsequently increasing the likelihood of obtaining a crystal complex of said analogues bound with AMA1. This work explores the chemical development of the 2-aminothiazole series via a 'SAR by NMR' approach in an attempt to further characterize this chemical class with the aim of understanding how these compounds bind AMA1.[11]

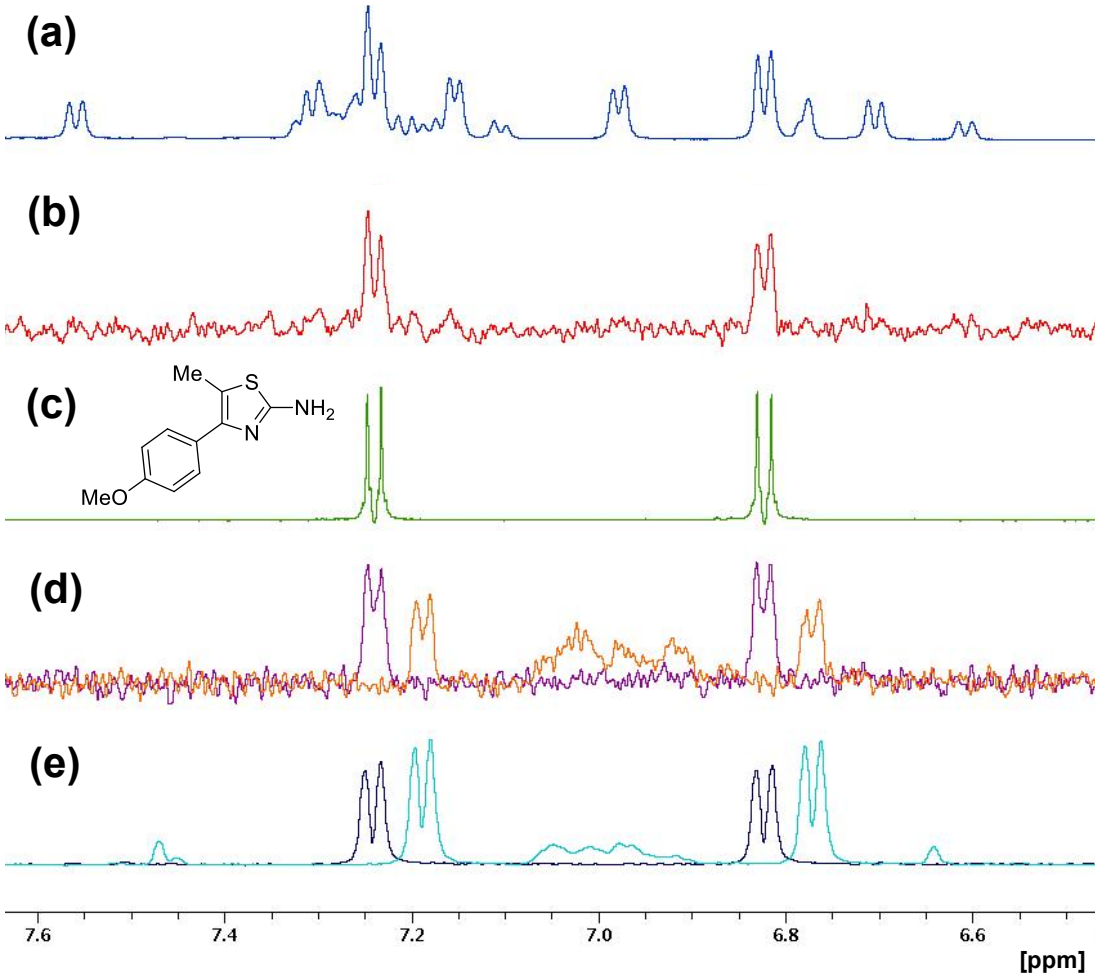


Figure 2.2 Example of saturation transfer difference (STD) results from our library screen.[9] (a) ^{1}H NMR spectrum for cocktail of six fragments. (b) STD spectrum showing that one fragment out of the cocktail was identified as a hit, with STD signal intensity of 1.3 % (STD signal intensity is the percentage of proton signal reduction relative to the off-resonance spectrum). (c) ^{1}H NMR reference spectrum for the fragment hit of compound **4**. (d) STD competition NMR experiments using R1. STD spectra were acquired in the absence (purple) and presence (orange) of R1. (e) CPMG spectra for the fragment hit in the presence of AMA1 and in the absence (grey) and presence (cyan) of R1. In both (d) and (e) spectra in the presence of R1 are offset by 0.07 ppm for clarity. Proton signals between 6.9 and 7.1 ppm are resonances of the R1 peptide.

2.2 Elaboration of the 4-aryl substituted aminothiazole series

The 4-aryl substituted aminothiazole scaffold was the simplest thiazole sub-type recognized from the screen (**Figure 2.1**). A thorough SAR probe of this sub-type was initiated through chemical manipulation. The lack of chemical diversity of thiazole containing compounds within the MIPS library led to an initial synthetic approach focusing on the optimization of the aryl ring at the 4-position; this position was considered an ideal starting point to efficiently introducing chemical diversity without altering the core structure. Investigation of each arene position of the phenyl ring (i.e. *o*-, *m*- and *p*-position) was investigated using numerous functional groups with varying physicochemical properties such as size and polarity to probe the yet-to-be characterized chemical space surrounding the scaffold when bound to AMA1 (**Figure 2.3**). Expansion of the 4-aryl substituted series would aide the determination of the contribution of various functional groups on the ring structure in order to maximize the binding potential of analogues in the search for higher affinity strain-transcending compounds. We also were interested in maintaining ideal LE whilst taking into consideration the hydrophobicity, steric descriptors such as steric effects, and electronic descriptors such as resonance when developing this series. Due to the lack of structural feedback, synthesis of differentially functionalized analogues was proposed to uncover favorable substituents on the 2-aminothiazole core, which enhance the scaffold's ability to form new interactions with the backbone of AMA1.

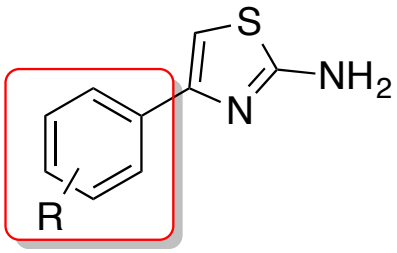
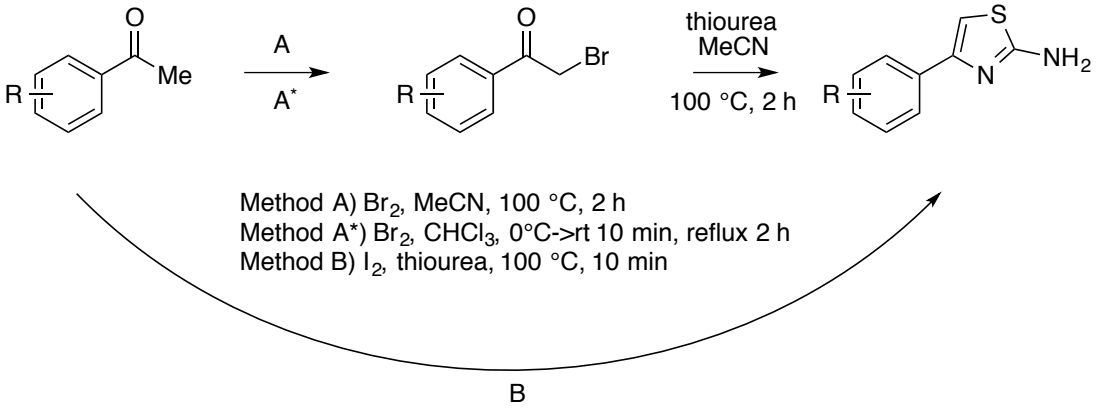


Figure 2.3 Proposed initial investigation of the 2-aminothiazole core. Our focus was the optimization of the 4-aryl group (circled in red) to establish preliminary SAR.

Synthesis of the 4-aryl substituted series utilized the Hantzsch thiazole synthesis, reacting thiourea with substituted haloketones to form the thiazole ring.[12] For the purpose of this study, the initial halogenation of substituted acetophenones to form the α-haloketone required optimization owing to the variable reactivity of substituents present on the aryl ring. The halogenation of selected acetophenones was conducted with reference to literature procedures;[13,14] the conditions employed for halogenation reaction was determined by the type of substituent attached to the aryl ring. A general procedure was initially adopted using bromine (Br_2) or *N*-bromosuccinimide (NBS) heated

in MeCN at 100 °C for 2 h (**Scheme 1, method A**). The bromination reaction was the first step of a two-step synthetic procedure that was followed by the resultant cyclization step. Following the bromination, the crude material was analyzed via ^{1}H NMR and LCMS to confirm whether the formation of the brominated intermediate was successful. The brominated intermediate was subsequently dissolved in MeCN and thiourea to the reaction to be heated to 100 °C for 2 h to allow the S-alkylation and resultant condensation between the remaining free amine and carbonyl to occur. The ability of activating substituents on the phenyl ring (i.e. amino (-NH_2) or hydroxy (-OH)) to form resonance structures in solution meant excessive bromination was plausible at multiple regions of the acetophenone reagent. This is a result of the re-arrangement of electron density around the ring system; method A required optimization to ensure selective reactivity at the α-carbon of the acetophenone was maintained. For compounds **19**-**21** and **24**, each α-haloketone was initially prepared using an adapted version of method A (**Scheme 1, method A***), based upon a literature procedure.[15] This procedure was completed in chloroform ($CHCl_3$) and required the drop-wise addition of Br_2 to a mixture at 0 °C. The drop-wise addition of Br_2 at 0 °C was employed to control the stoichiometry of the reaction by keeping the concentration of available bromine low to avoid any subsequent over bromination. The use of less polar solvent in $CHCl_3$ was shown to minimize the unwanted excessive halogenation. These two methods successfully achieved the synthesis of the desired brominated α-haloketone, which was a valuable intermediate for further elaboration.

A more efficient synthesis of 4-aryl substituted aminothiazoles was adapted from the literature by Cáceres-Castillo *et al.*[14] This procedure implemented the use of a neat solution of substituted acetophenone, iodine (I_2) and thiourea, heated for 2 h at 100 °C to minimize the amount of un-wanted byproduct (**Scheme 1, method B**). This procedure was attempted to improve the yield of the desired α-haloketone particularly for those acetophenones with electron-donating groups, which had limited success when synthesized by method A*. By comparing the yields of both method A and A* to the newly adopted method B, the one-pot synthesis of 4-aryl substituted 2-aminothiazole analogues resulted in improved yields and easier isolation of the desired product. Method B was considered the preferred method due to the improved reaction outcome with the success attributed to the simplified work-up required to afford the desired material owing to the absence of solvent involved. Albeit the yield for each method described remained low yielding (below 50%), the reaction provided enough material for biological assessment via SPR and NMR testing stage of the study.



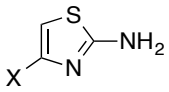
	Method	Final Yield (%)		Method	Final Yield (%)
1 R=H	-	78	**19** R=*o*-OMe	A*	25
10 R=*m*-F	A	16	**20** R=*p*-OMe	A*	13
11 R=*o*-F	A	11	**21** R=*m*-OMe	A*	14
12 R=*p*-Me	A	14	**22** R=*p*-NH_2	B	21
13 R=*o*-Me	A	30	**23** R=*o*-OH	B	23
14 R=*m*-Me	A	22	**24** R=*m*-NH_2	A*	5
15 R=*m*-NO_2	A	34			
16 R=*p*-CF_3	A	15			
17 R=*o*-CF_3	A	20			
18 R=*m*-CF_3	A	23			

Scheme 1 Synthesis of 4-aryl substituted aminothiazole analogues. A* represents an adapted method of A.[15] The brominated reagent was available for the synthesis of **1**.

Screening of all new analogues was performed by SPR against 3D7 *Pf*AMA1$_{[104-442]}$ immobilized on a CM5 chip using an amide-coupling technique. The known AMA1 binder R1 was included for all SPR experiments as a positive control, tested with analogues to confirm whether the protein surface remained active throughout experiments. Each of the analogues synthesized in this body of work were screened at 12.5, 25, 50, 100 and 200 μM with 30 s contact times forming a dose-response curve to determine compound affinity. The binding affinities of fragment hits were estimated based on the response (RU) at the five different concentrations of compound with a fixed fragment R_{max} (maximal binding capacity of AMA1 protein surface). In an effort to rank the elaborated analogues, sensorgram responses were normalized using the scheme described by Giannetti *et al.*[16] The normalized maximal binding response (R_{max}), predicted from the average of the kinetic fit of the control (R1 peptide) sensorgram, was applied to response curves obtained with fragments at non-saturating concentrations. To determine K_D values, the Langmuir adsorption isotherm (Langmuir adsorption model explains adsorption by assuming an atom, ion or molecule behaves as an ideal gas at a constant temperature) was then fitted to the normalized compound

binding data. Unlike the original MIPS fragment screen, we initially opted to analyze all new aminothiazole analogues via SPR; by performing SPR first we were able to more efficiently screen a larger number of compounds whilst using a minimal amount of the highly valuable AMA1 sample. Following SPR analysis, all analogues shown to display strong affinity (K_D > 1 mM) via SPR were subsequently screened using the more protein-intensive STD and CPMG competition NMR experiments. The X-ray crystallography of the enhanced binders was also attempted with the aide of co-workers at Monash University.

In total, eighteen 4-aryl substituted aminothiazole analogues made up our preliminary SAR study against AMA1 (**Table 2.4**). When tested by SPR the strongly electron-withdrawing substituents on the 4-aryl group, like the trifluoromethyl group, were least tolerated, with the *m*-CF_3 analogue **18** demonstrating slightly better binding affinity of 2.6 mM, over both *p*-CF_3 and *o*-CF_3 derivatives, **16** (5.7 mM) and **17** (3.2 mM). When compared to the parent fragment **1** (1.6 mM) identified from the MIPS library, interestingly, the replacement of a hydrogen atom with a fluorine atom at each position around the aryl ring afforded a range of affinities including 1.0 mM **(3**, *p*-fluoro**),** 1.5 mM (**10,** *m*-fluoro) and 5.2 mM (**11**, *o*-fluoro) analogue; a fluorine atom is often used as an isostere of hydrogen atom given its similarity in volume with the fluorine atom (1.47 Å and 1.20 Å van der Waals radius respectively).[17] The difference in affinities could be associated with atom size as observed when comparing **3** and **11** to **1**, suggesting a tighter fit for **3** and a restricted fit for **11** possibly due to the limited chemical space available to occupy this position. In the case of the methyl- and methoxy-substituted analogues (**12**-**14** and **19**-**21**, respectively), each substituent elaborated at the *meta*-position demonstrated the strongest affinity of all the arene positions investigated, with binding affinities of 660 μM for **14** and 1.2 mM for **21** respectively. Considering the hydrophobic nature of the cleft, it was conceivable for a non-polar functional group, such as the *m*-methyl analogue **14**, to display improved affinity when compared to the parent compound **1**. The affinity for the methyl-group was also characteristic of the steric bulk related to a comparison of the STERIMOL length parameter *L* for each substituent.[18] This parameter describes the length from the point of attachment (P) to the projected van der Waals (vdW) surface, which calculated the methyl-group (L = 2.87) to be the smaller group when compared to the methoxy-group (L = 3.98). Since the poorer affinity for the larger methoxy-group suggested a more hindered fit, this result alluded to limited chemical space available for probing at certain positions around the aryl ring.



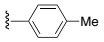
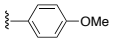
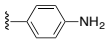
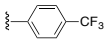
	X	K_D (mM)	LE		X	K_D (mM)	LE
1		1.0	0.32	**17**		3.2	0.21
10		1.5	0.30	**18**		2.6	0.22
11		5.2	0.24	**19**		1.7	0.27
12		1.6	0.29	**20**		2.6	0.25
13		1.7	0.29	**21**		1.2	0.28
14		0.66	0.33	**22**		0.25	0.37
15*		-	-	**23**		0.35	0.36
16		5.7	0.19	**24**		0.94	0.32

Table 2.4 Synthesized 2-aminothiazole series with binding affinities (K_D) determined by surface plasmon resonance (SPR) and ligand efficiencies (LE) indicated for each analogue. *The affinity of **15** could not be determined owing to solubility issues in running buffer.

The strong electron-donating effects of the amino- and hydroxy-groups were also able to display strong affinity as observed for **22**, **23** and **24** (**Table 2.4**). As hypothesized, the electron-donating groups displayed the strongest affinity due to their capability to form strong intermolecular forces, such as hydrogen bonding within the binding domain. As initially proposed, the compounds that displayed strong affinity, such as the *m*-methyl analogue **14**, were followed up by competition NMR; each analogue tested was competed off by R1, suggesting a similar interaction to that of the original aminothiazole library fragments. Nevertheless the inability to crystallize any analogues of interest during crystallography studies meant no key interactions formed by specific functional groups or structural features (i.e. size, shape and charge) could be characterized and as a result the SAR was considered largely flat around the 4-aryl ring. In addition to the MIPS fragment library

screen described by Lim *et al.*[9], the preliminary results of the 4-aryl substituted aminothiazole SAR study were also published describing the initial work conducted for the FBDD investigation of AMA1 (Refer to Appendix A, section A.1).

Although slight variations in affinity were identified for the series tested, the SPR behaviour exhibited by the 4-aryl substituted aminothiazoles raised a number of concerns regarding overall reliability of the data obtained. Upon analysis of the eighteen 4-aryl analogues, it was evident that no compound reached saturation over the concentration range initially investigated. This behaviour was considered non-optimal, leading to efforts to reach saturation to verify the true binding affinity of the 4-aryl aminothiazoles, however when compounds were made up to higher concentrations (> 250 μM) the solubility of these compounds was affected resulting in precipitation. When testing fragments (M_r <300) it is widely acknowledged that non-specific binding may result from flowing small fragments with low affinity over a protein target, as witnessed with AMA1. Concentration-dependent aggregation was another plausible theory when attempting to explain the SPR data observed for the series. A study by Giannetti *et al.*[16] described this as certain compounds behaving in a 1:1 manner when tested at lower concentrations but subsequently behaving in a manner beyond a 1:1 ratio when tested at higher concentrations. Giannetti[16] identified concentration-dependent aggregation to be fairly common for performed experiments, which was the case upon review of the SPR data for the 4-aryl series revealing clear evidence of concentration-dependent aggregation for the 18 compounds tested. All compounds tested were able to display fast on/off kinetics, as was anticipated for small molecules, but the continual presence of possible non-stoichiometric binding or aggregation when higher concentrations were tested made it difficult to accurately calculate the affinity of each compound (**Figure 2.5**). It was hypothesized that the elaborated series of analogues would be consistently around the millimolar (mM) affinity range but this hypothesis could not be supported since the compound binding never reached saturation within the limited concentration range permissible for testing.

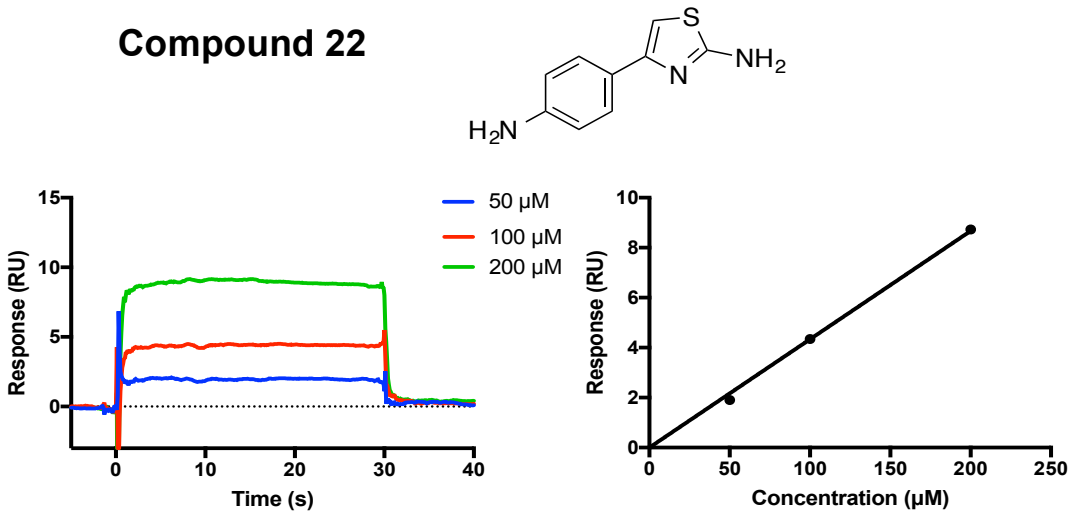


Figure 2.5 An SPR sensorgram of compound **22**. Owing to solubility issues, the highest concentration tested was limited to 200 µM. Saturation was never achieved within this concentration range, hence a straight-line fit resulted when fitted against a 1:1 binding model.

2.3 Development of the 4-Methyl-5-aryl substituted 2-aminothiazole series

The failure to elucidate any reliable SAR from the 4-aryl series led to a new SAR study based on compounds **4** and **5** (**Figure 2.1**). These compounds were representative compounds of the second largest aminothiazole sub-type recognized from the fragment hits. The presence of a methyl group at the 5' position on the thiazole core was hypothesized to increase lipophilicity, i.e. partition coefficient *P* (LogP), but water solubility improved due to a decrease of the crystal lattice energy. Thus, analogues were shown to possess a lower average LogP value for both **4** and **5** when compared to **1**. Compound water solubility was a consistent issue encountered during the screening campaign of the 4-aryl substituted series, therefore it became imperative future scaffolds of interest would remain in solution during testing. Only 2 additional 4-aryl-5-methyl aminothiazole analogues (*p*-methyl and *p*-trifluoromethyl) were synthesized utilizing method A of Scheme 1 to halogenate substituted propiophenones prior to cyclization to form the 4-aryl-5-methyl aminothiazole core (**Figure 2.6**). However early indications, based upon subsequent SPR testing, suggested no dramatic improvement in solubility and affinity would result for this series as indicated by insolubility displayed by **25** and the poor affinity and reduced LE of **26**, remaining a millimolar binder. This series was eventually abandoned.

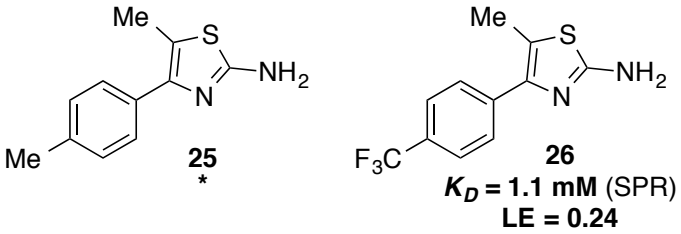
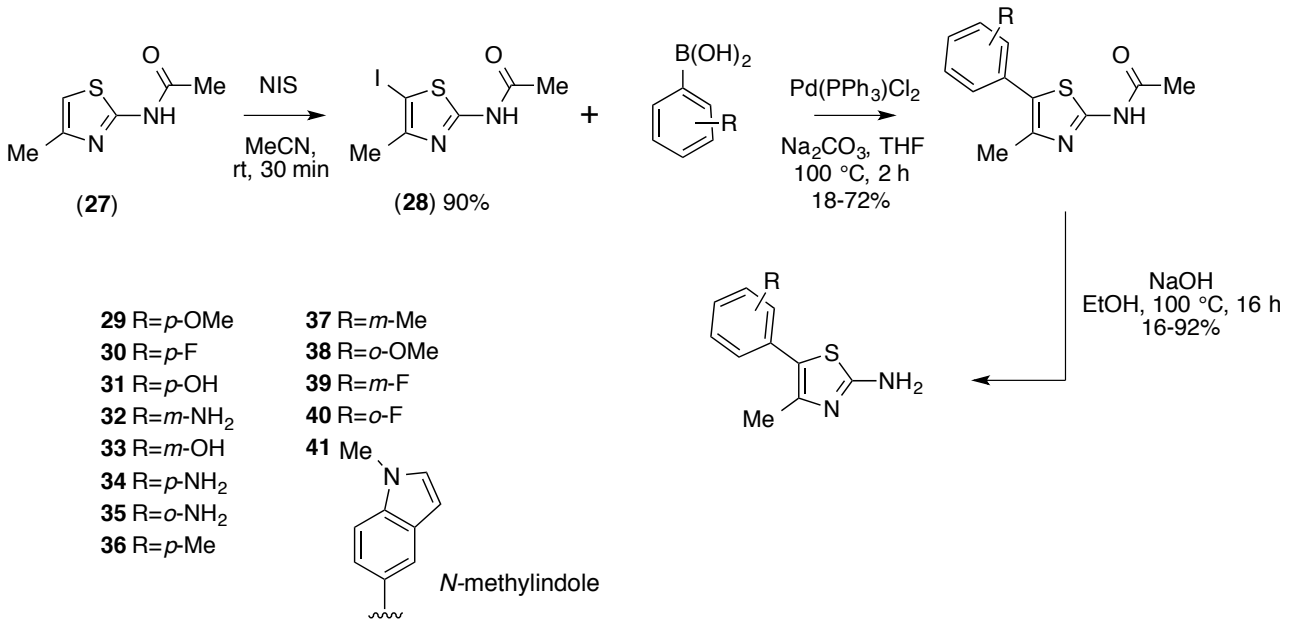


Figure 2.6 Synthesized 4-aryl-5-methyl aminothiazole analogues. *****The binding affinity of **25** could not be determined resulting in an SPR response of zero for each tested sample.

In parallel, the spatial arrangement of the methyl group was also investigated in which a 'flipped' version of compounds similar to **4** and **5** were developed to determine whether binding would be affected. 4-Methyl-2-aminothiazole was commercially sourced to provide rapid SPR testing data, this compound subsequently assisted in the identification of the necessity of the substituted aryl substituent as critical for binding to AMA1. 4-Methyl-2-aminothiazole's inability to bind AMA1 confirmed this finding. Given the lack of structural information surrounding 2-aminothiazole binding, we were interested in exploring whether the elaboration of substituents surrounding the aryl substituent of the 4-methyl-5-aryl series would lead to compounds demonstrating higher affinity for AMA1. We were also interested in determining whether we could generate some SAR based on this new scaffold through the introduction of substituents at each of the arene positions (i.e. *o*-, *m*- and *p*-position). We incorporated both polar and non-polar substituents along with probing substituents with varying steric and electronegativity effects to grasp an improved understanding of the chemical space demonstrated to interact with the hydrophobic cleft of AMA1 (**Scheme 2**).



Scheme 2 Synthesis of the 4-methyl-5-aryl-2-aminothiazoles.[19]

The synthetic strategy for the synthesis of the new 4-methyl-aryl-2-aminothiazoles series would incorporate Suzuki coupling conditions and as a result these conditions would require the protection of the amino starting material. To avoid any cross reactivity, the amino starting material was purchased from commercial sources as the acetyl-protected reagent **27**. The acetyl group was believed to be the most favourable protecting group owing to the mildly basic conditions required to successfully deprotect the amine and furnish the free amino group for further elaboration (**Scheme 2**). The key iodo-containing intermediate **28** was synthesised in excellent yield using *N*-iodosuccinimide stirred in MeCN at rt to install the iodo-group at the 5-position giving *N*-(5-iodo-4-methylthiazol-2-yl)acetamide **28** in an isolated yield of 90%. This was, confirmed by the loss of the proton signal at ^{1}H NMR 6.54 ppm reminiscent of the replacement of an aromatic proton with another atom in this case an iodo group. The introduction of an iodo group on the aromatic ring provides a substrate, which can undergo cross-coupling reactions in the presence of palladium catalyst. Other functional groups, which are also suitable for this step, include bromo and triflate-substituted aromatics. Iodination of **27** was conducted since palladium insertion into the carbon halide or pseudohalide bond tends to be more efficient when the halide is an iodo.

Suzuki reactions were adapted from a literature procedure[19] utilising the palladium catalyst bis(triphenylphosphine)palladium(II) dichloride, aqueous Na_2CO_3 (1 M) in THF. After the addition of all reagents the solution was degassed with nitrogen and subsequently heated at 100 °C for 2 h. Deoxygenation of the resultant solution was imperative for the oxidative addition of palladium

(Pd(0) into the carbon-halogen bond becoming Pd(II)) in the process whilst also reducing the oxidation of the catalyst to the inactive triphenylphosphine oxide. Following the formation of the organopalladium complex, transmetalation ensued with the boron-ate complex (produced from boronic acid reacting with base) forming an intermediate species and subsequent reductive elimination of the organopalladium complex results in the formation of the new carbon-carbon bond and regenerates palladium (Pd(0), thus completing the catalytic cycle. All coupled acetyl products were subsequently de-protected via base hydrolysis; the use of 5M NaOH solution at 100 °C for 2-3 h gave the best result. The final yields were observed to range between 16-77%; resultant yields were impacted namely by substituent(s) on the aryl ring. Novel compounds generated in this series were characterised using ^{1}H NMR, ^{13}C NMR and were required ^{19}F NMR.

The 4-methyl-5-aryl aminothiazole series was screened via SPR against 3D7 *Pf*AMA1$_{[104-442]}$ immobilized on a CM5 chip using an amide-coupling methodology. Each elaborated analogue was initially screened via SPR as previously described in examination of compound affinity (Section 2.2). In total, 13 analogues were screened via SPR with the main emphasis being the generation of SAR around the 5-aryl ring. Unfortunately the SAR did not demonstrate any clear indication of the attributes leading to improved affinity with the emergence of non-stoichiometric binding which was commonly observed for most derivatives except 2 analogues, which did not demonstrate binding to AMA1. For the 11 analogues that bound to the protein surface, we rationalized this observation as a result of poor compound solubility in buffer and given all analogues belonging to this series precipitated out of solution at concentrations above 300 μM, a rational decision was made not to test analogues beyond 200 μM. As a result data collected for the 4-methyl-5-aryl series was problematic to interpret since neither analogue displayed saturable binding within the limited concentration range tested. Of the 11 analogues that bound the protein surface 10 displayed a similar shape to that of compound **33** with the exception of compound **34**, which consistently displayed slow dissociation kinetics (a slow on/off rate) (**Figure 2.7**).

When analysing the response levels of each analogue screened, analogue interaction with the sensor surface was ameliorated as each analogue was reference subtracted removing all possible response that could be attributed to the running of buffer (+DMSO). The reduced solubility of analogues above 200 μM, inability of each analogue to reach saturation, coupled with the inability to crystallize analogues of interest with AMA1 suggested that all analogues, which bound to AMA1 remained millimolar (mM) binders of AMA1. In relation to the sensorgram of **34**, the presence of slow dissociation kinetics is typically not associated with fragments that display low affinity for a target.[20] In practice, a slow dissociation rate is typically attributed to kinetic behaviour of larger

molecules (M_r > 300) that possess a strong affinity for a protein target or region; R1 is one example known to display slow dissociation kinetics displaying a K_D of 100 nM for 3D7 AMA1. The existence of slow dissociation kinetics for a fragment is commonly attributed to aggregation, however this is difficult to identify given the same compound was identified to behaved when tested at lower concentrations.[16] This observation was true for compound **34,** since **34** seemingly displayed stoichiometric binding when tested at lower concentrations but bound in a non-stoichiometric manner when tested at higher concentrations (>150 μM). Such behaviour is commonly described as that of a concentration-dependent aggregator. Consequently the non-optimal behaviour displayed by each of the 11 bound analogues was characteristic of small molecules binding a large protein surface in which multiple binding sites may exist. Although many analogues initially demonstrated the ability to bind at sub-saturating levels when tested at lower concentrations, the inability to accurately determine affinity without reaching a saturation point meant the SAR would remain undefined around the 5-aryl group.

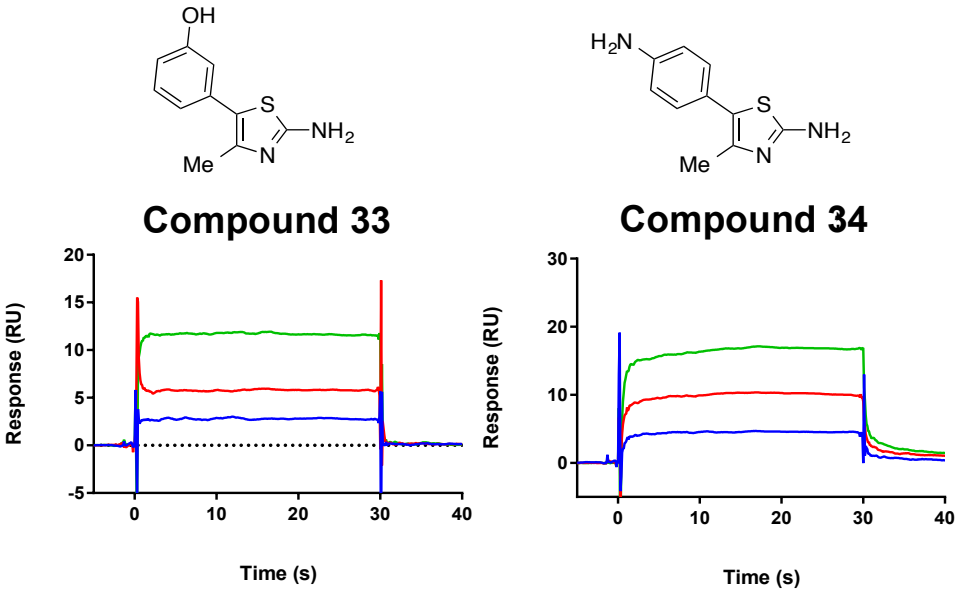
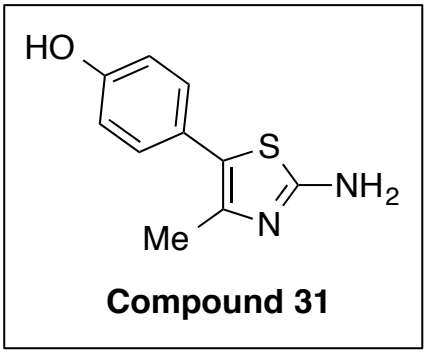


Figure 2.7 Classes of non-ideal binding behaviour for analogues of the 4-methyl-5-aryl aminothiazole series.

Although the SAR by SPR remained flat for this chemical class, analogues of the 4-methyl-5-aryl series had still been shown to interact in or around the hydrophobic cleft of AMA1 via R1 competition NMR. Following these studies, attempts to map the possible binding site(s) of newly synthesized 4-methyl-5-aryl analogues were accomplished using 2D [^{1}H-^{15}N]-transverse relaxation

optimized spectroscopy (TROSY) NMR. This work was published including efforts to assign the backbone residues of both 3D7 and FVO AMA1 (Refer to Appendix A, section A.3). TROSY NMR was designed to help identify residues which interact with the molecule bound to the protein target; FVO AMA1 was used for this purpose given the assignment of the backbone residues was more complete when compared to 3D7 AMA1.[21] A small selection of 4-methyl-5-aryl analogues were investigated as part of this study, utilizing analogues which displayed the most consistent response levels when screened via SPR such as compound **31**, 4-(2-amino-4-methylthiazol-5-yl)phenol (**Figure 2.8**). Similarly to SPR results, the NMR assay was able to identify **31** as a non stoichiometric binder due to the presence of multiple perturbation sites, including more than one possible site along the length of the hydrophobic cleft (**Figure 2.8**). Although **31** was still regarded as a fragment in relation to molecular mass (M_r < 300), the lack of the improvement in affinity coupled with no saturable binding identified (no change in CSP values across all sites) during TROSY NMR testing meant these compounds were deemed to provide no opportunity for further development towards high-affinity binders. No further studies were conducted on this series other than compound **34,** which was included in another study focusing on the use of ^{19}F NMR to detect the ligand-induced conformational changes of AMA1 (Refer to Appendix A, section A.2).[22]

A



Compound 31

B

Binding Regions	Compound **35**
Site 1	T171 (0.01) L176 (0.02) K177 (0.02) D178 (0.01) G179 (0.01) G180 (0.02) F183 (0.01)
Site 2	S191 (0.01) Q255 (0.01)
others	L307 (0.01) E308 (0.01) N309 (0.01) A310 (0.01) I363 (0.01) N241 (0.01)

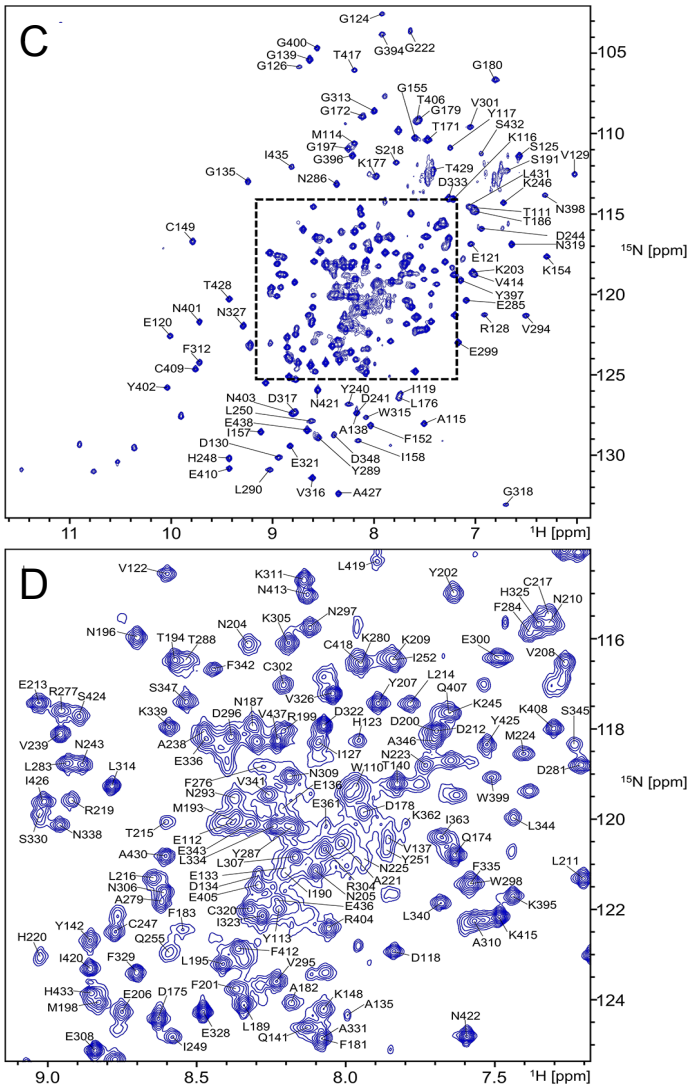


Figure 2.8 Example of NMR study to identify binding site on FVO AMA1.[21] A) The structure of the 2-aminothiazole analogue, 31. B) Specific binding sites of 31 determined using 2D [^{1}H-^{15}N]-TROSY perturbations. Site 1 and 2 correspond to the conserved and polymorphic ends, respectively, of the FVO *Pf*AMA1 hydrophobic cleft. The CSP values (ppm) are shown in brackets (polymorphic residue is highlighted red). Two-dimensional [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectrum of FVO *Plasmodium falciparum* apical membrane antigen 1 (DI + DII) with resonance assignments shown. (C) Full spectrum. (D) Enlarged view of the middle region (dotted box in C). The experiments were performed at 35 °C, pH 6.8, on a Bruker Avance III 600-MHz spectrometer.

2.4 Promiscuous 2-aminothiazoles (PrATs): A Frequent Hitting Scaffold

The unreliable data obtained for both 2-aminothiazole series hampered the progress towards the development of higher affinity binders of AMA1. In parallel to the 2-aminothiazole FBDD campaign, screening of the MIPS library against separate structurally diverse proteins for other projects continued to be performed at Monash University. Upon routine analysis of these individual screens within the MIPS cohort, the frequent identification of positive hits of fragments belonging to the 2-aminothazole chemical class for became apparent, signified by a 100% hit rate for 14/14 protein targets. The outcome of these screens coupled with the failure to identify any viable 2-aminothiazole SAR led to a separate study, under both a FBDD and HTS setting, exploring the possible promiscuous binding nature of the 2-aminothiazole scaffold. This work is included within this chapter due to my involvement in the assembling and preparation of SPR samples and subsequent SPR fragment testing, involving chip and plate preparation, Biacore instrument usage and the calculation of % occupancy for fragments screened. This culminated in the preparation of a publication in the Journal of Medicinal Chemistry of which I am a co-author. This work was pivotal to the preliminary investigation into the promiscuity of the 2-aminothiazole scaffold, subsequently resulting in in the eventual decision to abandon the development of 2-aminothiazole series for this body of work and eliminate such scaffolds from the MIPS fragment library.

Foreword Promiscuous 2-Aminothiazoles (PrATs): A Frequent Hitting Scaffold

This section focuses on the outcomes of the study into the examination of the 2-aminothiazole scaffold in both an FBDD and HTS screening context to identify whether the 2-aminothiazole core is a frequent hitting scaffold. The analysis of this scaffold looked at an academic fragment library used for FBDD and two larger compound libraries used for HTS. Analysis of the two larger compound libraries was completed in collaboration with J. Willem M. Nissink of AstraZeneca. This study revealed PrATs to behave as frequent hitters in both FBDD and HTS settings, though the problem was more prominent for fragment-based studies. In our opinion, the combination of their promiscuity and the difficulties associated with optimizing aminothiazoles, made them poor scaffolds for fragment libraries. However the 2-aminothiazole scaffold cannot be necessarily deemed undesirable considering a raft of 2-aminothiazoles are present as known drugs.

Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

Promiscuous 2-Aminothiazoles (PrATs): A Frequent Hitting Scaffold

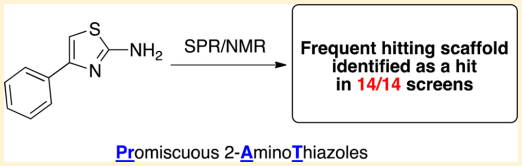
Shane M. Devine,[†] Mark D. Mulcair,[†] Cael O. Debono,[†] Eleanor W. W. Leung,[†] J. Willem M. Nissink,[‡] San Sui Lim,[†] Indu R. Chandrashekaran,[†] Mansha Vazirani,[†] Biswaranjan Mohanty,[†] Jamie S. Simpson,[†] Jonathan B. Baell,[†] Peter J. Scammells,[†] Raymond S. Norton,[†] and Martin J. Scanlon*[,†]

[†]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

[‡]Oncology iMed, AstraZeneca, Unit 310 − Darwin Building, Cambridge Science Park, Milton Road, Cambridge CB4 0WG, United Kingdom

Supporting Information

ABSTRACT: We have identified a class of molecules, known as 2-aminothiazoles (2-ATs), as frequent-hitting fragments in biophysical binding assays. This was exemplified by 4-phenylthiazol-2-amine being identified as a hit in 14/14 screens against a diverse range of protein targets, suggesting that this scaffold is a poor starting point for fragment-based drug discovery. This prompted us to analyze this scaffold in the context of an academic fragment library used for fragment-based drug discovery (FBDD) and two larger compound libraries used for high-throughput screening (HTS). This analysis revealed that such "promiscuous 2-aminothiazoles" (PrATs) behaved as frequent hitters under both FBDD and HTS settings, although the problem was more pronounced in the fragment-based studies. As 2-ATs are present in known drugs, they cannot necessarily be deemed undesirable, but the combination of their promiscuity and difficulties associated with optimizing them into a lead compound makes them, in our opinion, poor scaffolds for fragment libraries.



INTRODUCTION

Fragment-based drug discovery (FBDD) is becoming a widely used technique in drug discovery as part of a medicinal chemist's arsenal.[1,2] FBDD strategies utilize small molecules (typically ~200 Da), known as "fragments", that often have relatively low affinities for the target of interest. Nonetheless, numerous FBDD campaigns have demonstrated that it is feasible to elaborate fragment hits to achieve clinically useful compounds.[3] Furthermore, there are published examples where FBDD has enabled the development of potent compounds against targets where HTS of a large library did not yield any useful hits[4] and a review of internal projects by scientists at AstraZeneca revealed that FBDD represents a powerful tool to assess the likelihood of finding highly potent ligands for any given target.[5] This highlights one attraction of FBDD, which is that fragment libraries contain relatively small numbers of compounds (a few hundred to thousands) but are very effective at finding hits because they are able to sample chemical space more efficiently than the larger compounds that are typically found in HTS libraries.[6] One way of illustrating this is to consider that a library of druglike molecules of 30 heavy atoms might need up to 10^{60} members to efficiently cover chemical space, whereas this number is significantly smaller (around 10^{7}) for a library of fragments with around 12 heavy atoms.[7,8] As fragments are smaller, they typically bind with lower affinity to the relevant target, with K_D values ranging from high micromolar to millimolar. In contrast, druglike hits from HTS typically have K_D values in the high nanomolar to low micromolar range, although in both cases the binding energy per heavy atom (or "ligand efficiency") can be comparable. Fragments can also be developed into lead compounds that are smaller and less lipophilic than those generated from HTS.[9] A highlight of FBDD to date was the development of vemurafenib, a BRAF kinase inhibitor used for the treatment of late-stage melanoma, which, via medicinal chemistry optimization, became the first FBDD-derived compound to reach the clinic.[10,11] Since FBDD libraries are usually small, it is essential to ensure that the library is populated with high quality fragments. Some of the key considerations in the design of fragment libraries have been discussed previously.[12,13]

PAINS

A matter of growing concern with screening libraries, for FBDD or otherwise, is the inclusion of compounds that may act promiscuously and display activity both across different target classes and via a number of different assays or biological readouts. These types of compounds, which are widely referred to as pan assay interference compounds (PAINS), were first described by Baell and Holloway[14] and are identified by the presence of substructural features that promote frequent-hitting behavior. It has been suggested that they should be excluded from screening libraries for that reason. While in many cases PAINS may appear to give optimizable hits, elaboration often results in flat or confusing structure−activity relationships (SARs).[15] The reasons for their promiscuity are varied and

Received: September 12, 2014
Published: January 5, 2015

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

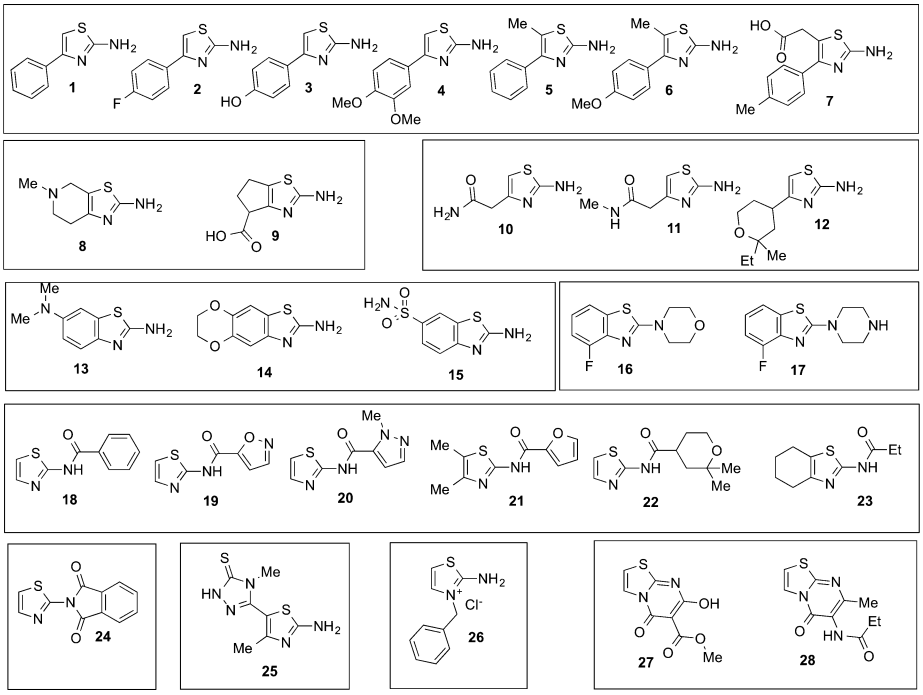


Figure 1. 2-Aminothiazole (2-AT) and 2-AT-like compounds in the MIPS fragment library.[12]

include the presence of potential Michael acceptors, chelation, redox activity, and strong chromophoric interference.[16] Some PAINS scaffolds have easily identifiable problems associated with them, but the chemical basis of the observed promiscuity for many PAINS is unknown. Confounding the complexity of PAINS identification is the fact that there are a small number of cases where seemingly "unprogressable" PAINS have in fact been developed into potent and selective molecules. One example is the PI3Kγ inhibitor (AS-604580), which is based on an alkylidene rhodanine hit containing this known PAINS motif.[17] However, the fact that certain PAINS or PAINS-like motifs are present in some potent and selective molecules does not imply that the PAIN is a viable starting point. In general, the chances that a PAIN will be progressed to a useful lead compound are overwhelmingly smaller than the chances it will not.[16]

There are 14 subclasses of 2-ATs that have been categorized as PAINS.[14,16] There are a number of possible reasons for their reported promiscuity, such as their potential photoreactivity or the presence of impurities generated by their chemical precursors, for example, where 2-ATs are prepared from bromomethyl ketones.[14] 2-ATs have also been identified as potentially thiol-reactive, which is another mechanism that produces problematic screening hits. For example a subclass of 2-ATs was identified in the ALARM-NMR assay as being thiol-reactive,[18,19] although in this case it is possible that their reactive precursors were responsible.[20]

That 2-ATs can be progressed to generate useful compounds is testified by a number of marketed drugs, including antibiotics such as carumonam, cefcapene, cefdinir, cefditoren, cefepime, cefetamet, cefoselis, cefotaxime, cefotiam, cefpodoxime, cefpirome, ceftazidime, ceftibuten, ceftriaxone; talipexole, and pramipexole, dopamine agonists for the treatment of Parkinson's disease; mirabegron, a β_3-adrenoceptor agonist used to treat overactive bladder; and riluzole, a 2-aminobenzothiazole used to treat acute myeloid leukemia (AML). Conversely, 2-ATs have displayed cytotoxicity and metabolic instability as antimycobacterial and antiplasmodial agents.[21] Thus, the value of 2-ATs in screening collections is currently unclear. Herein, we describe our efforts to determine if 2-ATs are promiscuous binders by both FBDD and HTS techniques.

RESULTS

FBDD Screening. FBDD screens were undertaken at the Monash Institute of Pharmaceutical Sciences (MIPS) using a library of 1137 fragments comprising molecules that pass both biophysical and chemical filters, are chemically and structurally diverse, are soluble at 1 mM in aqueous buffer, and can be chemically elaborated from readily accessible precursors.[12,22] Chemical filters that were applied in assembling the fragment library include the removal of PAINS,[14] unwanted functionality,[13] and reactive groups.[23] A review of 14 fragment screening campaigns with this library, in which the primary readout of fragment binding was saturation transfer difference nuclear magnetic resonance spectroscopy (STD-NMR),[24] revealed that at least one 2-AT from the library had been identified as a hit in every case.[12] This led to our investigation of their role as potential promiscuous binders using an orthogonal biophysical technique. We report here the results of our study and analysis of all 2-ATs in our fragment library (Figure 1) undertaken using surface plasmon resonance (SPR) against six different protein targets. Target proteins were the *Plasmodium falciparum*

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

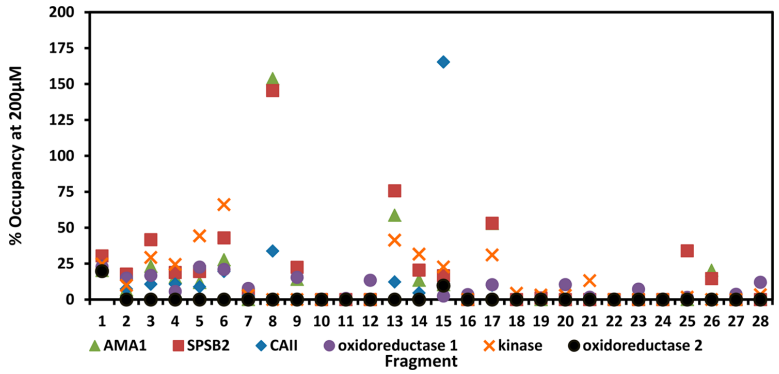


Figure 2. Binding of 2-ATs to AMA1, SPSB2, CAII, oxidoreductase 1, a kinase, and oxidoreductase 2. Compounds were tested at a concentration of 200 μM, and the raw response was converted to percentage occupancy.

apical membrane antigen AMA1,[22] the E3 ubiquitin ligase adapter protein SPSB2,[25] two DsbA oxidoreductases from different bacterial species (oxidoreductase 1 and 2), carbonic anhydrase II, and a kinase. These proteins were selected, as they exhibit diversity in function and have little structural similarity across their binding sites. All except oxidoreductase 2 had known positive controls that were used in the SPR assays. In addition, the fragment screen contained a number of internal controls. Four fragments used in this study were identified previously as binders to SPSB2 by SPR (**1**, **3**, **15**, **17**; our unpublished data). Carbonic anhydrase II was expected to show a very strong preference for compounds containing a terminal sulfonamide, and a prior fragment screen of oxidoreductase 2 (our unpublished data) had revealed only very weak binding for any fragment. Thus, oxidoreductase 2 served as a negative control.

Dose−response determinations were undertaken for positive controls (available for all proteins except oxidoreductase 2) and gave K_D values that were within the expected ranges for each target protein (Supporting Information Figures S1 and S2), confirming that all proteins were stable and active on the sensor chip under the conditions used. Binding of 2-ATs was carried out at a single concentration (200 μM) in duplicate, and the average raw response was converted to percentage occupancy of binding, corrected for the molecular weight of each compound (Figure 2). The sensorgrams for selected 2-ATs can be found in Supporting Information Figures S3−S5. Injections of buffer blank were included between each compound, both to provide for double referencing and to minimize carryover problems from poorly behaved compounds.

Figure 2 reveals a consistency in the pattern of protein binding. Only two 2-ATs (**1** and **15**) demonstrated any binding to oxidoreductase 2, which is consistent with previous observations that this protein has very limited capacity to bind small molecules. For the other five proteins, there were two cases (fragment **8** and **15**) where a large range in the fractional occupancy was observed, ranging from ∼0% to >150%. Otherwise, most fragments showed little discrimination in binding, with fragments typically binding to either all five proteins or none at all. While there were some exceptions, there was also a general tendency for fragments to bind each of the five proteins at similar occupancy, suggesting approximately the same affinity and thus providing very little clear SAR.

To examine the SAR in a more systematic manner, 2-ATs from the MIPS library were clustered based on linear fingerprints and key chemical features, and the screening results were examined for each grouping. Compounds **1**−**7** feature 4-aryl substitutions with a free 2-amino group, with or without a substituent at the 5-position. Fragments **8**−**12** maintain the free 2-amino group, with **8** and **9** containing fused aliphatic rings, whereas **10**−**12** have aliphatic substituents at the 4-position. Fragments **13**−**15** integrate benzothiazoles with their 2-amino group intact, while fragments **16** and **17** have morpholino or piperazino attachments through their 2-amino group, respectively. Fragments **18**−**24** are 2-amido containing thiazoles, and **27** and **28** are thiazolo[3,2-*a*]pyrimidines, in which the 2-amino group is incorporated in the heterocycle. The triazolothione (**25**) and thiazolium (**26**) complete the selection of 2-ATs from the MIPS library. These molecules contain diverse substituents around the 2-AT core. Thus, despite the chemical diversity in the 28 2-ATs in the library, evaluation of their binding to the different targets provided little clear SAR.

As an example, 4-phenylthiazol-2-amine (**1**) highlights the problem with this class of molecule, showing binding to all six proteins examined, all with similar occupancy levels (∼25%). As this fragment contains 12 heavy atoms, this level of binding corresponds to a favorable ligand efficiency (i.e., binding energy per heavy atom) in each case.[26] Thus, if this fragment were tested in an isolated setting against one particular protein, the likelihood of follow-up by medicinal chemists would be high, demonstrating the insidious behavior associated with this class. To rule out an impurity in the commercial preparation of fragment **1**, it was resynthesized and purified. This gave similar occupancy levels to the commercial product across all protein targets, indicating that the problem is inherent to the molecule itself, as opposed to reactive precursors or side products potentially present in the commercial source. Furthermore, the initial STD-NMR screening data demonstrated binding of **1** rather than precursors or side products and also revealed that the binding was noncovalent (Supporting Information Figure S6). This rules out several of the potential mechanisms for promiscuity, including photoreactivity, thiol reactivity, and the presence of reactive precursors.

Fragments **2**−**6** show binding to all of the proteins examined except oxidoreductase 2. The addition of the 5-methyl group appears to have little effect, as does the nature of the

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

levels of >100% of the theoretical R_{max} are observed for two proteins (AMA1 and SPSB2). Fragment **15** is the only sulfonamide in the series and demonstrates affinity to CAII, as expected for this protein since it specifically binds terminal sulfonamides, but several other fragments (for example **3**, **8**, **13**) also demonstrate binding to CAII.[27] The 2-substituted aminothiazoles are an interesting test case; morpholino (**16**) shows little to no binding to any protein tested, whereas piperazino (**17**) binds to five of the six proteins. The 2-amido and heterocyclic thiazoles (**18**−**24**, **27**, and **28**) show little or no binding to any of the proteins and do not appear to be inherently promiscuous. Triazolothione **25** and thiazolium **26** showed no binding to five and four of the targets, respectively. In summary, 2-ATs appear to be generally promiscuous at fragment screening concentrations, whereas the corresponding amides are not. Within the set of 2-ATs, there are no clear characteristics that distinguish the more promiscuous binders from their less promiscuous counterparts, and thus there is no clear SAR.

In order to extend the SPR studies a sample of 2-ATs was evaluated for binding to oxidoreductase 1 using 2D heteronuclear single quantum correlation (HSQC-NMR) studies. Oxidoreductase 1 was chosen for this analysis, as it is relatively small (21 kDa), expressed at high levels in bacteria, and amenable to isotope labeling and analysis by protein-detected NMR. Four of the fragments (**1**, **5**, **6**, and **14**) were tested for binding to oxidoreductase 1 by recording $^{1}H-^{15}N$ HSQC NMR in the presence and absence of each fragment. Binding was analyzed by measuring chemical shift perturbations (CSPs) in the spectrum of the protein upon addition of the fragment (Figure 3).

The SPR data revealed that **1**, **5**, and **6** showed similar levels of binding to the protein, whereas no binding was observed by SPR for **14**. In contrast, the extent of CSP observed in the NMR data suggested that while **1**, **5**, and **6** showed similar extents of CSP, fragment **14** showed the largest CSP in the spectra. This would normally suggest that **14** is the fragment that binds to oxidoreductase 1 with the highest affinity. However, analysis of the location and extent of CSP on the structure of oxidoreductase 1 (Figure 3E and Figure 3F) revealed clusters of strong CSP at distinct sites on the protein and a concentration dependence of the CSP that was inconsistent with 1:1 binding stoichiometry (Supporting Information Figure S7). Taken together, the NMR data suggest that **14** binds to oxidoreductase 1 at more than one site. For fragments **1**, **5**, and **6**, the smaller CSP observed makes it difficult to determine whether they are binding at more than one site.

Protein-detected NMR was also used to investigate the binding of 2-ATs to a second test protein SPSB2. In this case, ^{19}F NMR studies on 5-F-Trp-SPSB2 have proven effective as an analytical tool for determining binding to the active site of the protein.[28] Specific binding of peptides and ligands to the functionally important binding site is characterized by a substantial downfield shift of the fluorine resonance corresponding to Trp207 (Figure 4). In contrast, no downfield shift of this peak was observed for the 2-ATs that were observed to bind SPSB2 by SPR, despite their reasonable binding occupancy at 200 μM. This suggests that these 2-ATs do not bind to the active site of SPSB2 and that binding probably occurs at a secondary site on the protein. The ability to bind to more than one site on a protein may underpin the promiscuity that we have observed with some 2-ATs.

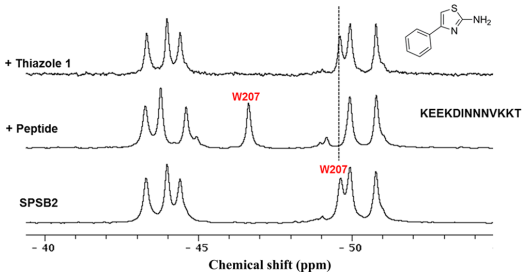


Figure 4. ^{19}F NMR spectra of 5-F-Trp-SPSB2 alone (bottom), with 500 μM control peptide (middle) and 3 mM **1** (top). Specific binding of the control peptide to the active site is characterized by a downfield shift of the peak corresponding to Trp207.[25] No such shift is observed for the thiazole fragment, suggesting a different mode of binding to SPSB2.

We have previously reported analysis of binding of a small set of 2-ATs to AMA1.[22] In the case of AMA1, 4-phenylthiazol-2-amine (**1**) was identified as a hit in the primary screen and a series of 4-aryl substituted 2-ATs was synthesized to generate preliminary SAR. The structures and percentage occupancy data of an extended series of 4-aryl substituted 2-AT analogues and their binding to AMA1 as determined by SPR are presented in Supporting Information Figures S8 and S9. All of the analogues were found to bind with low affinity to AMA1, and no clear SAR was derived from the series.

HTS Screening: Academic and Corporate Collections. We were interested in assessing the apparent promiscuity of 2-AT derivatives in fragment-based vs high-throughput screening. To analyze the latter, we adopted two approaches. The first was to analyze the relative prevalence of 2-AT-based PAINS in a HTS library of 93 000 compounds housed at the Walter and Eliza Hall Institute (WEHI).[29] The second was to analyze the full AstraZeneca HTS screening deck. With respect to the first approach, a search of the WEHI HTS library revealed 989 2-AT-based compounds and an additional 1012 compounds where the amine was acylated. Among these are 14 subclasses of PAINS, the structural definitions of which are given in Supporting Information Figure S10. This analysis was concordant with data from our fragment library in that no PAINS class contains 2-ATs in which the 2-amino group is acylated.

The incidence of frequent hitting 2-ATs was then investigated with greater statistical power using a large corporate collection (AstraZeneca, January 2014). These compounds have been tested in a range of HTS campaigns with concentrations typically around 10 μM. To assess the promiscuity of 2-ATs, we employed a descriptor that has been designed to flag potential frequent-hitter behavior for all compounds with sufficient data.[30] This descriptor is calculated by first designating a compound as active/inactive in each available screen using data from all historical AstraZeneca HTS campaigns. The body of HTS data that is available varies from compound to compound for a variety of reasons (age of compound, sample availability, membership of screen sets, manual collation), so potentially anomalous binding behavior was designated by comparing observed incidence of activity for a given compound to the expected activity for an "average" compound. In this way a descriptor (termed the pBSF score) was derived that indicates whether the compound is more

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

active than expected based on historical observations.[30] This knowledge-based descriptor is calculated as the negative logarithm of the probability of encountering the observed level of activity for a compound that behaves normally by chance. A low probability, and hence a high score, indicates that it is very unlikely that the observed level of activity would be observed for a well-behaved compound, which therefore suggests that the compound is a promiscuous hitter. We first analyzed the number of 2-ATs that had a pBSF score above a threshold score of 2.0. Second, we divided the 2-ATs into classes based on substructures and determined the proportion of each substructure that had pBSF scores of >2.0.

A comparison was made of the pBSF score for the 61 040 2-ATs within the AstraZeneca collection. For a visual comparison, a set of the same size was selected at random from the library, as documented in Figure 5. For the 2-ATs there were 6122

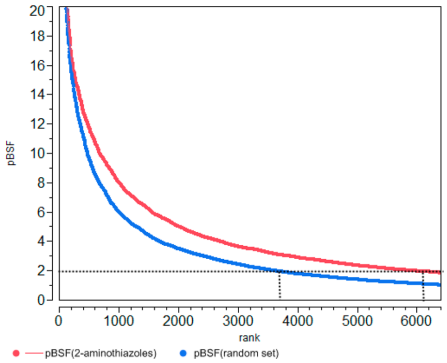


Figure 5. Rank-ordered frequent hitter scores for the first 6122 2-ATs (red line) from a total set of 61 040 2-ATs retrieved by substructure from the AstraZeneca screening collection. Dotted lines show the numbers of compounds at a threshold value of >2, which is indicative of anomalous binding behavior.

compounds with a score of >2 (incidence of 10.0%), whereas a randomly picked set of the same size has, on average, 3953 ± 61 compounds with a score of >2 (incidence of 6.5% ± 0.1%, as determined by a bootstrapping analysis). It is evident that the incidence of frequent hitters in the random set is significantly lower than for the 2-ATs (confidence $P > 99.9\%$, as derived from bootstrapping analysis), which suggests that the 2-ATs show a tendency to be promiscuous that is significantly higher than seen in a diverse set of randomly selected compounds.

We then divided the 2-AT structural classes using the substructures shown in Table 1 and counted the incidence of suspicious compounds using a threshold of >2 to distinguish potential frequent hitters from "clean" compounds and those lacking sufficient data. The data in Table 1 are consistent with the patterns observed with the MIPS FBDD and WEHI HTS data, namely, that 2-ATs show a higher incidence of anomalous binding behavior across a number of substructure classes. Typically, structures across the AstraZeneca collection that have on average ≤6% of the compounds within the class with a pBSF score of >2.0 are not considered to be frequent hitters. Here, an elevated incidence (approximately 10%) of such behavior can be seen for the overall class of 2-ATs and for subclasses 1, 3, 4, 5, and 7, which is in line with observations for the WEHI HTS library.

It is notable that promiscuity can be observed in the HTS data even within sets of low-MW compounds For example, within class 3, there are 1781 primary amines with MW < 300, of which 8% are classified as frequent hitters. Although these compounds are "fragment-like" in terms of their size, they are observed to be promiscuous in the HTS data despite being screened at a typical HTS concentration of 10 μM, which suggests that their promiscuity in the FBDD data is not due solely to the higher concentrations used in the fragment screen.

Subclass 9 comprises the acylated 2-ATs, which show a lower incidence of frequent-hitter behavior than the entire class of 2-ATs for both the low-MW and high-MW set. Nonetheless, the incidence of frequent-hitter behavior in the high-MW set is somewhat higher than the expected "normal" level of 6%, although this can be explained by biases that are apparent in target-specific analyses of the screening data. Inspection of a target-specific frequent hitter descriptor (results not shown) reveals that subclasses 3, 4, and 9 show an increased incidence of kinase activity. As kinases are typical drug targets pursued by pharma companies, a bias may be observed in the data where kinase-like motifs have been screened preferentially in kinase-targeted screens. This skews the descriptor results based on such data to some extent; i.e., some of the frequent hitters identified may actually be kinase-frequent hitters and not necessarily problematic. Within the 1657 potential frequent hitters of kinase-like class 9 (MW ≥ 300), a subset of 556 molecules appears to be kinase-frequent hitters. Subclass 7 contains a structural motif less likely to hit kinases, yet appears to have a high frequent-hitter incidence and contains 960 suspicious compounds (MW ≥ 300). Of these, a smaller fraction of only 136 displays kinase-related frequent hitter behavior, suggesting that the origin of frequent-hitting behavior in the remaining compounds is mostly not related to kinases. We suggest therefore that the somewhat increased incidence of frequent hitters we observe for the acylated 2-ATs in this data set is an artifact resulting from a kinase-activity bias and that the acylated 2-ATs as a whole do not meet the criteria to be categorized as frequent hitters.

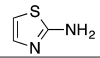
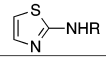
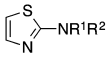
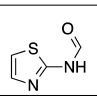
Using the AstraZeneca 2-AT set of 61 040 compounds and data derived from the AstraZeneca corporate collection, we investigated overall correlations of frequent hitter incidence with lipophilicity (experimental log *D*, clogP), experimental solubility (pSol, calculated as log[Sol/uM]), calculated polar and nonpolar surface area (PSA and NPSA in Å^2 and %PSA in %), and molecular volume.[31,32] Correlations are observed with ion class, PSA, %PSA, as well as donor and acceptor counts (Supporting Information Figure S11), suggesting that there could be a relation with polarity (as each of these properties reflects polarity and they are interrelated). A weak trend with log *D* (experimental octanol–water partitioning coefficient) was observed, with low log *D* compounds somewhat more likely to be frequent hitters. No relation is observed with experimental aqueous solubility, but a trend is evident with increasing clogP. Although there are no categorical reasons for promiscuity in these compounds, the observation that ionizable groups and high lipophilicity may increase anomalous behavior in the 2-AT class is in line with recent observations by Tarcsay and Keserű,[33] who observed similar trends.

DISCUSSION

We have demonstrated that certain members of the class of compounds containing a 2-AT substructure are frequent hitting and promiscuous fragments in the context of FBDD, where

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

Table 1. Nine Simplified 2-AT Classes Were Inspected for Promiscuous Behavior across the AstraZeneca Corporate Collection[a]

Class		Substructure[a]	N	N with data MW <300	N with data MW >=300	N(FH) MW <300	N(FH) MW >=300	fraction pBSF>2 LowMW	fraction pBSF>2 highMW
1	*aminothiazoles*		77,826	8,070	52,970	559	5,563	6.9	10.5
2	*fused aminothiazoles*		3,259	284	2,190	17	144	6.0	6.6
3	*primary amines*		8,678	1,781	4,425	143	564	8.0	12.8
4	*secondary amines*		55,445	5,336	38,337	372	3,970	7.0	10.4
5	*tertiary amines*		13,864	978	10,312	53	1,068	5.4	10.4
6	*tertiary amines, acyclic*		3,517	304	2.795	15	108	4.9	3.9
7	*tertiary amines, cyclic*		10,307	661	7,495	37	960	5.6	12.8
8	*thiazole cyclic amines ('embedded')*		59	21	25	0	0	0.0	0.0
9	*acylated thiazole amines*		29,816	2,613	21,447	99	1,657	3.8	7.7

[a]Counts are shown for subsets of small molecules (MW < 300) and larger compounds (MW ≥ 300). Structure legend: A, any atom; rn, ring bond; ch, chain bond; s*n*, substitution count *n*; r*n*, number of ring bonds

screening is undertaken using biophysical binding assays. We have dubbed these fragments promiscuous 2-aminothiazoles (PrATs). Analysis of binding of the 28 2-ATs in the MIPS fragment library was undertaken by SPR and NMR against six unrelated protein targets. Although some patterns have emerged, such as the presence of the free 2-amino group seeming to promote promiscuity, a clear mechanism of action has not been identified at this point. Nonetheless, our SPR analysis produced flat and confusing SAR against several of the targets, which has previously been noted in the characterization of PAINS identified in HTS.

As several approved drugs contain a 2-AT, it is clearly possible to accommodate this structure in a therapeutically useful compound. However, it is noteworthy that a number of PAINS subclasses also contain the 2-AT chemotype, which suggests that the 2-AT may carry some risks if selected for development. For example, in common with many aromatic amines, 2-ATs can be Ames positive dependent on the substitution pattern. Further, aminothiazoles have been associated with liver toxicity via bioactivation of the double bond, leading to formation of thioureas that are further bioactivated. Problems relating to reactivity and covalent modification, however, are unlikely to contribute to the observed promiscuity in biophysical binding assays, as the ligand-detected NMR assays used in FBDD provide evidence of both compound identity and binding simultaneously, while both NMR and SPR discriminate between covalent and noncovalent interactions. Therefore, the FBDD data suggest that the binding observed is noncovalent and can be attributed to the 2-ATs themselves rather than reactive precursors or impurities in the samples.

Analysis of the AstraZeneca HTS data reveals that 2-ATs show elevated frequency as screening hits relative to the compound library as a whole, although the percentage (10%) is significantly lower than the corresponding value for other known PAINS classes, which can be around 15−20%.[30] The AstraZeneca data also suggest that acylation of the 2-amino group ameliorates the problem of frequent hitting, which indicates that not all 2-ATs are problematic in the context of HTS. Similarly, only 3.2% of 2-AT-containing compounds in the WEHI HTS library are defined as PAINS, which does not provide a strong case for exclusion of all 2-ATs from HTS collections.[14] Thus, the two HTS analyses are in broad agreement with each other and suggest that the majority of 2-ATs do not seem to be promiscuous at HTS-relevant concentrations in the 10−25 μM range. Even at HTS-relevant concentrations, some optimizable hits can be relatively promiscuous and we have previously identified a 2-amino-

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

benzothiazole derivative[16] that hits four out of the six HTS assays selected for the PAINS analysis,[14] yet led to a highly selective and potent compound.[34]

Nonetheless, many 2-ATs were observed to be problematic in both HTS and FBDD, suggesting that the promiscuity observed in FBDD is not due solely to the higher concentration used in the fragment screens. While the higher hit rates that are expected in FBDD dictate that many fragments are likely to hit more than one target, these hits are only useful if they can be elaborated into more potent compounds. On this basis, it is possible that certain substructures such as 2-ATs, which show some limited promiscuity in HTS assays but would not be flagged for exclusion from a fragment library by analysis of chemical or physical properties, are in fact undesirable as members of a fragment library based on a retrospective analysis of their behavior in screening assays. A similar approach has been described previously in the HTS setting for deprioritizing compounds that are known to be promiscuous.[35]

CONCLUSION

We have identified 2-ATs as a promiscuous substructure in screens of our fragment library. Each of the 28 fragments containing a 2-AT substructure passed all of the biophysical and chemical filters that were used in designing the library and were demonstrated to have appropriate purity and aqueous solubility. However, on the basis of our findings reported here and our unsuccessful attempts to optimize these fragments against different targets, we have removed 2-ATs from the fragment library at MIPS, on the basis that we judge the prospects for such compounds to be progressable as much smaller than the chances that they will not.

EXPERIMENTAL SECTION

General Procedures. Biacore sensor chips, *N*-hydroxysuccinimide (NHS), *N*-ethyl-*N*′-(3-dimethylaminopropyl)carbodiimide (EDC), ethanolamine HCl, HBS-P running buffer, and P-20 surfactant were obtained from GE Healthcare. Carbonic anhydrase II and 4-carboxybenzylsulfonamide were obtained from Sigma-Aldrich. Isabelle Lucet (WEHI) provided the kinase protein and positive control H4. Doubly His-tagged oxidoreductases 1 and 2 were provided from Martin Williams (MIPS). All 2-AT fragments were obtained from Chembridge or Life. The aqueous solubility, identity, and purity (>95%) of the purchased compounds were verified by analysis of 1D ^{1}H NMR spectra as described previously.[12]

Surface Plasmon Resonance (SPR) Conditions. Expression and purification of hexahistidine-tagged AMA1[36] and GST-tagged SPSB2[25] were as described previously. Binding of 2-ATs to the target proteins was measured on a Biacore T200 instrument (GE Healthcare). AMA1, SPSB2, and CAII were immobilized onto a CM5 sensor chip (Biacore) by standard amine coupling chemistry using sodium acetate at pH 4.5 in running buffer A. Doubly His_6-tagged oxidoreductase proteins were immobilized onto an NTA chip (Biacore) charged with Ni^{2+} according to the manufacturer's instructions, in running buffer B. Singly His_6-tagged kinase was immobilized to an NTA sensor chip charged with Ni^{2+} using the capture−couple method,[37] in running buffer C. Immobilization levels were typically 9000 RU for AMA1, 6000 RU for SPSB2 and CAII, and 2000−3000 RU for both oxidoreductases and the kinase.

Running buffer A: 25 mM HEPES, 150 mM NaCl, 5% DMSO, 0.005% P-20, pH 7.4.

Running buffer B: 25 mM HEPES, 200 mM NaCl, 5% DMSO, 0.005% Tween-20, pH 8.0.

Running buffer C: 25 mM HEPES, 150 mM NaCl, 4 mM $MgCl_2$, 2 mM TCEP, 3% DMSO, 0.005% Tween-20, pH 7.5.

SPR Screening of 2-ATs. Binding assays were carried out at 25 °C using a flow rate of 100 μL/min in running buffer A (AMA1, SPSB2, CAII), running buffer B (oxidoreductase 1 and 2), or running buffer C (kinase). 200 mM fragment stocks were diluted in the appropriate running buffer to obtain 200 μM working concentrations. The association and dissociation phases of binding were each followed for 30 s. An identical injection of running buffer was carried out between each fragment injection, and the average of adjacent blanks was subtracted from the raw fragment response. Raw sensorgram data were reduced, solvent-corrected, and double-referenced using BIAEvaluation Software (GE Healthcare). Raw responses were converted to percentage occupancy by the method of Giannetti[38] using an R_{max} value based on the response measured with a saturating concentration of the appropriate control and the molecular weight of each fragment. For oxidoreductase 2, there is no positive control available, and binding is reported as percentage of the theoretical R_{max}.

SPR Positive Controls. The R1 peptide[39] was used as a control for AMA1. A peptide corresponding to residues within the N-terminus of inducible nitric oxide synthase was used as a control for SPSB2.[25] 4-Carboxybenzylsulfonamide was used as a control for CAII. An elaborated fragment "H4" was used as a control for the kinase. An elaborated fragment "LA010" (our unpublished data) was used as a control for oxidoreductase 1. As oxidoreductase 2 has demonstrated a weak ability to bind fragments (unpublished data), no positive control was available for this protein. Dose−response determinations were carried out for all positive controls to confirm protein activity. A concentration series of 2-fold dilutions in running buffer was used, with a top concentration of 1 μM (R1), 5 μM (iNOS peptide), 10 μM (4-CBS), 20 μM (H4), or 200 μM (LA010) (see Supporting Information Figure S1).

Screening by STD-NMR. Screening was performed by recording STD NMR experiments on cocktails of up to six fragments in the presence of the target protein (1−10 μM). Screens were undertaken at 600 or 800 MHz on spectrometers equipped with cryogenically cooled probes. Where possible, screens were carried out in phosphate buffer prepared in >90% ^2H_2O, to minimize spectral interference from either buffer signals or the ^1H_2O resonance, and at a pH close to 7.0 and temperature of 10 °C. Other buffer conditions were used where necessary as dictated by the stability of the target protein. The resulting data were processed in Topspin (Bruker Biospin) and analyzed manually. STD signal intensities in the spectra were qualitatively classified as strong, medium, or weak based on the relative intensity of signals observed in the low-field region (>5.0 ppm). Relative intensities were based on the most intense STD signal (I_{max}) identified across all the STD spectra for a particular target. A positive STD signal was categorized as strong where the intensity was >50% I_{max}, medium where the intensity was >25% I_{max} and <50% I_{max}, or weak where the intensity was <25% I_{max}. If the fragment contained no resonances in the low-field region of the ^{1}H NMR spectrum, the aliphatic region (<4.5 ppm) was analyzed and such fragments were considered hits if any positive STD signal was observed. No attempt was made to categorize aliphatic STD signals because of the potential for direct excitation by the on-resonance saturating pulse in the STD experiment and/or interference from overlapping protein resonances (see Supporting Information Figure S6).

Analysis of Binding by HSQC-NMR. To analyze the location of the fragment binding site on oxidoreductase 1, 2D ^{1}H−^{15}N HSQC NMR spectra were recorded in the absence and presence of fragments **1**, **5**, **6**, and **14**. A reference ^{1}H−^{15}N HSQC of the protein (125 μM [U−^{15}N]-labeled protein; 2% DMSO-d_6; 50 mM HEPES; 50 mM NaCl; pH 6.8) was acquired and compared with a spectrum acquired under the same conditions for a sample containing the oxidized oxidoreductase 1 in the presence of each fragment (125 μM [U−^{15}N]-labeled protein; 1−1.5 mM fragment; 1−2% DMSO-d_6; 50 mM HEPES; 50 mM NaCl; pH 6.8). Subsequently, the protein was titrated with **14** (0.09−1.5 mM). 3 mm thick-walled tubes of sample volume ∼160 μL were used for NMR data collection. All data were acquired on a Bruker 600 MHz spectrometer equipped with autosampler and CryoProbe at 300 K. Standard acquisition and processing parameters were used throughout (see Supporting Information Figure S7).

^{19}F NMR Studies of 5-Trp-SPSB2. Construction of the 5-F-Trp-SPSB2 protein and ^{19}F NMR determination of thiazole binding were

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

carried out as described previously.[28] ^{19}F NMR studies were carried out at 30 °C in 50 mM phosphate, 50 mM NaCl (pH 7.4), with a 5-F-Trp-SPSB2 concentration of 100 μM, with or without 500 μM control peptide or 3 mM **1**, as indicated.

■ ASSOCIATED CONTENT

Supporting Information

SPR sensorgrams of positive controls and selected compounds, STD, HSQC-NMR, and SPR-derived binding data for the different targets, 2-AT PAINS subclasses, frequent hitter incidence, and SMILES string identifiers in csv format. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: martin.scanlon@monash.edu. Phone: +61 3 99039540.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Martin Williams for the oxidoreductases 1 and 2, Luke Adams for the oxidoreductase 1 positive control, Hiromasa Yagi for assistance on the ^{19}F NMR work, and Isabelle Lucet for the kinase protein sample and positive control. This work was supported in part by grants from the Australian National Health and Medical Research Council (NHMRC) (Grants 1045995, 1009785, 1022693, 1025150). J.B.B. and R.S.N. acknowledge fellowship support from the NHMRC.

■ ABBREVIATIONS USED

2-AT, 2-aminothiazole; AMA1, apical membrane antigen 1; CAII, carbonic anhydrase II; CSP, chemical shift perturbation; FBDD, fragment-based drug discovery; HSQC, heteronuclear single quantum correlation; HTS, high-throughput screening; MIPS, Monash Institute of Pharmaceutical Sciences; NMR, nuclear magnetic resonance; PAINS, pan assay interference compounds; PrATs, promiscuous 2-aminothiazoles; SPSB2, SPRY domain-containing SOCS box protein 2; SAR, structure−activity relationship; STD, saturation transfer difference; WEHI, Walter and Eliza Hall Institute

■ REFERENCES

(1) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274*, 1531−1534.

(2) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-Based Drug Discovery. *J. Med. Chem.* **2004**, *47*, 3463−3482.

(3) Tsai, J.; Lee, J. T.; Wang, W.; Zhang, J.; Cho, H.; Mamo, S.; Bremer, R.; Gillette, S.; Kong, J.; Haass, N. K.; Sproesser, K.; Li, L.; Smalley, K. S. M.; Fong, D.; Zhu, Y.-L.; Marimuthu, A.; Nguyen, H.; Lam, B.; Liu, J.; Cheung, I.; Rice, J.; Suzuki, Y.; Luu, C.; Settachatgul, C.; Shellooe, R.; Cantwell, J.; Kim, S.-H.; Schlessinger, J.; Zhang, K. Y. J.; West, B. L.; Powell, B.; Habets, G.; Zhang, C.; Ibrahim, P. N.; Hirth, P.; Artis, D. R.; Herlyn, M.; Bollag, G. Discovery of a Selective Inhibitor of Oncogenic B-Raf Kinase with Potent Antimelanoma Activity. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 3041−3046.

(4) LaPlante, S. R.; Padyana, A. K.; Abeywardane, A.; Bonneau, P.; Cartier, M.; Coulombe, R.; Jakalian, A.; Wildeson-Jones, J.; Li, X.; Liang, S.; McKercher, G.; White, P.; Zhang, Q.; Taylor, S. J. Integrated Strategies for Identifying Leads That Target the NS3 Helicase of the Hepatitis C Virus. *J. Med. Chem.* **2014**, *57*, 2074−2090.

(5) Edfeldt, F. N. B.; Folmer, R. H. A.; Breeze, A. L. Fragment Screening To Predict Druggability (Ligandability) and Lead Discovery Success. *Drug Discovery Today* **2011**, *16*, 284−287.

(6) Fink, T.; Bruggesser, H.; Reymond, J.-L. Virtual Exploration of the Small-Molecule Chemical Universe below 160 Da. *Angew. Chem., Int. Ed.* **2005**, *44*, 1504−1508.

(7) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent Developments in Fragment-Based Drug Discovery. *J. Med. Chem.* **2008**, *51*, 3661−3680.

(8) Fink, T.; Reymond, J.-L. Virtual Exploration of the Chemical Universe up to 11 Atoms of C, N, O, F: Assembly of 26.4 Million Structures (110.9 Million Stereoisomers) and Analysis for New Ring Systems, Stereochemistry, Physicochemical Properties, Compound Classes, and Drug Discovery. *J. Chem. Inf. Model.* **2007**, *47*, 342−353.

(9) Murray, C. W.; Carr, M. G.; Callaghan, O.; Chessari, G.; Congreve, M.; Cowan, S.; Coyle, J. E.; Downham, R.; Figueroa, E.; Frederickson, M.; Graham, B.; McMenamin, R.; O'Brien, M. A.; Patel, S.; Phillips, T. R.; Williams, G.; Woodhead, A. J.; Woolford, A. J.-A. Fragment-Based Drug Discovery Applied to Hsp90. Discovery of Two Lead Series with High Ligand Efficiency. *J. Med. Chem.* **2010**, *53*, 5942−5955.

(10) Bollag, G.; Hirth, P.; Tsai, J.; Zhang, J.; Ibrahim, P. N.; Cho, H.; Spevak, W.; Zhang, C.; Zhang, Y.; Habets, G.; Burton, E. A.; Wong, B.; Tsang, G.; West, B. L.; Powell, B.; Shellooe, R.; Marimuthu, A.; Nguyen, H.; Zhang, K. Y. J.; Artis, D. R.; Schlessinger, J.; Su, F.; Higgins, B.; Iyer, R.; D'Andrea, K.; Koehler, A.; Stumm, M.; Lin, P. S.; Lee, R. J.; Grippo, J.; Puzanov, I.; Kim, K. B.; Ribas, A.; McArthur, G. A.; Sosman, J. A.; Chapman, P. B.; Flaherty, K. T.; Xu, X.; Nathanson, K. L.; Nolop, K. Clinical Efficacy of a RAF Inhibitor Needs Broad Target Blockade in BRAF-Mutant Melanoma. *Nature* **2010**, *467*, 596−599.

(11) Bollag, G.; Tsai, J.; Zhang, J.; Zhang, C.; Ibrahim, P.; Nolop, K.; Hirth, P. Vemurafenib: the First Drug Approved for BRAF-Mutant Cancer. *Nat. Rev. Drug Discovery* **2012**, *11*, 873−886.

(12) Doak, B. C.; Morton, C. J.; Simpson, J. S.; Scanlon, M. J. Design and Evaluation of the Performance of an NMR Screening Fragment Library. *Aust. J. Chem.* **2014**, *66*, 1465−1472.

(13) Baurin, N.; Aboul-Ela, F.; Barril, X.; Ben Davis; Drysdale, M.; Dymock, B.; Finch, H.; Fromont, C.; Richardson, C.; Simmonite, H.; Hubbard, R. E. Design and Characterization of Libraries of Molecular Fragments for Use in NMR Screening against Protein Targets. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2157−2166.

(14) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719−2740.

(15) Hansch, C.; Fujita, T. P-Σ-Π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J. Am. Chem. Soc.* **1964**, *86*, 1616−1626.

(16) Baell, J. B.; Ferrins, L.; Falk, H.; Nikolakopoulos, G. PAINS: Relevance to Tool Compound Discovery and Fragment-Based Screening. *Aust. J. Chem.* **2013**, *66*, 1483−1494.

(17) Camps, M.; Rückle, T.; Ji, H.; Ardissone, V.; Rintelen, F.; Shaw, J.; Ferrandi, C.; Chabert, C.; Gillieron, C.; Françon, B.; Martin, T.; Gretener, D.; Perrin, D.; Leroy, D.; Vitte, P.-A.; Hirsch, E.; Wymann, M. P.; Cirillo, R.; Schwarz, M. K.; Rommel, C. Blockade of PI3Kgamma Suppresses Joint Inflammation and Damage in Mouse Models of Rheumatoid Arthritis. *Nat. Med.* **2005**, *11*, 936−943.

(18) Huth, J. R.; Mendoza, R.; Olejniczak, E. T.; Johnson, R. W.; Cothron, D. A.; Liu, Y.; Lerner, C. G.; Chen, J.; Hajduk, P. J. ALARM NMR: A Rapid and Robust Experimental Method To Detect Reactive False Positives in Biochemical Screens. *J. Am. Chem. Soc.* **2005**, *127*, 217−224.

(19) Huth, J. R.; Song, D.; Mendoza, R. R.; Black-Schaefer, C. L.; Mack, J. C.; Dorwin, S. A.; Ladror, U. S.; Severin, J. M.; Walter, K. A.; Bartley, D. M.; Hajduk, P. J. Toxicological Evaluation of Thiol-Reactive Compounds Identified Using a La Assay To Detect Reactive Molecules by Nuclear Magnetic Resonance. *Chem. Res. Toxicol.* **2007**, *20*, 1752−1759.

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

(20) Metz, J. T.; Huth, J. R.; Hajduk, P. J. Enhancement of Chemical Rules for Predicting Compound Reactivity towards Protein Thiol Groups. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 139−144.
(21) Mjambili, F.; Njoroge, M.; Naran, K.; De Kock, C.; Smith, P. J.; Mizrahi, V.; Warner, D.; Chibale, K. Synthesis and Biological Evaluation of 2-Aminothiazole Derivatives as Antimycobacterial and Antiplasmodial Agents. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 560−564.
(22) Lim, S. S.; Debono, C. O.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening. *Aust. J. Chem.* **2013**, *66*, 1530−1536.
(23) Chuprina, A.; Lukin, O.; Demoiseaux, R.; Buzko, A.; Shivanyuk, A. Drug- and Lead-Likeness, Target Class, and Molecular Diversity Analysis of 7.9 Million Commercially Available Organic Compounds Provided by 29 Suppliers. *J. Chem. Inf. Model.* **2010**, *50*, 470−479.
(24) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784−1788.
(25) Kuang, Z.; Lewis, R. S.; Curtis, J. M.; Zhan, Y.; Saunders, B. M.; Babon, J. J.; Kolesnik, T. B.; Low, A.; Masters, S. L.; Willson, T. A.; Kedzierski, L.; Yao, S.; Handman, E.; Norton, R. S.; Nicholson, S. E. The SPRY Domain-Containing SOCS Box Protein SPSB2 Targets iNOS for Proteasomal Degradation. *J. Cell Biol.* **2010**, *190*, 129−141.
(26) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand Efficiency: A Useful Metric for Lead Selection. *Drug Discovery Today* **2004**, *9*, 430−431.
(27) Supuran, C. T.; Scozzafava, A. Carbonic Anhydrase Inhibitors and Their Therapeutic Potential. *Expert Opin. Ther. Pat.* **2000**, *10*, 575−600.
(28) Leung, E. W. W.; Yagi, H.; Harjani, J. R.; Mulcair, M. D.; Scanlon, M. J.; Baell, J. B.; Norton, R. S. (19) F NMR as a Probe of Ligand Interactions with the iNOS Binding Site of SPRY Domain-Containing SOCS Box Protein 2. *Chem. Biol. Drug Des.* **2014**, *84*, 616−625.
(29) Baell, J. B. Broad Coverage of Commercially Available Lead-like Screening Space with Fewer Than 350,000 Compounds. *J. Chem. Inf. Model.* **2013**, *53*, 39−55.
(30) M Nissink, J. W.; Blackburn, S. Quantification of Frequent-Hitter Behavior Based on Historical High-Throughput Screening Data. *Future Med. Chem.* **2014**, *6*, 1113−1126.
(31) Bruneau, P. Search for Predictive Generic Model of Aqueous Solubility Using Bayesian Neural Nets. *J. Chem. Inf. Model.* **2001**, *41*, 1605−1616.
(32) Bruneau, P.; McElroy, N. R. logD7.4 Modeling Using Bayesian Regularized Neural Networks. Assessment and Correction of the Errors of Prediction. *J. Chem. Inf. Model.* **2006**, *46*, 1379−1387.
(33) Tarcsay, Á.; Keserü, G. M. Contributions of Molecular Properties to Drug Promiscuity. *J. Med. Chem.* **2013**, *56*, 1789−1795.
(34) Lessene, G.; Czabotar, P. E.; Sleebs, B. E.; Zobel, K.; Lowes, K. N.; Adams, J. M.; Baell, J. B.; Colman, P. M.; Deshayes, K.; Fairbrother, W. J.; Flygare, J. A.; Gibbons, P.; Kersten, W. J. A.; Kulasegaram, S.; Moss, R. M.; Parisot, J. P.; Smith, B. J.; Street, I. P.; Yang, H.; Huang, D. C. S.; Watson, K. G. Structure-Guided Design of a Selective BCL-X(L) Inhibitor. *Nat. Chem. Biol.* **2013**, *9*, 390−397.
(35) Nissink, J. W. M.; Schmitt, S.; Blackburn, S.; Peters, S. Stratified High-Throughput Screening Sets Enable Flexible Screening Strategies from a Single Plated Collection. *J. Biomol. Screening* **2014**, *19*, 369−378.
(36) Hodder, A. N.; Crewther, P. E.; Anders, R. F. Specificity of the Protective Antibody Response to Apical Membrane Antigen 1. *Infect. Immun.* **2001**, *69*, 3286−3294.
(37) Rich, R. L.; Errey, J.; Marshall, F.; Myszka, D. G. Biacore Analysis with Stabilized G-Protein-Coupled Receptors. *Anal. Biochem.* **2011**, *409*, 267−272.
(38) Giannetti, A. M. From Experimental Design to Validated Hits a Comprehensive Walk-Through of Fragment Lead Identification Using Surface Plasmon Resonance. *Methods Enzymol.* **2011**, *493*, 169−218.
(39) Harris, K. S.; Casey, J. L.; Coley, A. M.; Masciantonio, R.; Sabo, J. K.; Keizer, D. W.; Lee, E. F.; McMahon, A.; Norton, R. S.; Anders, R. F.; Foley, M. Binding Hot Spot for Invasion Inhibitory Molecules on *Plasmodium falciparum* Apical Membrane Antigen 1. *Infect. Immun.* **2005**, *73*, 6981−6989.

DOI: 10.1021/jm501402x
J. Med. Chem. 2015, 58, 1205−1214

2.5 Conclusion

Based upon 9 thiazole-based hits identified from the MIPS library, the rational design and synthesis of two independent series of 2-aminthiazoles with the aim of revealing novel 2-aminothiazole derivatives demonstrating higher affinity binding to AMA1. These studies involved the chemical elaboration of both the 4- and 5-aryl groups. For each of the series investigated, initial screening by SPR was considered the primary screen which would consequently been followed by further analysis through R1 competition NMR studies which would focus on the improved binders identified from the primary SPR screen.

Screening results of SPR did not reveal any improvement in affinity for both series when compared to the original MIPS library fragments. The common observation of poor compound solubility and concentration-dependent aggregation for a majority of the in-house synthesized analogues tested led to the identification of non-stoichiometric binders by SPR. Efforts to improve the solubility of optimized analogues was unsuccessful resulting in analogues, which could never reach saturable binding since the maximal testing concentration was limited to 200 μM due to solubility issues. We subsequently identified the 2-aminothiazole as a promiscuous substructure in screens against a variety of protein targets. Although the scaffold was able to pass all of the biophysical and chemical filters used to design the MIPS library utilized for this body of work, the findings of the study supported the removal of all 2-aminothiazole fragments from the fragment library. As suggested by this study, the 2-aminothiazole scaffold was eventually abandoned owing to the inability to develop meaningful SAR or crystallize any compounds shown to be competed off by R1.

Although no higher affinity specific binders of AMA1 resulted from this traditional FBDD investigation, the scaffold remained a prominent figure for several studies involving NMR techniques, such as 2D [^{1}H-^{15}N]-TROSY NMR and ^{19}F NMR. This work was completed in collaboration with other members of the group leading to a further two publications (Refer to Appendix A, sections A.2 and A.3).

2.6 Bibliography

(1) Cantrell, A. S.; Hoegberg, M.; Jaskunas, S. R.; Johansson, N. G.; Jordan, C. L.; Kinnick, M. D.; Lind, P.; Morin Jr, J. M.; Bell, F. W. Phenethylthiazolethiourea (PETT) Compounds, a New Class of HIV-1 Reverse Transcriptase Inhibitors. 1. Synthesis and Basic Structure-Activity Relationship Studies of PETT Analogs. *Journal of Medicinal Chemistry* **2001**, *38*, 1–8.
(2) Liu, Z.-Y.; Wang, Y.-M.; Li, Z.-R.; Jiang, J.-D.; Boykin, D. W. Synthesis and anticancer activity of novel 3,4-diarylthiazol-2(3H)-ones (imines). *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 5661–5664.
(3) Li, M.; Sim, Y.-J.; Ham, S.-W. Discovery of 2-Aminothiazole Derivatives as Antitumor Agents. *Bulletin of the Korean Chemical Society* **2010**, *31*, 1463–1464.
(4) Satoh, A.; Nagatomi, Y.; Hirata, Y.; Ito, S.; Suzuki, G.; Kimura, T.; Maehara, S.; Hikichi, H.; Satow, A.; Hata, M.; Ohta, H.; Kawamoto, H. Discovery and in vitro and in vivo profiles of 4-fluoro-*N*-[4-[6-(isopropylamino)pyrimidin-4-yl]-1,3-thiazol-2-yl]-*N*-methylbenzamide as novel class of an orally active metabotropic glutamate receptor 1 (mGluR1) antagonist. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 5464–5468.
(5) Laine, L.; Kivitz, A. J.; Bello, A. E.; Grahn, A. Y.; Schiff, M. H.; Taha, A. S. Double-Blind Randomized Trials of Single-Tablet Ibuprofen/High-Dose Famotidine vs. Ibuprofen Alone for Reduction of Gastric and Duodenal Ulcers. *The American Journal of Gastroenterology* **2011**, *107*, 379–386.
(6) Borelli, C.; Schaller, M.; Niewerth, M.; Nocker, K.; Baasner, B.; Berg, D.; Tiemann, R.; Tietjen, K.; Fugmann, B.; Lang-Fugmann, S.; Korting, H. C. Modes of Action of the New Arylguanidine Abafungin beyond Interference with Ergosterol Biosynthesis and *in vitro* Activity against Medically Important Fungi. *Chemotherapy* **2008**, *54*, 245–259.
(7) Gonzàlez Cabrera, D.; Douelle, F.; Feng, T.-S.; Nchinda, A. T.; Younis, Y.; White, K. L.; Wu, Q.; Ryan, E.; Burrows, J. N.; Waterson, D.; Witty, M. J.; Wittlin, S.; Charman, S. A.; Chibale, K. Novel Orally Active Antimalarial Thiazoles. *Journal of Medicinal Chemistry* **2011**, *54*, 7713–7719.
(8) Cohen, A.; Verhaeghe, P.; Crozet, M. D.; Hutter, S.; Rathelot, P.; Vanelle, P.; Azas, N. Tandem synthesis and *in vitro* antiplasmodial evaluation of new naphtho[2,1-d]thiazole derivatives. *European Journal of Medicinal Chemistry* **2012**, *55*, 315–324.
(9) Lim, S. S.; Debono, C. O.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening. *Australian Journal of Chemistry* **2013**, *66*, 1530–1536.
(10) Hajduk, P. J. Fragment-Based Drug Design: How Big Is Too Big?. *Journal of Medicinal Chemistry* **2006**, *49*, 6972–6976.
(11) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531–1534.
(12) Hantzsch, A.; Weber, J. H. Ueber Verbindungen des Thiazols (Pyridins der Thiophenreihe). *Berichte der Deutschen Chemischen Gesellschaft.* **1887**, *20*, 3118–3132.
(13) Dodson, R. M.; King, L. C. The reaction of ketones with halogens and thiourea. *Journal of the American Chemical Society* **1945**, *67*, 2242.
(14) Ceres-Castillo, D. C.; Carballo, R. M.; Tzec-Interián, J. A.; Mena-Rejón, G. J. Solvent-free synthesis of 2-amino-4-arylthiazoles under microwave irradiation. *Tetrahedron Letters* **2012**, *53*, 3934–3936.

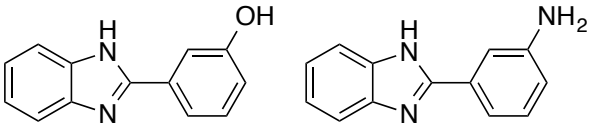
(15) Skoumbourdis, A. P.; Huang, R.; Southall, N.; Leister, W.; Guo, V.; Cho, M.-H.; Inglese, J.; Nirenberg, M.; Austin, C. P.; Xia, M.; Thomas, C. J. Identification of a potent new chemotype for the selective inhibition of PDE4. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 1297–1303.
(16) Giannetti, A. M.; Koch, B. D.; Browner, M. F. Surface Plasmon Resonance Based Assay for the Detection and Characterization of Promiscuous Inhibitors. *Journal of Medicinal Chemistry* **2008**, *51*, 574–580.
(17) Bondi, A. van der Waals Volumes and Radii. *The Journal of physical chemistry* **2001**, *68*, 441–451.
(18) Krogsgaard-larsen, P.; Liljefors, T.; Madsen, U. Qantitative structure-activity relationships and experimental design. *Textbook of Drug Design and Discovery.* CRC press, **2002**, 3, 117-156.
(19) Miyaura, N.; Yanagi, T.; Suzuki, A. The Palladium-Catalyzed Cross-Coupling Reaction of Phenylboronic Acid with Haloarenes in the Presence of Bases. *Synthetic Communications* **2006**, *11*, 513–519.
(20) Davis, B.; Erlanson, D. A. Learning from our mistakes: the "unknown knowns" in fragment screening. *Bioorganic & Medicinal Chemistry Letters* **2013**, *23*, 2844–2852.
(21) Krishnarjuna, B.; Lim, S. S.; Devine, S. M.; Debono, C. O.; Lam, R.; Chandrashekaran, I. R.; Jaipuria, G.; Yagi, H.; Atreya, H. S.; Scanlon, M. J.; MacRaild, C. A.; Scammells, P. J.; Norton, R. S. Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials. *Journal of Molecular Recognition* **2016**, *29*, 281–291.
(22) Ge, X.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Wang, G.; Scammells, P. J.; Scanlon, M. J.; Anders, R. F.; Foley, M.; Norton, R. S. Ligand-Induced Conformational Change of *Plasmodium falciparum* AMA1 Detected Using ^{19}F NMR. *Journal of Medicinal Chemistry* **2014**, *57*, 6419–6427.

Chapter 3 – Evolution of the benzimidazole scaffold

3.1 Development of the 2-aryl benzimidazoles

3.1.1 Background

Benzimidazoles have been widely employed as antimicrobial agents over the last few decades owing to their ability to elicit a broad range of pharmacological effects.[1] Because of their synthetic utility and broad range of pharmacological effects, compounds exhibiting the benzimidazole pharmacophore have been shown to possess anticancer,[2] antiallergic,[3] antiparasitic[4] and antimalarial activity.[5,6] The structural and therapeutic diversity, coupled with broad-spectrum commercial availability of benzimidazole-based molecules, has amplified the examination of this scaffold within FBDD. From the original MIPS library screen, two 2-substituted aryl benzimidazole fragments were identified as positive hits, which were validated via R1 competition STD and CPMG NMR.[7] Further investigation via SPR of the benzimidazole positive hits identified benzimidazole fragments **51** and **52** (**Figure 3.1**) as strong binders for AMA1, with compound **52**, demonstrating an affinity of 500 μM against both 3D7 and FVO AMA1 whilst displaying favourable LE.[8] Given the undefined SAR resulting from our investigation of the 2-aminothiazole series, the chemical elaboration of another key chemical class from the MIPS screen, being the 2-substituted aryl benzimidazole scaffold, was initiated. The 2-substituted aryl benzimidazole scaffold was another chemical class to satisfy the key selection criteria of strong affinity, high LE, good physicochemical properties (i.e. solubility) and synthetic feasibility (i.e. ease of optimization). These characteristics were used to rank the eight chemical classes, identified via the fragment screen in search of an ideal candidate for an FBDD study (Chapter 1, Figure 1.14). This chemical class thereby became a suitable scaffold to pursue through chemical elaboration in search of higher-affinity, strain-transcending, binders of AMA1. In addition, prior to the optimization of benzimidazole fragments identified from the library screen, Ramachandran *et al*.[9] reported that benzimidazole derivatives of structural similarity to **51** and **52** demonstrated potent antimalarial activity (IC_{50}). This work was also able to investigate a wide range of malarial strains, including strains known to display resistance to current clinical antimalarials drugs. These resistant strains include the chloroquine-resistant strains, K1 and Dd2. Although the benzimidazole chemical class was considered one of the smallest classes to be recognized from the MIPS library, the advantages of exploring the SAR of the benzimidazole series were favourable for the improvement in binding affinity that had eluded the investigation into the 2-aminothiazole scaffold.



51
K_D = 800 μM
LE = 0.26

52
K_D = 500 μM
LE = 0.28

Figure 3.1 Benzimidazole fragments identified through fragment-based screening.[7]

3.1.2 Screening

Screening of the MIPS fragment library revealed the 2-substituted aryl benzimidazoles to display competition in the presence of R1 via competition NMR. Unfortunately only 2 of the 35 benzimidazole-containing fragments within the MIPS library were 2-substituted aryl benzimidazoles (**Figure 3.1**). Considering the poor representation of the 2-substituted aryl benzimidazole scaffold within the MIPS fragment library, we focused on the chemical synthesis of this compound class to generate a focused series of analogues with the aim of developing an SAR profile for this scaffold. 2-substituted aryl benzimidazoles possessing various functional groups of differing spatial distribution and physicochemical properties around the 2'-aryl ring were synthesized. The first generation 2-substituted aryl benzimidazole analogues were designed with importance given to properties including solubility (ionization), reactivity and steric effects. Any commercially available derivatives there-of were purchased to enhance the SAR for the given scaffold.

Prior to the chemical expansion of the benzimidazole scaffold, an in-house library screen was initiated utilizing the Wellcome Trust. A systematic examination of the Wellcome Trust Library was initiated to study additional compounds demonstrating derivatization of the aniline moiety and to identify higher affinity binders with good LE. The Wellcome Trust library was configured to investigate the optimization of the 2'-position of the benzimidazole core. Based upon the base structure of 4-(1*H*-benzo[*d*]imidazol-2-yl)aniline (**Figure 3.2**), four different chemical classes designated as either belonging to a 2-substituted aryl, amide-linked, urea-linked or *N*-alkyl-linked class were investigated. Each compound was functionalized through the *p*-amino substituent of the base structure (**Figure 3.2**). The screening of the Wellcome Trust library was considered a cost-effective strategy employed to efficiently screen for SAR, providing a greater opportunity to cover larger chemical space without completing extensive chemical elaboration on the scaffold so early in the study.

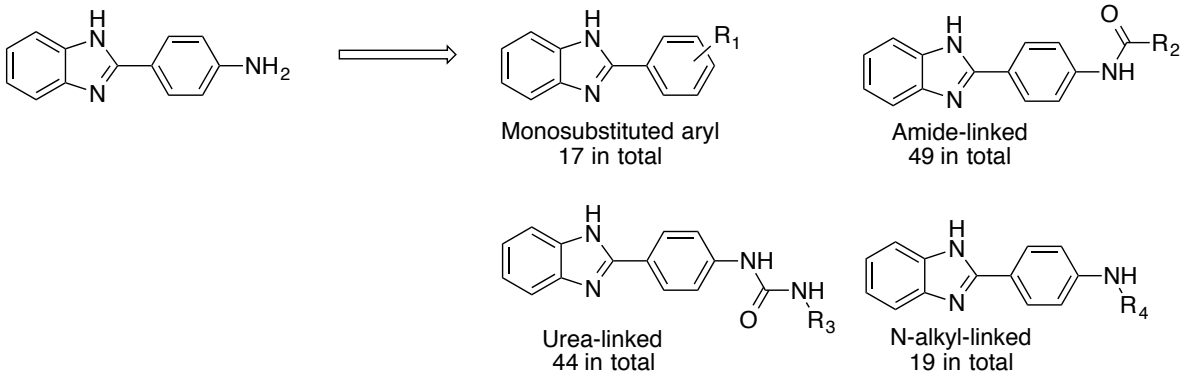


Figure 3.2 Make-up of Wellcome Trust library displaying the *p*-amino base structure for the development of the four different functional classes investigated. The number of molecules per class is also shown.

In total, 63 Wellcome Trust library compounds were tested using the original stock samples that were made up to 200 mM. Any precipitated stocks were excluded. The screening regime was slightly altered from previously completed campaigns as viable Wellcome Trust fragments were initially only tested at single point concentration via SPR. By implementing this strategy, the efficiency of the screening campaign was improved since all library fragments could be screened in the one run limiting the impact of experimental error on assay results that often result when attempting multiple runs.[10] In addition, SPR was now considered the primary screen for this study in place of the protein-intensive NMR techniques employed during earlier validation campaigns involving MIPS library fragments. Each positive hit of the initial single dose study was subsequently followed up by an SPR dose-response (5-point) investigation. All attractive compounds resulting from the dose response screen were then individually tested via competition STD and CPMG NMR[7] to identify whether hits were competed off by R1 at the hydrophobic cleft of AMA1.

The Wellcome Trust library was screened via SPR against 3D7 *Pf*AMA1$_{[104-442]}$ corresponding to domains I and II of the AMA1 ectodomain. The protein was immobilized on a CM5 chip using amine coupling.[11] Each compound was screened as a single dose concentration of 200 μM maintaining 30 s contact times; compounds that demonstrated a response were consequently tested at five different concentrations of 12.5, 25, 50, 100 and 200 μM also with 30 s contact times to determine compound affinity. The known AMA1 binder R1 was utilized as a positive control, which was screened alongside library compounds to confirm whether AMA1 remained active throughout run. All binding affinities were derived as previously discussed (Chapter 2, section 2.2) for hits of the MIPS library. Unlike previous screening data which were ranked principally based

upon LE, each positive hit from the Wellcome Trust library was ranked based upon the fractional surface occupancy (θ),[10] which was used to convert compound response units (RU) to a percentage of available specific binding sites occupied. Converting to a percentage occupancy, was accomplished by equation 3.3,

$$\theta = (R_{med}[100C/(C + K_D)])/R_{control} \qquad \textbf{(Equation 3.3)}$$

where C is the control concentration, K_D is the control's equilibrium dissociation constant determined by the dose-response curve for each compound, R_{med} is a given compound's baseline-corrected value in RU, and $R_{control}$ is the response of the control at concentration C in RU. For all binders from the Wellcome Trust library, the % occupancy (converted from θ) was calculated based on the 50 μM concentration of each compound; this concentration was selected to avoid non-stoichiometric binding which may result at higher concentrations. To efficiently analyze data recorded and calculated, a filter was introduced to help focus on the seemingly higher affinity binding compounds. As a general rule, compounds that displayed a binding affinity weaker than 1 mM were designated as inactive and were not investigated further via competition NMR.

Overall 24 of the 63 benzimidazole analogues tested displayed no affinity for AMA1. Of the remaining 39 positive hits identified from the Wellcome Trust library, only 19 compounds displayed a binding affinity stronger than 1 mM (K_D < 1 mM) (**Figure 3.4**). The calculated occupancy of these 19 compounds ranged from as low as 7% to as high as 91%, although any compounds with a percentage occupancy > 70% at 50 μM prompted a review of the original sensorgram data to identify whether non-stoichiometric binding was present.

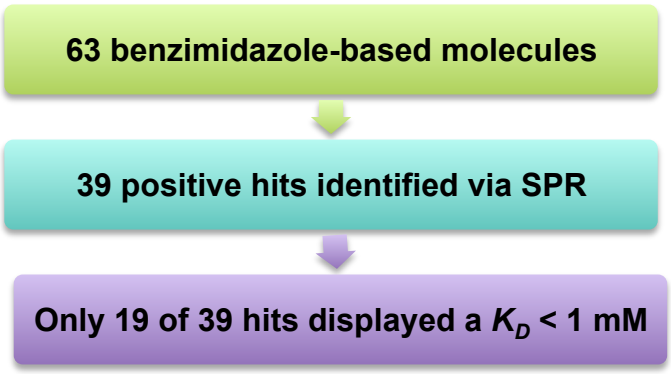


Figure 3.4 Breakdown of Wellcome Trust screening campaign.

A vast majority of the 20 compounds that displayed a binding affinity weaker than 1 mM (K_D > 1 mM) were considered concentration-dependent aggregators identified to behave beyond that expected of a 1:1 interaction as compound concentration was increased (**Figure 3.5**). Upon review of the original sensorgram data, the lack of stoichiometric binding for the majority of compounds within the concentration range testing suggested the possible presence of multisite binding for compounds against AMA1. Overall, no positive hits were calculated to display percentage occupancy greater than 100%, showing no signs of promiscuous behaviour for the benzimidazole scaffold. Owing to the seemingly promiscuous nature of the 2-aminothiazole series previously discussed,[12] this was considered a major finding in regards to proceeding to investigate and elaborate the benzimidazole scaffold.

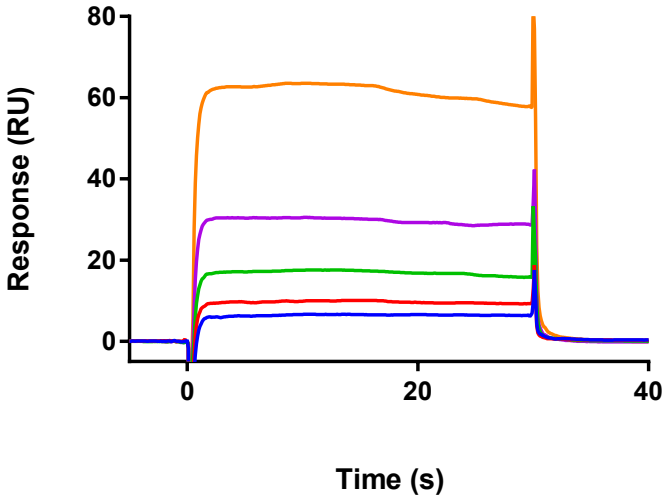


Figure 3.5 SPR sensorgram of a Wellcome Trust compound displaying concentration-dependent aggregation. All compounds shown to display K_D > 1 mM displayed a similar binding profile.

STD and CPMG competition NMR were employed to verify whether the better binders of the Wellcome Trust library were competed by R1. Considering the protein-intensive nature of both STD and CPMG NMR, as a precaution each of the 19 hits underwent a 1D ^{1}H NMR aggregation study to determine whether any of the hits had a tendency to self-aggregate in aqueous media;[13,14] each compound was tested at a range of concentrations (5, 10, 20, 40 and 80 µM) using 1D ^{1}H NMR spectroscopy. The concept of the NMR dilution assay came from the concept that aggregates should be sensitive to changes in concentration, and that the resultant ^{1}H NMR resonances should report structural changes in the aggregate and its environment (**Figure 3.6**). Compounds in aqueous solution adopt equilibria between three broad states, described as soluble single molecules, soluble aggregate entities and solid form (precipitate). These states can be clearly identified and distinguished using 1D ^{1}H NMR.

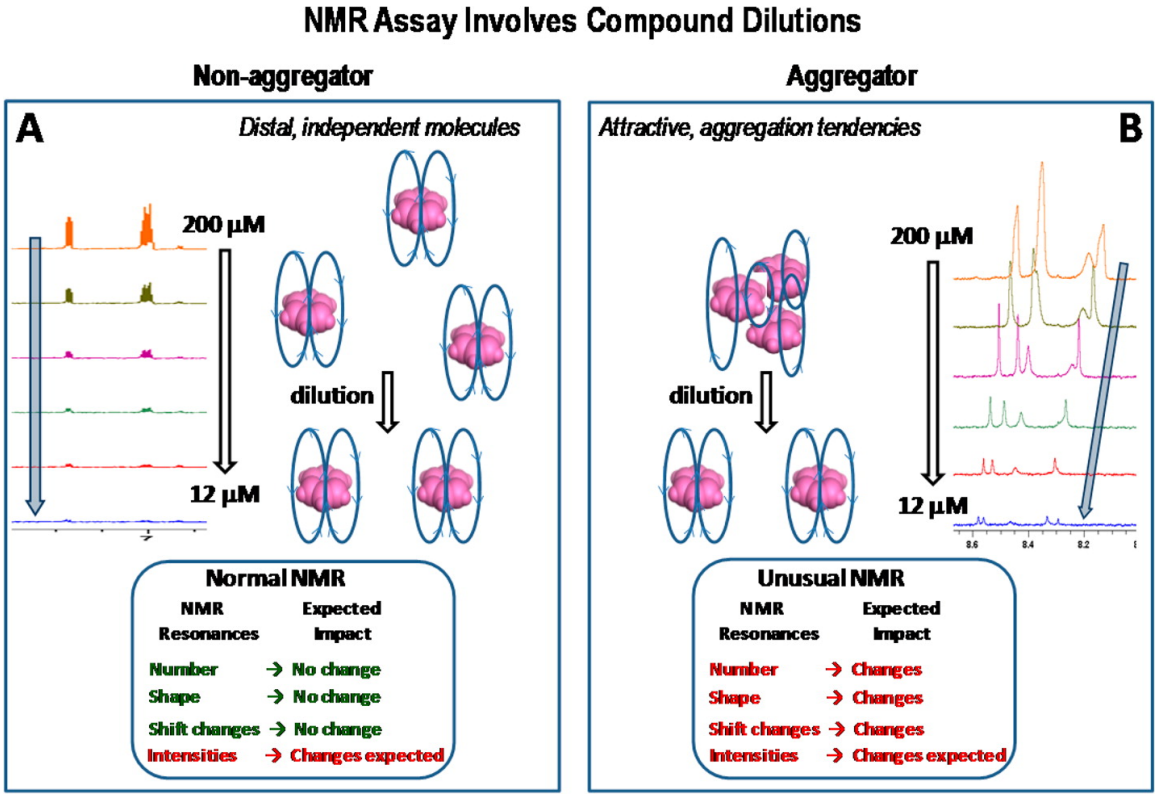


Figure 3.6 An illustration of how non-aggregating and aggregating molecules may behave in solution.[13] (A) Non-aggregating compounds are expected to remain distal, tumble rapidly, and move independently at all concentrations leading to sharp signal peaks. (B) Aggregating compounds are expected to have unusual NMR spectral attributes due to attractive tendencies at higher concentration that dissociate upon dilution.[13]

All compounds were assessed at a range of concentrations (5, 10, 20, 40, 80 and for some compounds 120 μM) using 1D ^{1}H NMR spectroscopy. Overall 17 of the 19 compounds selected for competition NMR studies showed clear evidence of aggregation in aqueous media (**Figure 3.7**). The bottom of Figure 3.6 lists the NMR parameters considered sensitive to these concentration studies, such as resonance number, shape, shifts, and intensities. As described by LaPlante *et al.*,[13] non-aggregating compounds are expected to tumble rapidly and not be impacted by dilution in aqueous buffer. As observed, changes to the shape and resonance number, as shown by SYN-428 of Figure 3.7 B, clearly identified each of the 17 compounds to display aggregation when tested at higher concentrations.

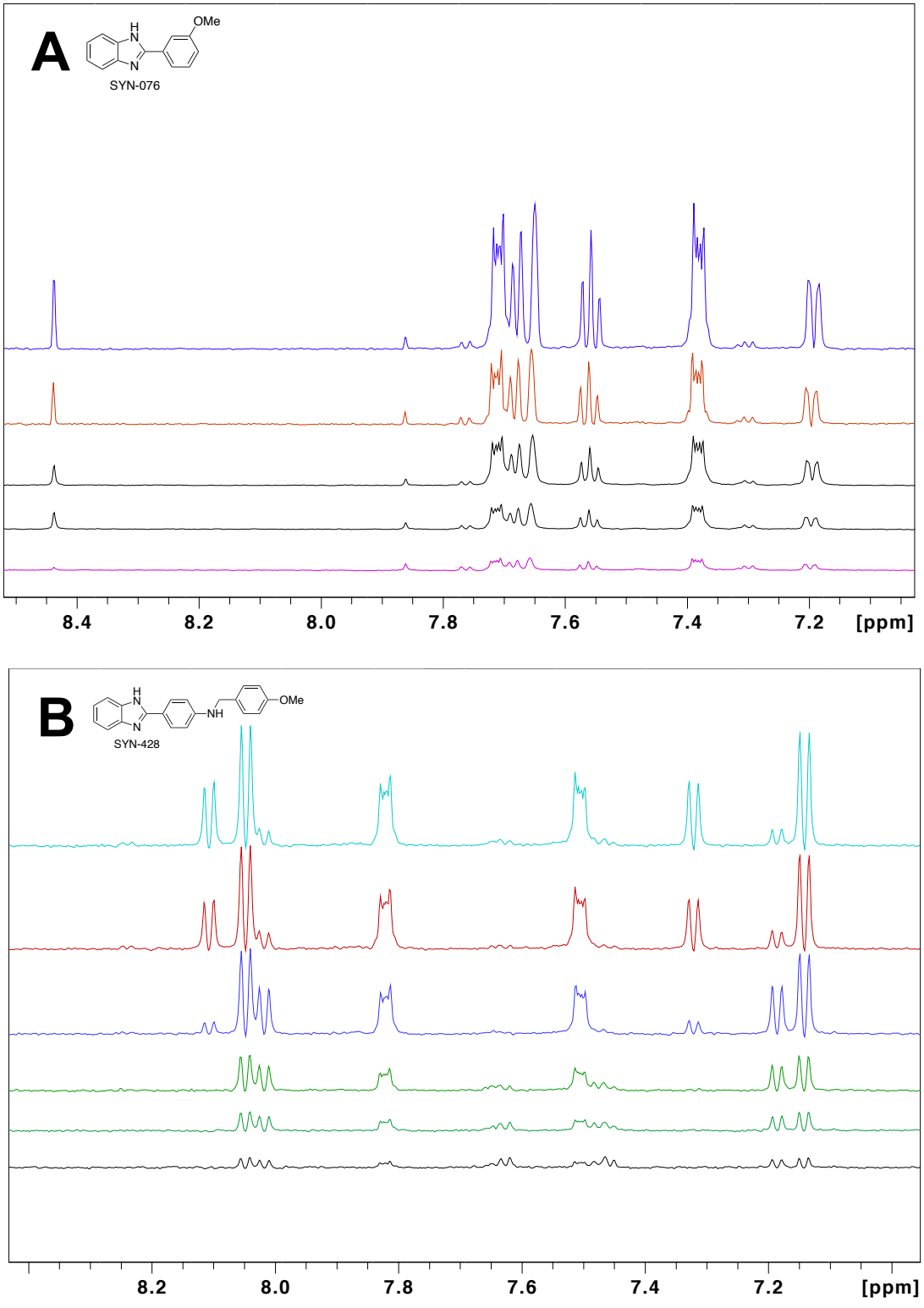


Figure 3.7 1D ^{1}H NMR spectra of (A) a non-aggregator, SYN-076, shown to display increasing peak intensity with increasing compound concentrations. (B) An aggregator, SYN-428 (Refer to **Figure 3.8**), shown to lose and gain certain signal intensities with increasing compound concentrations.

All 17 compounds (**Figure 3.8**) displayed very weak signals that showed no improvement when higher concentration samples were tested. Although compound solubility remained an issue when setting up NMR samples, a majority of the resultant NMR samples appeared to be homogeneous with no visible precipitate identified by the naked eye, suggesting that compounds could be self-associating to form colloidal aggregates over the entire concentration range tested.[13] Surfactant is often used to dissociate large assemblies of compound into smaller entities improving relaxation properties for NMR detection.[14] Thus the addition of surfactant (Tween 20) to NMR samples suggested to be aggregating was performed confirming the presence of colloidal aggregates in NMR samples. Tween 20 was able to improve NMR resonances slightly in which signals were observed to be consistent with what was expected for the compound in question, albeit with significant residual line broadening.

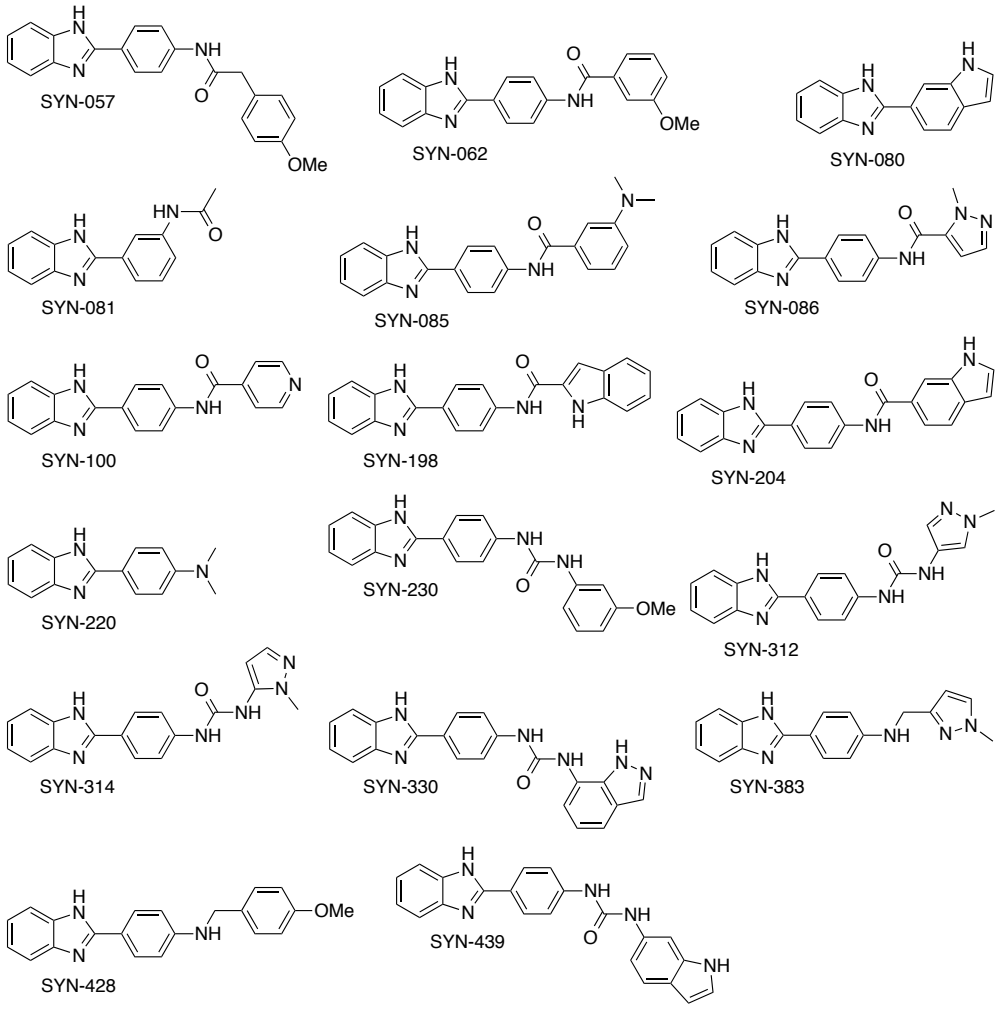
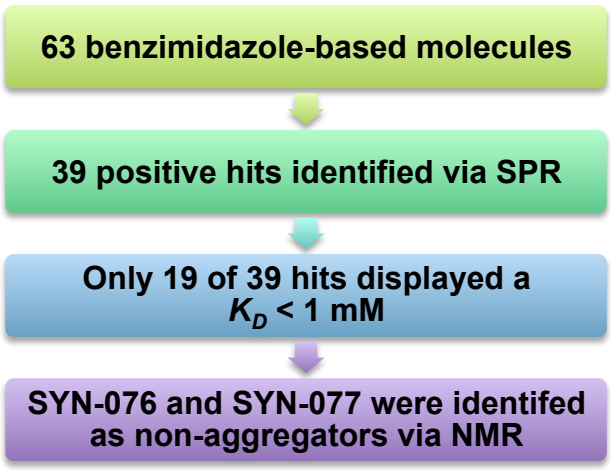


Figure 3.8 Wellcome Trust molecules (17 in total) shown to aggregate during NMR testing.

Only compounds SYN-076 and SYN-077 of the Wellcome Trust library (3% hit rate) showed no sign of aggregation in aqueous media. In addition, both compounds displayed R1 competition via STD and CPMG NMR and showed similar binding affinity for 3D7 and FVO AMA1, and thus were verified as true binders of AMA1 (**Figure 3.9**). Both analogues belonged to the 2-substituted aryl benzimidazole class, further confirming the scaffold as an ideal starting point for an FBDD study.



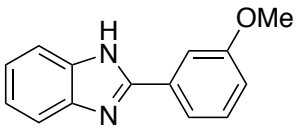
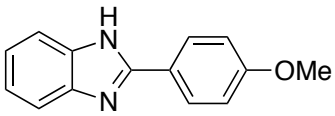
		
Compound ID	SYN-076	SYN-077
LogP	3.1	3.1
LE	0.31	0.27
K_D (mM)	0.12	0.63
θ*	46	23
1D NMR aggregation	No	No
STD NMR	Strong STD NMR	Strong STD NMR
CPMG NMR	Increased relaxation rate	Increased relaxation rate
Competed by R1	Yes	Yes
Binds to FVO	Yes	Yes

Figure 3.9 Screen flow-chart and results of screening campaign. *Surface occupancy (θ) is represented as a percentage occupancy (%).

At this stage, two independent compound libraries confirmed the validity of the 2-substituted aryl benzimidazole scaffold as a suitable FBDD candidate, with the *m*-methoxy substituted aryl benzimidazole, SYN-076, demonstrating concentration-dependent binding (**Figure 3.10**). Importantly, SYN-076 was also shown to bind to FVO AMA1, suggesting the presence of strain-transcending affinity. In relation to the originally proposed addition of functionality via the amino moiety of **52** of the MIPS library**,** the solubility issues encountered when screening Wellcome Trust benzimidazoles discouraged the requirement for functionalized benzimidazoles through the *m*-amino moiety of **52**. This observation directed the chemical elaboration of the benzimidazole scaffold to concentrate on the direct functionalization of the 2-aryl ring for valid SAR.

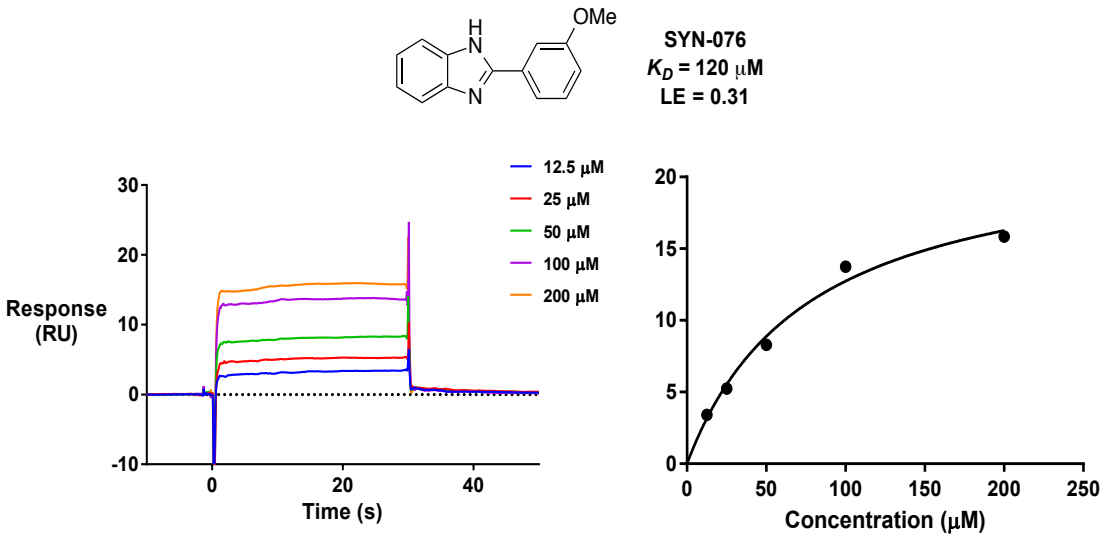


Figure 3.10 SPR sensorgram and curve for SYN-076 tested at five different concentrations (12.5, 25, 50, 100 and 200 µM).

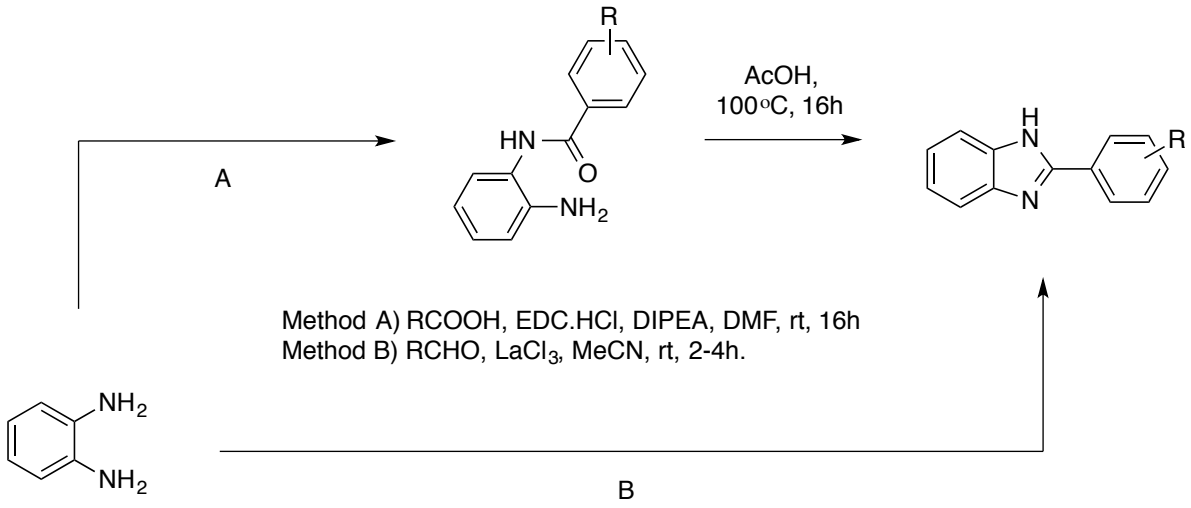
3.1.3 2-Substituted aryl benzimidazole series

Validation of the benzimidazole scaffold was completed via the combination of SPR, R1-competition NMR and TROSY NMR prior to the elaboration of the 2-substituted aryl benzimidazole series. The benzimidazole analogues, **51** and **52** (**Figure 3.1**) in parallel to the SAR investigation of the benzimidazole series, were shown to perturb the spectrum of AMA1 using two-dimensional [^{1}H-^{15}N]-TROSY NMR (Refer to Appendix A, section A.3).[15] These experiments focused on the sequence-specific backbone assignments for *Pf*AMA1 for two *Pf* strains, 3D7 and FVO. These experiments were completed to assign protein resonances to help map the binding sites for small molecules in search of structural information to guide further chemical synthesis. Selective labelling and unlabelling strategies were implemented to complement triple-resonance

experiments in order to facilitate the assignment of each AMA1 protein. Analogues **51** and **52** were able to bind at the more conserved DII end of the hydrophobic cleft, identified as the more promising site to target with regards to developing analogues with strain-transcending affinity for *Pf*AMA1.

The TROSY NMR experiment became an informative tool for the development of strain-transcending high-affinity binders for AMA1 since all attempts at obtaining crystal structures of the MIPS or Wellcome Trust compounds bound to either 3D7 or FVO AMA1 failed. Included within this study was a screen of multiple binding chemical classes identified from the MIPS library to rank the classes; this study also identified the 2-substituted aryl benzimidazoles as the next best candidate for chemical elaboration since 2-aminothiazole series were revealed as frequent hitters. Coworkers and myself completed the synthesis of all required compounds utilized during testing.

An SAR profile based on the 2-substituted aryl benzimidazole scaffold was initiated with the preliminary aim being the incorporation of additional functionality around the aryl group (**Scheme 1**). This work was guided by the TROSY NMR data compiled for **51** and **52**; this data encouraged the search for stronger interactions with the backbone residues of both 3D7 and FVO AMA1. The investigation of each arene position (*o*-, *m*- and *p*-position) around the 2'-position phenyl group was initially considered a top priority, with the introduction of various functional groups around the aryl ring. Owing to pronounced chemical shift perturbations of a 4-residue section, 'KDGG' (Lys177-Asp178-Gly179-Gly180) by **51** and **52**,[15] functional groups of variable molecular size and electronegativity were investigated to maximize binding potential in and around the perturbed region. Considering the charged nature of the 'KDGG' binding pocket, and the ability to form salt bridges via the presence of free amines and carboxylates, an emphasis on the exploration of functional groups such as the amino ($-NH_2$) and hydroxyl (–OH) groups which possess the ability to form both H-bonds and electrostatic interactions was evident. Additionally, more lipophilic functional groups such as the trifluoromethyl ($-CF_3$) group were also investigated considering the close vicinity of the 'KDGG' binding pocket to the hydrophobic cleft.



	Method	Yield (%)		Method	Yield (%)		Method	Yield (%)
42 R=*o*-OH	B	7	**51** R=*m*-OH	A	23	**61** R=*p*-OH	B	7
43 R=*o*-NH_2	*	30	**52** R=*m*-NH_2	A	15	**62** R=*p*-NH_2	B	22
44 R=*o*-OMe	B	20	**53** R=*m*-OMe	A	38	**63** R=*p*-OMe	B	13
45 R=*o*-CF_3	A	19	**54** R=*m*-CF_3	A	61	**64** R=*p*-CF_3	A	51
46 R=*o*-Me	A	28	**55** R=*m*-Me	A	15	**65** R=*p*-Me	A	6
47 R=*o*-Br	A	97	**56** R=*m*-Br	A	8	**66** R=*p*-Br	A	4
48 R=*o*-Cl	B	13	**57** R=*m*-Cl	B	29	**67** R=*p*-Cl	B	4
49 R=*o*-F	B	11	**58** R=*m*-F	A	4	**68** R=*p*-F	B	21
50 R=*o*-NO_2	B	28	**59** R=*m*-OCF_3	A	49			
			60 R=*m*-OEt	A	26			

Scheme 1 Synthesis of 2-substituted aryl benzimidazole analogues.[16,17] *Synthesized via hydrogenation of **50** - H_2, Pd/C (5%), MeOH, rt, 16 h.

2-Substituted aryl benzimidazole analogues were prepared from the condensation of *o*-phenylenediamine and various substituted carbonyl-containing compounds (**Scheme 1**).[16,17] Benzimidazole compounds originating from compound libraries were re-synthesised for further testing.

For method A of Scheme 1, the initial coupling step involved the utilization of the highly soluble coupling reagent, *N*-(3-dimethylaminopropyl)-*N*′-ethylcarbodiimide (EDC), to act as a carboxyl activating reagent between *o*-phenylenediamine and substituted benzoic acids. The coupling step was allowed to stir overnight at rt to ensure all *o*-phenylenediamine starting material had reacted. Although EDC is considered a common and highly effective coupling reagent, the likelihood of overactivity directed towards the additional amine moiety of *o*-phenylenediamine was limited owing to steric factors and the relatively mild reaction conditions implemented. The coupled intermediate was extracted via sat. $NaHCO_3$ solution to remove any residual benzoic acid starting

material from the organic phase. No further purification was required. The afforded crude material was tested via LCMS to identify the presence of the coupled intermediate prior to initiating the cyclization step. Upon confirmation, the crude material was dissolved in glacial AcOH and heated at 100 °C for 16 h to initiate the free amine attack of the carbonyl group, resulting in the loss of H_2O (condensation) when cyclized to form the imidazole ring.[18,19] The resultant isolated yield via method A of Scheme 1 for synthesized analogues varied from 4-97 % isolated yield; where deactivating group such as the halide containing derivatives generally tended to display poorer yields when either containing a fluoro- or bromo-substituent (<10 % yield).

Optimization of method A was conducted via the implementation of method B of Scheme 1 which utilized lanthanum chloride ($LaCl_3$) to react substituted aryl aldehydes with *o*-phenylenediamine.[20] Although no significant improvement in yield was noted (4-29 %), Method B was able to more efficiently afford desired product resulting in a reduced reaction time; 2-3 h at rt compared to Method A which required longer reaction times and higher temperatures at the expense of more difficult isolations and increased reaction by-products. Method B used the mild Lewis acid, $LaCl_3$ to activate the carbonyl, effectively allowing for the oxidative cyclodehydrogenation of Schiff bases.[20] Formation of the Schiff base via carbonyl activation leads to the cyclization of the imidazole ring via condensation.[17] Although method B of Scheme 1 could be employed for each of the functional groups investigated, preference was placed on its use in the synthesis of amino-substituted analogues to avoid exposing amino-containing benzoic acids to coupling conditions that may result in possible amide polymerization.

The only derivative, which was not synthesised by either method A or B, was the synthesis of **43** (*o*-amino analogue), which was completed by virtue of the formation of the *o*-nitro analogue (**50**) and subsequent reduction to the amino group. This reduction was completed using palladium activated on carbon (5% Pd/C) in MeOH at rt for 16 h. Each analogue was further characterised by ^{13}C NMR and HRMS and were consistent with the desired material.

A total of 18 2-substituted aryl benzimidazole analogues made up the preliminary SAR study against AMA1. The compounds identified from the aforementioned screening libraries that were resynthesised (**51** *m*-OH, **52** *m*-NH_2 and **53** *m*-OMe) were also screened as positive controls to examine the degree of reproducibility of SPR data previously obtained. Screening was performed against the 3D7 *Pf*AMA1$_{[104-442]}$, with a 5-point dose response completed for each analogue. All analogues were screened in duplicates. Overall the SPR data for all 18 benzimidazole compounds was indicative of non-stoichiometric binding owing to resultant analogue binding being characterized by a straight line fit when analysed thus analogues were shown to bind beyond a 1:1

manner (**Figure 3.11**). The R_{max} was anticipated to vary between runs depending on the amount of active protein immobilized on the chip surface, but the overall shape of the data set and K_D for each sample was expected to remain consistent.[10]

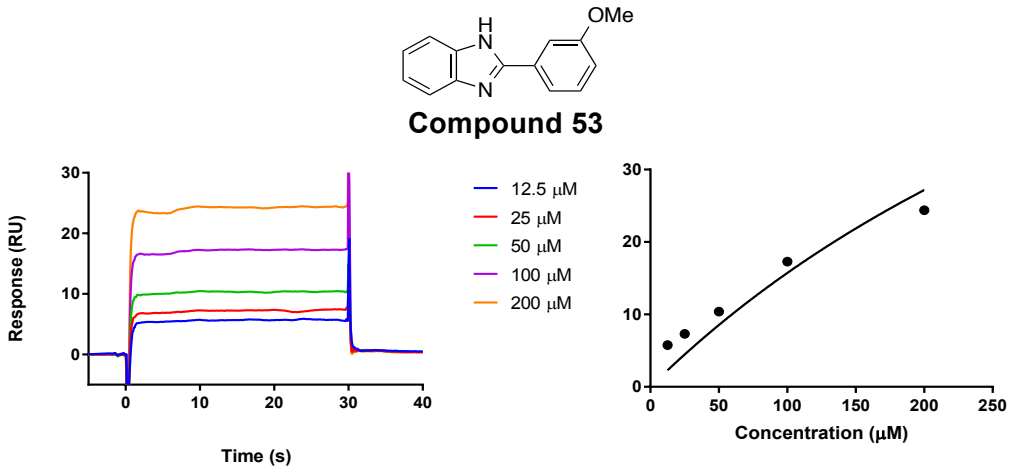


Figure 3.11 SPR sensorgram (left) and subsequent fit (right) of compound **53**, which displays non-stoichiometric binding. The curvature previously identified for SYN-076 is less pronounced for the data set of **53**.

The in-house synthesised 2-(3-methoxyphenyl)-1*H*-benzo[*d*]imidazole analogue **53** (also known as SYN-076) was considered one of the strongest binders to AMA1 (K_D of 120 µM), but a comparison of data between **53** (SYN-076) from the Wellcome Trust Library and the in-house synthesised derivative **53** (**Figure 3.10** and **Figure 3.11** respectively) identified no overlap between the two sets of data in which the resynthesized analogue bound in a non-stoichiometric manner. By using an identical CM5 chip and AMA1 surface for this experiment, the lack of consistent data observed, raised questions over the reproducibility of our SPR assay. The inconsistency in recorded sensorgrams was initially presumed to be the result of degradation of the Wellcome Trust stock sample, although an analysis of the original SYN-076 stock sample via LCMS and HPLC revealed no degradation, with the dominant species demonstrating a m/z of $[M+H]^+$ = 225, which was consistent with the expected result for SYN-076. The combination of non-stoichiometric binding and the inconsistent results when comparing separate runs meant that no reasonable SAR could be assumed for the 2-substituted aryl benzimidazole series in these assays. SPR was subsequently abandoned as the frontline assay for the evaluation of analogues targeting AMA1.

3.1.4 2-Substituted aryl benzimidazole in vitro growth inhibition study

Benzimidazole-derived compounds have been widely investigated as potential antiplasmodial agents.[5,9,21,22] Herein, the previously investigated 2-substituted aryl benzimidazole series was screened using a *P. falciparum* asexual growth inhibition assay to evaluate the analogues anti-malarial activity against transgenic cell lines.

As previously discussed, analogues of the 2-substituted aryl benzimidazole class were shown to bind to both 3D7 and FVO AMA1 via TROSY NMR although analogue affinity could never be determined when screened via SPR and NMR data. In continuing with the investigation of the 2-substituted aryl benzimidazole series, analogues were evaluated in a malarial growth inhibition assay. The most widely used method for measuring parasitaemia within growth inhibition assays is light microscopy but the labour-intensive nature and poor reproducibility of parasitaemia between experimental runs limited its use, particularly when screening compounds for drug discovery and development.[23] For the purpose of this study, flow-cytometry-based growth inhibition was preferred, owing to a more accurate and higher reproducibility for parasitaemia counts between duplicate wells when compared to microscopic-based techniques.[23] Differences in parasitaemia counts between duplicate wells of the same sample have been recorded to be as high as 98% by microscopy.[23] Opting for flow-cytometry would remove any observer error from the analysis of antimalarial activity for novel compounds. This study was based upon the flow-cytometry-based growth assay developed and optimized at the Walter and Eliza Hall Institute of Medical Research and Burnet Institute.[23,24] This flow-cytometry-based approach employed transgenic W2Mef *P. falciparum* engineered to express alleles specific for the AMA1 protein in place of the endogenous W2Mef AMA1 allele.[24] The presence of transgenic W2Mef *P. falciparum* parasites expressing either the 3D7 AMA1 (W2-3D7) or FVO AMA1 (W2-FVO) allele was intended to maintain consistency with previously explored testing platforms such as SPR and NMR that used 3D7 and FVO AMA1. Dr. Damien Drew of the Burnet Institute performed the flow-cytometry-based evaluation of the 2-substituted aryl benzimidazole series; all experimental details are discussed below.

Parasite culture and genotyping

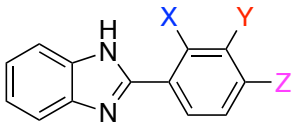
P. falciparum asexual blood stage parasites were maintained in culture in human erythrocytes (blood group type O^+) at a hematocrit of 4% in Roswell Park Memorial Institute medium-2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (RPMI-HEPES) supplemented with 0.25% (w/v) Albumax^{TM} (Invitrogen) and 5% (v/v) heat inactivated human serum. *P. falciparum* were synchronized using sorbitol and heparin treatments as described previously.[24] Transgenic W2Mef *P.*

falciparum strains expressing 3D7 (W2-3D7) and FVO (W2-FVO) AMA1 alleles were generated in previous studies.[24]

P. falciparum growth inhibition assay

All transgenic *P. falciparum* lines were synchronized at the mature pigmented trophozoite stage, and were adjusted to a starting parasitaemia of 0.1% and 2% hematocrit. Compounds to be tested were solubilized in 100% DMSO and adjusted to a common concentration of 1 M Compounds or DMSO controls were then adjusted to alternate molarities via serial dilution in Roswell Park Memorial Institute (RPMI) culture media. Diluted compounds or DMSO controls were added to infected erythrocytes and mixed to generate a final culture volume of 50 μL per well. Parasites were allowed to develop through two cycles of erythrocyte invasion for 72 h at 37 °C.[23,24] Glutaraldehyde (ProSciTech) was then added to the early trophozoite stage parasites after the second round of invasion, in order to fix them for 1 h at r.t (final concentration of 0.25% (v/v)). After fixation, the parasites were washed in human tonicity phosphate-buffered saline (HTPBS), stained with 106 SYBR green dye (Invitrogen) and approximately 50,000 red blood cells counted per well using a BD FACSCantoII flow-cytometer. Fluorescence Activated Cell Sorting (FACS) counts were analyzed using FlowJoTM (Ver 6.4.7) software (Treestar). Each growth inhibition assay (GIA) was run in a 96-well plate format with each compound concentration tested in duplicate wells.

In total, 27 of the 2-substituted aryl benzimidazole analogues, initially investigated as AMA1 hits, were screened for their antiplasmodial activity against transgenic W2Mef *P. falciparum* lines (**Table 3.12**). The initial testing of W2-*Pv*AMA1 wild type in the early stages of the screening campaign displayed comparable potency to transgenic strains (W2-3D7 and W2-FVO) for all analogues tested. Based on these results, this study was modified to focus on the developed transgenic lines, W2-3D7 and W2-FVO, excluding the wild-type from the evaluation of the 2-substituted aryl benzimidazole series. In total, eight different functional groups were investigated within the series; with substituents on the 2'-phenyl ring occupying each of the arene substitution (X, Y, Z).



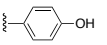
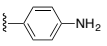
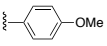
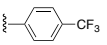
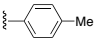
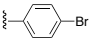
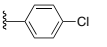
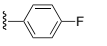
	X	IC_{50} (nM) 3D7	FVO		Y	IC_{50} (nM) 3D7	FVO		Z	IC_{50} (nM) 3D7	FVO
42		62	40	**51**		>50	<31	**61**		40	38
43		>250	>250	**52**		46	46	**62**		23	10
44		250	210	**53**		>50	>50	**63**		140	125
45		>250	>250	**54**		25	<31	**64**		10	<31
46		250	210	**55**		>50	<31	**65**		100	20
47		80	100	**56**		<31	<31	**66**		35	>31
48		>250	<250	**57**		200	175	**67**		>250	250
49		250	175	**58**		>250	>250	**68**		250	175

Table 3.12 Structure−Activity Relationship of Compounds **42-49, 51-58** and **61-68**. Compounds **50**, **59** and **60** are not shown and were insoluble in buffer.

For all 26 compounds tested, the IC_{50} values were consistent for both transgenic lines, with the exception of **65** (*p*-methyl), which displayed a 5-fold difference in potency between W2-3D7 (100 nM) and W2-FVO (20 nM). A comparison of each arene position identified the *ortho*-position in general as being least tolerated for both W2-3D7 and W2-FVO with the exception of the electron-donating hydroxy-substituent **42.** Compound **42** displayed a 4-fold (3D7) and 6-fold (FVO) better potency when compared to the seven other substituents investigated at *o*-position. The introduction of a hydroxy group at the *ortho*-position slightly improved potency but the position remained the least tolerated when compared to the *m*-hydroxy (**51**) and *p*-hydroxy (**61**) analogues, (3D7 IC_{50} >50 nM; and FVO IC_{50} <31.25 nM) and (3D7 IC_{50} 40 nM; and FVO IC_{50} 38 nM) respectively. This

observation could be rationalized by the increase in steric hindrance subsequently restricting the torsional bond angles of adjacent groups of the analogue. The *o*-bromo **47** also displayed slightly better potency akin to **42**, but was still considered the least favourable position investigated for the bromo-substituted analogues. With the exception of the fluoro- and chloro-substituted analogues that displayed IC_{50}'s greater than or equal to 175 nM, all six other substituents functionalized at either the *meta*- or *para*-position of the phenyl ring were considered favourable, in particular analogues **54** (*m*-trifluoromethyl), **61** (*p*-hydroxy), **62** (*p*-amino) and **64** (*p*-trifluoromethyl), which displayed the strongest potency of the series investigated.

Unlike the chloro- and fluoro-groups, the bromo substituent was better tolerated across all three positions investigated. Bromo-substituted analogues showed up to an 8-fold improvement in potency against both W2-3D7 and W2-FVO as observed when comparing **66** (3D7 IC_{50} 35 nM; and FVO IC_{50} <31 nM) to **67** (3D7 IC_{50} >250 nM; and FVO IC_{50} 250 nM) and **68** (3D7 IC_{50} 250 nM; and FVO IC_{50} 175 nM). Based upon atomic radius and STERIMOL parameters, particularly *L*,[25] an increase in substituent steric bulk seemed to correlate with better potency as displayed by the bromo-substituents when compared to the smaller fluoro- and chloro-substituents. The STERIMOL parameter *L* refers to the length of the bond between the parent skeleton (2-aryl benzimidazole scaffold) and the projected van der Waals surface along the bond of attachment.[26] This apparent preference for the bromo-substituent was also mirrored in the relative lipophilicity (described as π) of the substituents in which the addition of a bromo-group (π of 0.86) equating to a higher lipophilicity compared to chloro- (π of 0.71) and fluoro-groups (π of 0.14).[27] A similar trend was observed when comparing the trifluoromethyl- and methyl-substituents, such that the slightly larger trifluoromethyl-substituents of **64** (3D7 IC_{50} 10 nM; and FVO IC_{50} <31 nM) displayed a stronger potency compared to **65** (3D7 IC_{50} 100 nM; and FVO IC_{50} 20 nM). These factors suggest a likely target region involving a mode of action that requires the ability to interact with or cross membranes.

Comparison of the methoxy- and hydroxy-substituted analogues revealed the larger and slightly more lipophilic methoxy-substituent to display weaker potency across all three positions investigated. One explanation could be the hydrogen bond donor capability of the hydroxy-group that could explain the noticeable difference in potency across all arene positions optimized. Although the hydroxy- and methoxy-group are recognized as strong electron-donating groups, the hydrogen bond donor capability of the hydroxy-group allows for the greater potential to form strong intermolecular forces upon binding to target regions. Development of amino-substituted analogues was supported by the preference for electron-donating groups, with **62** (*p*-NH_2) showing the

strongest potency of all 26 analogues (3D7 IC_{50} 23 nM; and FVO IC_{50} 10 nM). Besides the *ortho*-position, the amino-substituted analogues showed on average a better potency for W2-3D7 and W2-FVO than any of the hydroxy-substituted analogues. This observation could relate to the amino-substituent's slightly stronger electron-donating effect across the *m*- and *p*-positions.[27] Although both electron-donating substituents were considered the better performers of the series, **64** (*p*-trifluoromethyl), was also able to display very strong potency against both W2-3D7 and W2-FVO. This data suggested that at the *p*-position there was no preference for electron-donating groups considering the electron-withdrawing substituents *p*-CF_3 (**64**) and *p*-Br (**66**) were also able to translate to a strong potency against both transgenic strains. The interpretation of these results could be influenced by the mode of action(s) attributed to these analogues which is yet to be determined.

Since the nanomolar potency observed for this assay did not match analogue binding affinities for AMA1 as measured by NMR or SPR, it was evident the 2-substituted aryl benzimidazole inhibitory action involved a different pathway other than inhibiting the formation of the MJ during invasion. Having screened all the compounds, the better-performed compounds will be prioritized for investigation in search of the mode of action for the scaffold: this work will be pursued in collaboration with other members of the group.

3.2 Benzimidazole covalent probes

3.2.1 Background

One of the challenges of FBDD is the initial low affinity associated with the majority of fragment leads, resulting in a heavy reliance on structure determination and structure-based design to help improve affinity and selectivity.[28,29] From the structure of a ligand bound to a target protein, valuable information about a ligand-binding site can be acquired, helping to identify an ideal starting point for the rational design of improved fragments. In parallel with conventional fragment hit optimization approaches, other alternative methods capable of developing fragments into higher affinity compounds, has and continues to be investigated within the FBDD field. The search for new methods has seen increasing interest in the design and synthesis of irreversible probes for the purpose of drug discovery.[30,31] A covalently binding modifying compound, also known as an irreversible probe or affinity probe, targets a specific binding site on a receptor or protein, forming a covalent bond at or near the recognition site. Irreversible probes possess a reactive group either described as a chemoreactive or photoreactive warhead (**Figure 3.13**). The addition of these warheads has been commonly employed for providing insight into the structural and functional

features of G protein-coupled receptors (GPCRs).[32-34] Irreversible probes have long been employed for the development of biomarkers, fluorescent probes and hetero- and homobivalent ligands.[35]

Over the years, the design of covalent binders in FBDD has been widely underappreciated owing to concerns over an increased likelihood of potential nonspecific or off-target activity.[36,37] Of late there has been renewed interest in covalently modifying molecules within the FBDD field, spurred on by a number of successful receptor structural studies[38-40] including the recently FDA approved covalent proteasome inhibitor carfilzomib and the covalent kinase inhibitors afatinib, and ibrutinib.[36]

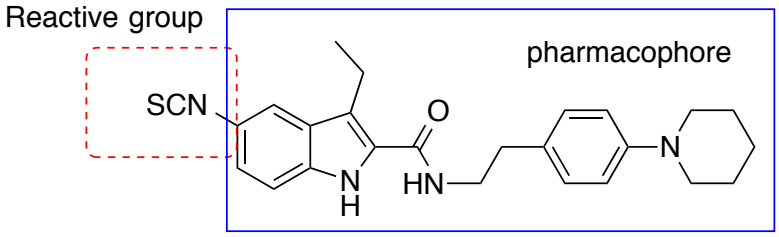


Figure 3.13 A literature example of a covalent and potent cannabinoid 1 (CB1) allosteric modulator.[34]

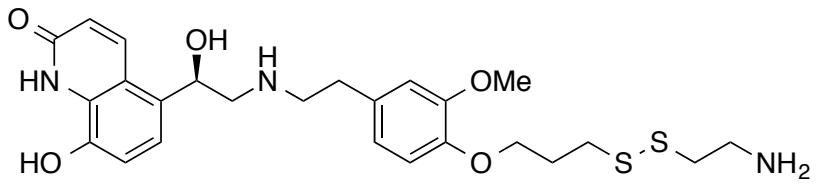
Our interest in this concept focused on the application of reactive groups to low affinity fragments developed based upon MIPS fragment screen results. The aim was to improve our search for higher affinity binders of AMA1 along with the development of covalent binders, which would focus on previously discussed 2-substitutued aryl benzimidazole analogues displaying micromolar binding affinity for the hydrophobic cleft of AMA1. An improvement in both affinity and selectivity of synthesized covalently modified probes was expected to help attain structural information regarding the binding regions of the previously identified AMA1 binders. The addition of reactive functional groups on the benzimidazole scaffold would also help to minimize the need for multiple iterations of synthetic optimization typically required for the conventional FBDD development of higher affinity molecules.

3.2.2 Covalent probes

To date, studies of GPCRs still present a major challenge because of their inherent flexibility and instability when examined during *in vitro* studies.[41] These studies are considered problematic owing to low binding affinity with rapid association and dissociation rates leading to conformational heterogeneity that prevents the formation of diffraction-quality crystals.[41] Over the past few

decades, studies involving diverse bifunctional molecular probes have helped to overcome these obstacles. A bifunctional probe is an antagonist or agonist molecule for a given GPCR, which is connected to a tag or functional group known to exhibit a specific property to enable the quantification and visualization of GPCRs. In particular, radioactive or fluorescent tags are frequently employed to quantify or visualize GPCRs.[40] To further improve the affinity of these bifunctional probes, covalently binding molecular probes were developed. The affinity of the original pharmacophore, the reactivity of the electrophilic moiety, and the subsequent positioning of the reactive moiety on the pharmacophore, determines the overall success of the resultant ligation reaction owing to these probes. The covalent modification of the active site is also dependent on a suitably positioned nucleophilic amino acid residue in order to elicit low nonspecific reactivity when incubated with the respective GPCR target.

One of the major requirements for protein crystallography is the existence of stable and conformationally homogeneous ligand-receptor complexes.[42] The high-affinity and specificity associated with covalent probes has helped to form such complexes, which have recently been shown to be useful tools for structural studies of GPCRs.[41,43] Through the formation of stable ligand-receptor complexes, researchers have been able to investigate the physiological function and prevalence of receptors, their subtypes, and the pharmacological effects mediated by various drugs.[44] Covalent probes have enabled the study of receptor reserve and turnover[45] and facilitated the purification of GPCRs from native tissue, using radioactive labels to track the progress of GPCR purification.[41] Nevertheless, reluctance within the FBDD landscape to employ these techniques remains evident owing to the difficulty associated with identifying the right balance between reactivity and selectivity for covalent binders. The abundance of nucleophilic groups in proteins dictated that new approaches were required in order to avoid low coupling selectivity. The disulfide-based cross-linking approach[43,46] was one of the earliest approaches investigated which takes advantage of the covalent binding of disulfide-containing compounds, chemoselective for cysteine. Through this approach, the ligation of the covalent probe is encouraged using the affinity of the ligand-pharmacophore rather than the electrophilicity of the attached reactive group.[41] The development of the covalent ligand FAUC50 (**Figure 3.14**) is a classic example of the disulfide-based cross-linking approach, forming what is considered to be the first agonist-bound GPCR structure.[43]



FAUC50

Figure 3.14 Structure of the covalent ligand, FAUC50.[43]

3.2.3 Selective reactive groups

A number of factors must be considered prior to the development of a covalently modified molecule. These factors include the electrophilic reactivity required of a specific reactive group to form a covalent bond with a selected amino acid residue,[47] the synthetic viability of introducing a reactive functionality onto the original binding pharmacophore and the intended purpose of installing a reactive moiety to either inhibit or probe the binding site of interest. In order to develop an irreversible binder, the original binding scaffold must be able to tolerate the introduction of a reactive group without incurring a significant loss in pharmacophore affinity or molecular structure.[30] The pre-existing functional groups of the pharmacophore also play a key role in guiding the synthesis of a reactive moiety on the binding scaffold. Linker units such as amides, esters and thioureas are commonly employed as chemical handles owing to their simplicity of synthesis when functionalizing the pharmacophore. These linkers are commonly installed to connect the reactive moiety with the binding scaffold. The reactive moieties used to form a covalent bond can either be described as photoreactive or chemoreactive.

Photoreactive groups

The theory of a covalent molecular probe was inspired by the photoreceptor rhodopsin and its native isomer retinal, 11-*cis*-retinal.[48] This isomer, upon absorption of a photon, isomerizes to all-*trans*-retinal, switching rhodopsin from its inactive form to an active conformation. The concept behind a photoreactive (also known as photoactivatable) irreversible group involves a chemically inert group, which upon photolysis via a specific activator source, is converted into a reactive species with the ability to form a covalent bond with a receptor. Considering that the photoreactive group is inert prior to photolysis, designing photoreactive probes allows for the reactive group to be incorporated at any stage within the biological assay. However, one issue associated with photoreactive group employment is the ability for photoreactive groups to form covalent bonds with amino acid residues of non-nucleophilic nature (unlike chemoreactive groups). To avoid this issue,

the photoactivatable group is incorporated early to ensure the bound inert probe is in equilibrium prior to activation[30] thus allowing the pharmacophore to be positioned more closely to its original binding pose.[44] The selectivity of the original pharmacophore also plays an important role in determining the success of a photoreactive probe binding to a particular target.[44] For common *in vitro* experiments, the existence of a highly selective pharmacophore is not always present thus limiting the prospective application of this technique to certain biological assays. Although recent successful examples of photoreactive irreversible probes are well documented,[49,50] a major drawback is the UV radiation required to activate photoreactive groups is capable of causing irreversible damage to the target and surrounding cells or tissue.[51]

A variety of photoreactive groups have been investigated previously, for example benzophenones, diazirines, and aryl azides (**Table 3.15**). Each of these groups is discussed in further detail, including their most important properties and respective chemical accessibility.

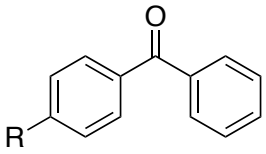
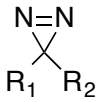
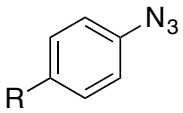
Reactive class	Structure	Photoactivation energy (λ nm)
benzophenone		350-360
diazirines		350-380
aryl azides		< 300

Table 3.15 Commonly employed photoreactive groups and their respective structures and activation wavelengths. Adapted from [30]

Benzophenone

One example of a universally employed photoreactive group is the benzophenone.[52] The benzophenone group is unique in a sense that upon exposure to UV light, a diradical is generated, which is able to form a covalent bond with the receptor.[52] A wavelength of 350-360 nm (**Table 3.15**) is generally used to activate the diradical; this wavelength is considered mild enough to avoid damage to biological material.[53] An advantage of benzophenone as a photoreactive group is the broad range of inexpensive, commercially available, functionalized benzophenones, allowing for synthetic practicability during probe design. Although they are considered chemically more stable then the azides and diazirines, the drawback of adding a benzophenone group is the overall impact on the solubility of the pharmacophore owing to the additional hydrophobicity via the addition of aryl rings. Benzophenones also add bulk and have a propensity to take longer to activate than other photoreactive groups.[54]

Diazirines

Another commonly used photoreactive group is the diazirines, which may be aliphatic or aromatic. A diazirine unit is small in comparison to the benzophenones, and lipophilic.[52] The activation of a diazirine unit using a wavelength of 350-380 nm (**Table 3.15**) affords a highly reactive carbene species that is inserted into an amino acid residue to form a covalent bond with the receptor. A drawback of diazirine activation is the formation of linear diazo isomers that are prone to protonation and thus carbocation formation;[55] these isomers are associated with non-specific binding resulting from the diffusion of molecules from the binding site of the receptor.[30] This diffusion is possible because of the longer lifetime of the isomers when compared to carbenes. While aliphatic diazirines are smaller in size, and less bulky, the photochemistry of the aromatic types is preferred owing to their ability to form more reactive carbenes than the aliphatic type.[55,56] In terms of synthetic feasibility, the aromatic diazirine is labour-intensive, requiring a long multiple step synthesis to install the photoreactive group, whilst the aliphatic diazirine generally requires three steps from commercially available alkyl halides and ketones.

Aryl azides

Aryl azides are widely used photoaffinity labels.[52] Aryl azide, when activated, forms a reactive nitrene species able to form covalent bond with an amino acid residue of the receptor. However, an extensive study of the phenyl azide has pointed out that the moiety rarely gives stoichiometric labeling[57] since upon forming a reactive singlet species via activation, the aryl nitrene is also prone to by-product formation, resulting in ring-expansion to form a ketenimine by-product.[58] This compound effectively reacts with available nucleophiles, although its long lifetime and stability is not conducive for rapid reaction thus resulting in lower cross-linking yields.[52] Aryl azides are

readily synthesized via anilines employing a variety of available procedures[59] but the disadvantage of azide-containing probes is that they require activation at wavelengths < 300 nm (**Table 3.15**); such a high energy is generally associated with an increased risk of damage to the receptor and/or the test system (e.g., cells, tissue).

Chemoreactive groups

Chemoreactive probes are defined by the presence of a reactive electrophilic functionality that is able to covalently bond with a nucleophilic amino acid residue of a receptor. Chemoreactive probes can be applied to both *in vivo* and *in vitro* experiments in circumstances where a pharmacophore possesses an established affinity for a target receptor.[47] The affinity of the original pharmacophore is thereby considered a controlling factor for non-specific binding. In contrast to the installation of photoreactive groups, chemoreactive groups are installed as late as possible in the synthetic scheme in order to avoid cross-reactivity with either the pharmacophore itself or other reagents employed for chemical synthesis. As with photoreactive moieties, a variety of chemoreactive moieties can be incorporated depending on the target amino acid residue in the binding site. The availability of structural information on the target site plays an important role in the selection of the chemoreactive moiety best suited to bind a receptor, although for many studies, structural information is not accessible, resulting in an experimentally determined chemoreactive selection based on a trial-and-error process.[30]

Although electrophilic groups are considered more synthetically accessible, the need to fine-tune the reactivity of the electrophilic group in order to limit non-specific binding is one drawback of the application of chemoreactive groups when compared to the photoactivatable groups.[47] Numerous electrophilic chemoreactive groups have been utilized in the past, including disulfides, Michael acceptors, alkylating agents, nitrogen mustards and isothiocyanates (**Table 3.16**).

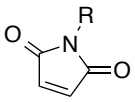
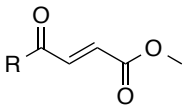
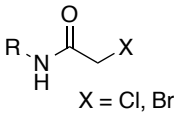
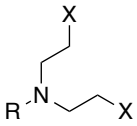
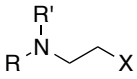
Reactive class	Functionality	Structure	Amino acid residues
Isothiocyanates/ thiocyanates/ disulfides	isothiocyanates	R–N=C=S	cysteine lysine
	thiocyanates	R–S–C≡N	cysteine
Michael acceptors	maleimides		cysteine lysine histidine
	fumarates		
Alkylating agents	haloacetamides	 X = Cl, Br	lysine serine histidine threonine
Nitrogen mustards	nitrogen mustards	 X = Cl, Br R' = alkyl chain	lysine glutamate aspartate cysteine
	2-haloalkylamines		

Table 3.16 An overview of the chemoreactive groups sorted by reactivity class displaying preferred amino acid counterparts. A list of functional classes of each chemoreactive group is also shown, including their chemical structure. Adapted from [30].

Disulfides

The reactive thiols, particularly a disulfide-based strategy, are considered one of the earliest electrophilic chemoreactive classes described.[60] The desirable property of this chemoreactive group is its selectivity for the sulfhydryl function of a proximate cysteine residue on the receptor, thus minimizing non-specific binding. If the activated thiol is positioned in close proximity to a cysteine anchor, the enhanced local concentration is the driving force for the conjugation of ligand and receptor. This positioning is largely dependent on an ideal linker length which is critical to cross-linking since the linker would be required to place the disulfide in a position in which binding

would be possible. The introduction of a disulfide moiety within a pharmacophore can be completed via multiple ways using halomethyl ketones. One example is the substitution of a halogen atom with potassium thiocyanate to afford the final thiocyanate product. This strategy however is somewhat limited by its preference for cysteine residues, which has led to the investigation of other electrophiles that bind a wider range of amino acid groups.

Michael acceptors

Substituted or unsubstituted acrylamides, fumarates and related *α,β*-unsaturated carbonyl-derived functions are all referred to as Michael acceptors. These electrophilic groups undergo a Michael-type addition with preferential binding to lysine or histidine residues.[61] Michael acceptors have also been observed to react with cysteine. When in solution, Michael acceptors tend to react slowly with nucleophiles thereby relying heavily on the affinity of the pharmacophore for the target receptor to drive specific covalent bond formation.[42] The maleimide and fumarate functionalities are considered popular choices owing to their chemical accessibility. The synthesis of these functionalities generally occurs via an amine moiety located on the respective pharmacophore. In particular, the maleimide functionality is converted directly from an amino functionality where as the fumarates initially require the formation of an amide bond prior to being synthesized.

Alkylating agents

Another chemoreactive group is the alkylating agents, particularly the haloacetamides (also described as halomethyl ketones). The more reactive bromo- and iodo-haloacetamides react with a variety of amino acid residues including serine, histidine, lysine and threonine.[42] Haloacetamides are readily available as carboxylic acids or acid chlorides; these reagents require a one-step procedure to be incorporated via a primary or secondary amine. Although the electrophilic functionalities are useful for targeting a variety of different amino acids, the ability to limit the amount of non-specific binding depends largely on the selectivity of the original pharmacophore for the target site or residue.

Nitrogen mustards and 2-haloalkylamines

Another important and unique cross-linking class is the nitrogen mustard analogues or 2-haloalkylamines. The electrophilic reactivity of this class results from the formation of the corresponding aziridinium ion, afforded during the intramolecular cyclization of the group in aqueous solution.[42] Although the mustards can be considered as subtypes of alkylating agents, mechanistically they are considered different from the nucleophilic substitution reactions associated with haloacetamide groups. For the mustard, the halogen atom determines the rate of formation for the aziridinium ion; the bromo-containing mustard is more reactive than the chloro-derivative.[62] The formation of the aziridinium ion makes mustards highly reactive, possessing the ability to react

with amines, alcohols, sulfides and carboxylates, which enables mustards to target a variety of amino acids such as lysine, glutamate, aspartate and cysteine.[63] Equally, the mustard class is considerably unstable and easily prone to degradation once incorporated into the pharmacophore. A major drawback of this reactive class is the high incidence of non-specific binding owing to the large variety of amino acids able to react with the aziridinium ion.

Isothiocyanates

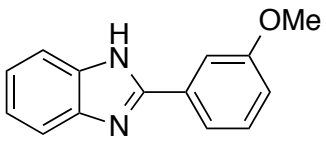
The final chemoreactive group is the isothiocyanates; this reactive class has been incorporated in irreversible ligands targeting cysteine and lysine.[32,64] The isothiocyanates are considered the smallest of the electrophilic groups employed. In addition, this group is also considered the most synthetically accessible, since it can be prepared from primary amines in one-step.[65] The conversion of alkyl or aryl amines generally is completed using carbon disulfide, thiophosgene or 'thiocarbonyl transfer' reagents such as thiocarbonylditriazole and thiocarbonyldiimidazole (TCDI). The incorporation of an isothiocyanate moiety has led to the development of electrophilic probes for a variety of receptors including muscarinic,[66] melatonin[67] and vasopressin receptors.[68]

In comparison to photoreactive classes, the vast range of chemoreactive binding molecules available for cross-linking makes them a valuable asset for both *in vivo* and *in vitro* experiments. The relatively straightforward synthetic accessibility of most chemoreactive groups and their ability to target specific amino acids has turned the tide on their feasibility in FBDD. The common presumption of the requirement of nanomolar affinity to guarantee successful cross-linking has slowly began to diminish within the FBDD landscape.

3.3 Design of covalently binding molecules as AMA1 probes

For the AMA1 FBDD study, a lack of structural information to guide fragment development has led to new approaches being considered in order to improve AMA1 hit affinity and selectivity. The concept of adding a reactive, covalently modifying functionality to previously identified AMA1 binding scaffolds was one approach investigated in order to probe potentially important binding sites of AMA1 previously identified to support analogue binding.[15] The addition of a covalently modifying functionality would help improve the affinity and specificity of previously investigated analogues with the overall outcome intended to map binding regions of AMA1 via X-ray crystallography. Various reactive functionalities were incorporated into a variety of 2-substituted aryl benzimidazole analogues that were rationally progressed owing to encouraging 2D [^{1}H-^{15}N] TROSY NMR perturbation data, acquired during a study by Krishnarjuna *et al.*[15] This study was able to recognize a dominant perturbation region in the conserved region of hydrophobic cleft of

AMA1 characterized by a four amino acid peptide sequence, 'KDGG', in which both **51** and **52** of the 2-substituted aryl benzimidazole series figured predominantly. To take advantage of this limited structural data available, the strongest viable binder of the benzimidazole series identified via SPR, **53** (*m*-OMe) (**Figure 3.17**) was considered the best analogue for the development of covalently modifying molecules designed to target the conserved 'KDGG' region of AMA1.



53
K_D = 120 μM
LE = 0.31

Figure 3.17 Strongest observed AMA1 binder via SPR screening.

Three key issues were considered when designing covalent probes based on **53**: (a) the target, (b) reactive group suitability based upon the pharmacophore and amino acid counterpart and (c) the synthetic accessibility of the reactive group. The possible damage to AMA1 from UV radiation and better suitability of chemoreactive probes for *in vitro* experiments favoured chemoreactive cross-linking classes. The suggested target region for this experiment was the KDGG sequence and the surrounding amino acids within 8 Å of Lys177. In total, 14 different amino acid types occur within this radius, of which cysteine (C), lysine (K), histidine (H), serine (S) and threonine (T) all stood out for preferential targeting owing to the synthetic accessibility of their electrophilic counterpart. Considering this, lysine had previously been perturbed by both **51** and **52** via 2D [^{1}H-^{15}N] TROSY NMR (Refer to Appendix A, section A.3)[15] which meant all reactive groups possessing preferential reactivity for lysine were investigated first. By introducing a variety of electrophilic functionalities to target amino acids within this region, it was possible to introduce variable reactivity for key residues to gauge the extent of non-specific binding that may occur during incubation period. Reactive classes such as isothiocyanates, Michael acceptors, in particular the maleimide functionality, and the haloacetamide alkylating agents, were all investigated owing to their preference for a lysine nucleophile. In addition to the selection of these reactive groups, the positioning of the reactive functional group on the pharmacophore required careful consideration owing to the lack of structural information on the binding mode of the scaffold. The positions, denoted by R_1 and R_2 (**Figure 3.18**), occupying site 1 and 2 respectively were of particular interest; each position was selected owing to their synthetic accessibility considering the reactive ligand

would need to be installed at the later stages of the synthetic pathway to avoid any cross-reactivity with other parts of the molecule or reaction reagents.

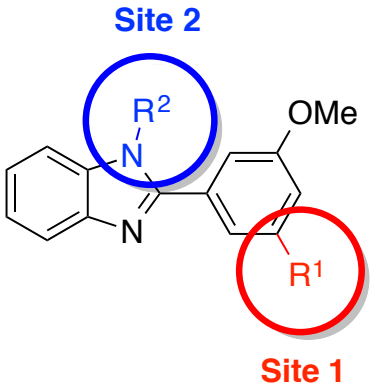


Figure 3.18 Proposed addition of reactive, covalently modifying functional groups to **53**.

Each position was also selected with the intention of avoiding key structural changes to the original pharmacophore that may affect the biological activity of the compound. The water solubility and CLogP of the covalent inhibitor was also considered during the development of covalent probes. Prior to the installation of an electrophilic functionality, elaboration at either sites 1 and 2 was assessed via the synthesis of a variety of benzimidazole analogues based upon **53**, maintaining the *m*-methoxy functionality to determine whether the installation of reactive groups at each proposed site would be tolerated. This was important since all benzimidazole analogues synthesized previously had only minor structural changes resulting from the mono-substitution of the 2'-aryl ring system.

The investigation of both sites resulted in the development of 16 analogues based upon **53** (**Figure 3.19**). The purpose of installing a reactive group was to improve the binding affinity of a pharmacophore whilst attempting to maintain the normal binding mode of the pharmacophore; to do this site 1 was ideal as the starting point. The optimization of site 1 focused on whether functionalization of the aryl ring whilst maintaining the *m*-OMe group would be tolerated. All additional optimization of the aryl group involved the combination of small functional groups to limit the amount of steric bulk added to the benzimidazole scaffold. In total, 8 analogues were synthesized to investigate site 1 using STD NMR (**Figure 3.19**). All analogues excluding **72** were synthesized using Method B described in Scheme 1 (Refer to section 3.1.3); the synthesis of **72** was accomplished via the reaction of acetic anhydride with **70** to install the acetyl-protecting group. When screened, 7 of the 8 analogues maintained AMA1 binding, particularly **70**, which contained an amino functionality at site 1. Analogue **76** displayed poor solubility and could not be tested via

STD NMR. The tolerance of an amino moiety at site 1 (**70**) was a key finding since functionalization of the primary amine would allow for incorporation of each of the chemoreactive groups selected to target lysine.[69] Functionalization of the benzene ring was also proposed as a potential site for the incorporation of a cross-linking moiety, but the poor solubility displayed by **76** during NMR screening suggested that this site was problematic and therefore it was abandoned.

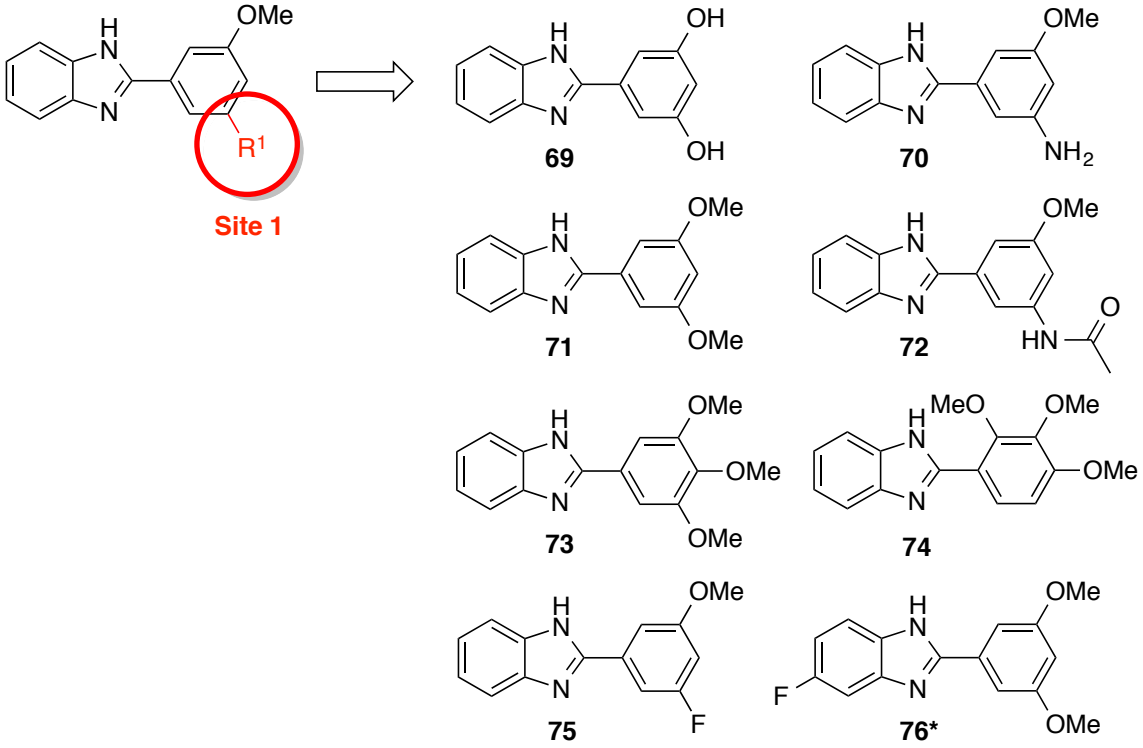


Figure 3.19 Analogues developed to investigate site 1 of the benzimidazole scaffold. *Compound **76** was insoluble in NMR buffer.

The strategy implemented for the investigation of site 2 involved the *N*-alkylation of elaborated benzimidazole analogues based upon **53** to determine the extent of probable chemical space available at this position. The combined optimization of site 1 and 2 was completed to improve solubility of benzimidazoles through reduction in the crystal packing forces and planarity of the ring system. Each alkylated benzimidazole analogue was synthesized via the addition of an appropriate alkylation reagent in methyl iodide, ethyl iodide or 1-chloro-3-iodopropane (used for analogue **84**) to K_2CO_3 in DMF and stirred for 16 h at rt (**Figure 3.20**). In total, 9 alkylated benzimidazole analogues were synthesized and subsequently screened via STD NMR. All analogues bound to AMA1 by STD NMR, particularly **84**, which contained an alkylated propyl chain. Since **84** was able to maintain the original affinity of the pharmacophore, this result

suggested that the surrounding chemical space at site 2 would tolerate further functionalization thereby conducive to the proposed addition of a reactive group at site 2.

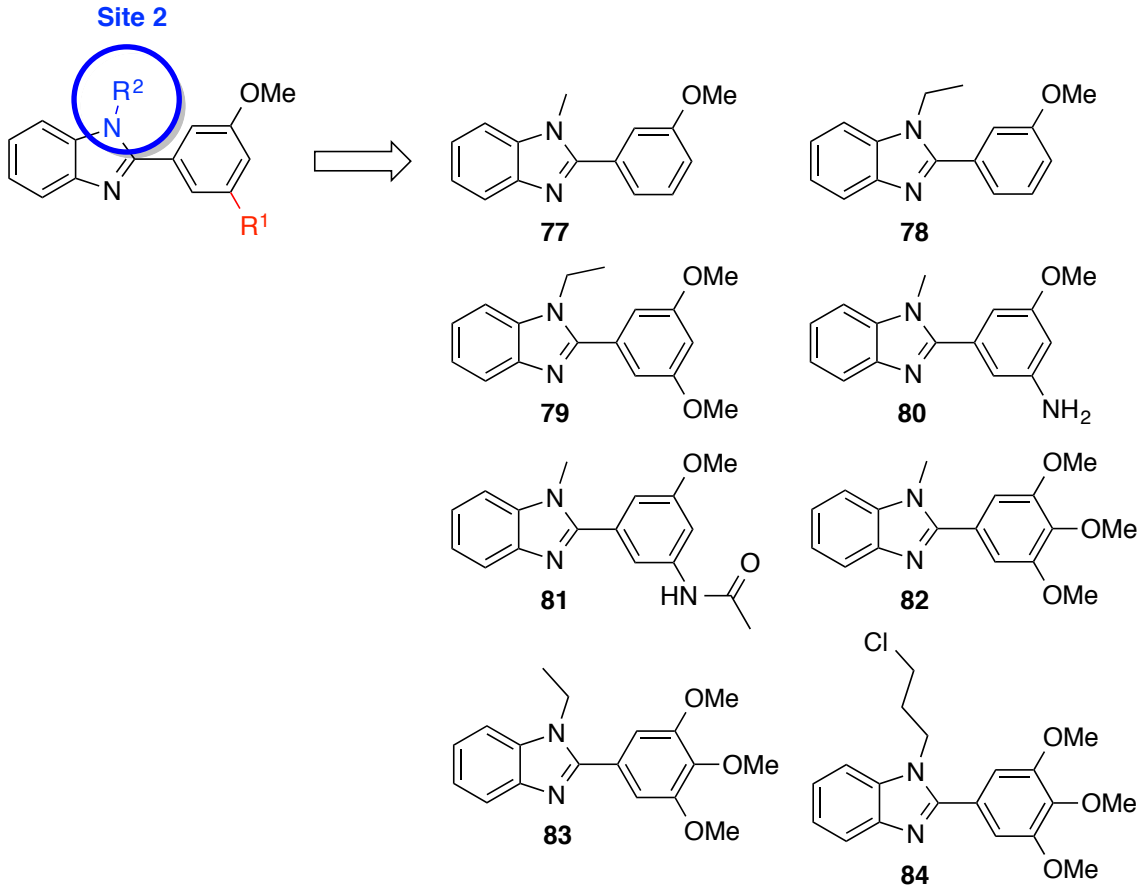


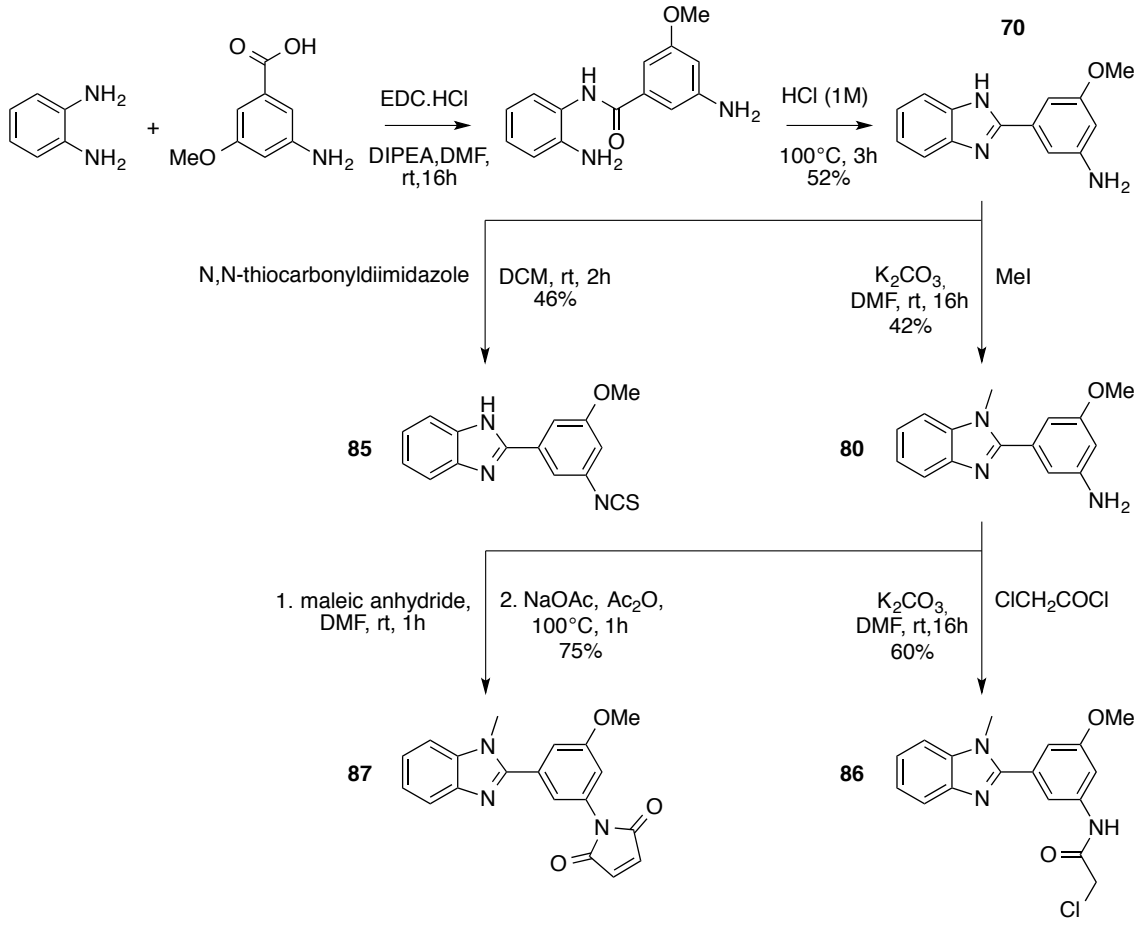
Figure 3.20 Analogues developed to investigate site 2 of the benzimidazole scaffold. Analogues **77-84** displayed a combination of optimization at both site 1 and 2.

3.3.1 Synthesis of covalently modifying benzimidazoles

Elaboration of site 1

The incorporation of a reactive cross-linking moieties at site 1 began with the synthesis of the key intermediate, 3-(1*H*-benzo[*d*]imidazol-2-yl)-5-methoxyaniline **70** from a modified literature procedure (**Scheme 2**).[16,17] Commercially available *o*-phenylenediamine and 3-amino-5-methoxybenzoic acid were coupled together using EDC.HCl and DIPEA in DMF at rt for 16 h. Prior to the cyclization step, the crude material was analyzed via LCMS to verify the presence of desired product of $[M+H]^+$ = 258. The afforded coupled intermediate was extracted in EtOAc, reduced *in vacuo* and subsequently re-solubilized and heated in acid (1 M HCl), allowing for imidazole ring formation via cyclization to give **70**. Reaction time was shortened upon the use of stronger acid (HCl), compared to the use of glacial acetic acid, which was previously to synthesize

2-substituted aryl benzimidazole analogues. The presence of an amino moiety at site 1 provided the necessary chemical handle for the incorporation of an electrophilic cross-linking functionality. Of the planned reactive classes to be installed, the isothiocyanate functionality was initially synthesized upon addition of *N*,*N*-thiocarbonyldiimidazole to a solution of **70** dissolved in DCM that was allowed to stir at rt for 2 h. The synthesis of **85** proceeded in moderate yield (46%) that was synthesized to target lysine residues. The isothiocyanates form covalent bonds by undergoing addition reactions with amines and thiols, but display lower reactivity toward alcohol groups in amino acids like S and T.[70]



Scheme 2 Synthesis of covalently modifying derivatives via amino moiety at site 1.

To avoid previously encountered solubility issues with the benzimidazole scaffold, the *N*-methylation of the benzimidazole ring of **70** was completed prior to installing the reactive group in order to improve the overall analogue solubility without affecting pharmacophore affinity. The *N*-methylation of the benzimidazole ring was completed using methyl iodide and K_2CO_3 in DMF at rt for 16 h allowing the nucleophilic substitution to occur and thereby forming the desired methylated

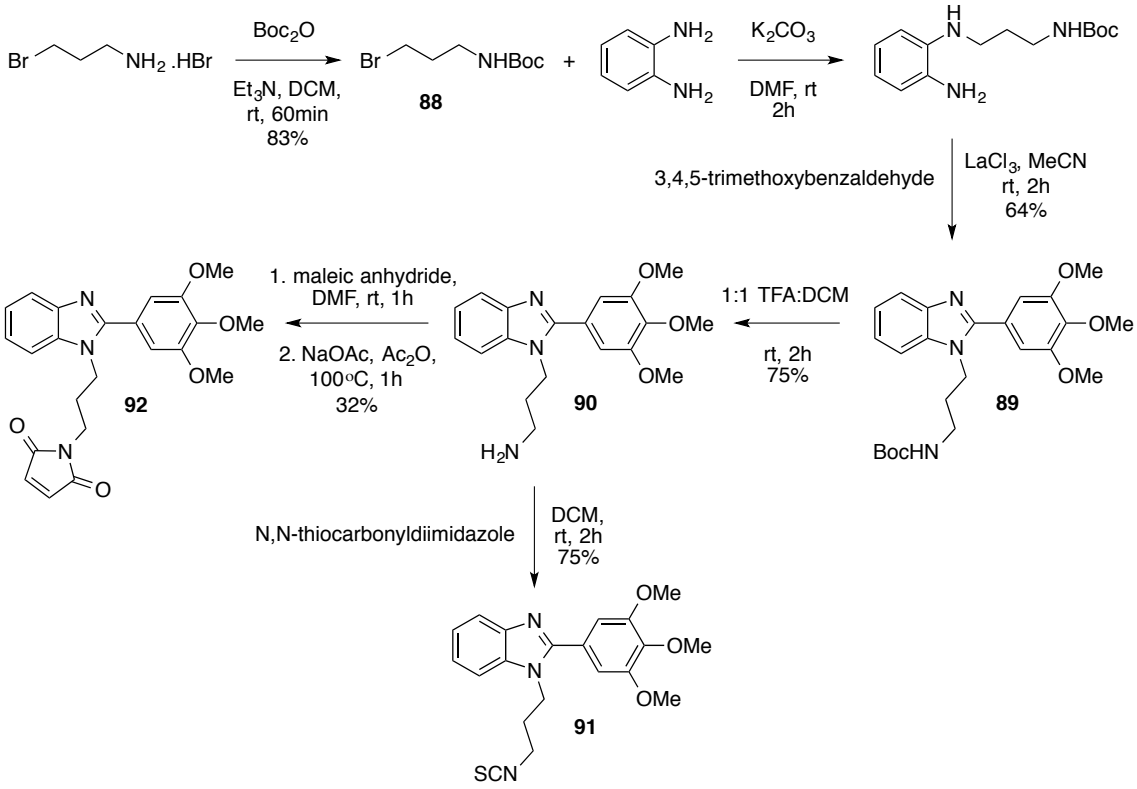
analogue (**80**) in moderate yield after purification via column chromatography (42%) (**Scheme 2**). An improvement in analogue solubility was critical since all protein-ligand-based experiments were performed in aqueous buffer. ^{1}H NMR spectra confirmed preferential *N*-methylation of the benzimidazole ring via a comparison of the NMR spectra for analogues **70** and **80**. This conversion was confirmed by the lost of proton signal at 12.66 ppm for **70**, characteristic of the N-H ^{1}H NMR chemical shift of the benzimidazole ring.

With the *N*-methyl substituent in place, two new covalently modifying probes were developed from **80**, displaying two differentially reactive electrophiles to target lysine residues (**Scheme 2**). Similar reaction conditions to those used in the synthesis of **80** were employed for the synthesis of **86**, with the installation of the reactive chloroacetamide electrophile via the amino-substituent at site 1. Alkylation using chloroacetyl chloride required the addition of K_2CO_3 in DMF to afford **86**. Once purified, the presence of **86** was confirmed by ^{1}H NMR, which showed a loss of the ^{1}H-proton signal at 5.35 ppm, which was reminiscent of the alkylation of the aromatic amine. Dialkylation of the amine moiety was avoided by completing this reaction at rt using a weak base in K_2CO_3. Compound **86** was the first of the alkylating agent compound class to be synthesized. The synthesis of the final covalent probe based on **80** involved the two-step reaction to form the maleimide group, which would be the first analogue from the Michael acceptor class. The first step involved the addition of maleic anhydride to a solution of **80** in DMF, left to stir at rt for 1 h to form the uncyclized *α,β*-unsaturated acid intermediate. After 1 h, the reaction mixture was reduced *in vacuo* and solubilized. This mixture was heated to 100°C for 1 h and upon completetion the mixture was neutralized using sodium bicarbonate solution, extracted in DCM and purified via column chromatography to afford **87** in good yield (75%). Similarly to the haloacetamide, maleimides form covalent bonds with cysteine, but occasionally react with lysine and histidine amino acid residues as well.[61] Nucleophilic amino acid residues are able to add across the maleimide double bond via a Michael addition-type reaction. Other larger Michael acceptor functionalities were not considered since previous benzimidazole analogues containing an amide functionality on the 2'-aryl ring displayed little to no binding, owing to poor solubility as observed during the Wellcome Trust screening campaign (Refer to section 3.1.2). For this purpose, the maleimide group was considered a better electrophilic group compared to other Michael acceptor functionalities like fumarates and bicyclo[2.2.2]octa-2,5-dienes which are incorporated via an amide bond formation.

Elaboration of site 2

The increased functionalization of the aryl ring was able to maintain binding when screened via R1-competition NMR as demonstrated by the trimethoxysubstituted analogue **73**. Analogue **73** (**Figure 3.19**) was also observed to display a strong STD NMR signal that was comparable to the original pharmacophore **53** (Refer to Experimental, section 5.4 for STD spectra). In order to limit potential solubility issues that may arise during testing, analogue **73** was the preferred analogue for the reactive group studies at site 2 as it was considered the most soluble (highest CLogP) of the three methoxy-substituted analogues (**53**, **71** and **73**) of interest for this study. Based upon STD NMR results recorded for **84** (**Figure 3.20**), the propyl chain was selected as an ideal linker length for the proposed installment of an aliphatic amino moiety (Refer to Experimental, section 5.4 for STD spectra). With the propyl substituent in place, the benzimidazole analogue was able to display a comparable STD signal to **53** and additionally its ease of synthesis was conducive for a less labour-intensive synthetic scheme in order to access the necessary primary amino moiety for reactive group formation.

Prior to cyclization to form the benzimidazole ring (**Scheme 3**), the mono-*N*-alkylation of *o*-phenylenediamine was performed to install a Boc-protected propyl amine. The alkylating reagent used, 3-bromopropylamine, was initially de-protected using triethylamine (Et_3N) in DCM since it was commercially available as the HBr salt, and reacted on with the addition of di-*tert*-butyl dicarbonate to afford the Boc-protected product **88**. This was subsequently reacted with *o*-phenylenediamine in DMF at rt for 2 h, giving the alkylated product. Upon isolation of the alkylated product, the reaction was carried forward to the next step without further purification. Purity was determined from UV trace analysis obtained from LCMS analysis, which also confirmed the synthesis of desired material ($[M+H]^+ = 266$). It was imperative that mild conditions were used for this reaction along with adding the alkylating agent drop-wise to the solution to ensure selective reactivity occurred to avoid the dialkylation of *o*-phenylenediamine. Later we discovered that even with the presence of the minor dialkylated product, no purification was required owing to cleavage of the second Boc protected propyl amine during *in situ* cyclization giving **89**. With reference to the work completed for the synthesis of the 2-substituted aryl benzimidazole series (Refer to section 3.1.3), the Lewis acid, $LaCl_3$ was employed in catalytic amounts to help promote ring formation, helping to form the *N*-alkylated benzimidazole upon addition of 3,4,5-trimethoxybenzaldehyde in MeCN at rt for 2 h. This reaction was able to proceed in moderate yield (64%) giving **89**.



Scheme 3 Synthesis of covalently modifying derivatives via *N*-alkylation.

To synthesize the covalent probes, Boc-deprotection of **89** was completed in a 1:1 mixture of TFA: DCM, stirred at rt for 2 h or until 100% converted (**Scheme 3**). The key intermediate **90** was obtained in good yield (75%). In an attempt to optimize the synthesis of key intermediate **90**, direct alkylation of the trimethoxy-substituted benzimidazole was proposed using 3-bromoproplyamine with K_2CO_3 or a stronger base in sodium hydride (NaH), however both attempts failed to afford any desired material. The inability to directly alkylate the scaffold post cyclization via traditional alkylation methods helped to recognize the adapted alkylation of *o*-phenylenediamine as the more effective procedure to access the key intermediate **90**. Two covalent derivatives of **90** were synthesized, employing electrophiles known to target lysine residues. The first probe focused on the installation of an isothiocyanate warhead; this was synthesized via the addition of *N*,*N*-thiocarbonyldiimidazole to a solution of **90** dissolved in DCM that was allowed to react at rt for 2 h, affording **91** in good yield (75%). The other covalent probe involved the formation of a maleimide warhead over a two-step reaction procedure. The initial step included the formation of the uncyclized *α,β*-unsaturated acid intermediate upon addition of maleic anhydride to a solution of **90** in DMF at rt for 1 h. Once completed, the resultant acid intermediate was taken up along with sodium acetate in acetic anhydride, allowing for the ring closure of the acid intermediate to form the

covalently modifying probe **92**.

3.3.2 Ligand binding study

Method

To test the reactivity profile of each probe, **85**, **86**, **87**, **91** and **92** (**Figure 3.21**) were individually incubated with AMA1 in aqueous buffer and monitored by MALDI-TOF mass spectrometry. These experiments were conducted to identify conditions conducive for the formation of protein-ligand adducts in regions of interest to the project. MALDI-TOF MS is an ionization technique that uses laser energy to absorb matrix (compound of crystallized molecules) to create ions from large molecules with minimal fragmentation. The matrix acts like a buffer between the sample and the laser helping to ionize the sample to carry along the flight tube so it can be detected. To generate the required $[M+H]^+$ ions for the large molecules presented, a counter ion source in trifluoroacetic acid (TFA) is added to the matrix and used during testing. The MALDI analysis was performed in collaboration with Taryn Guinan, a Research Fellow at Monash University. Mansura Akter, a PhD candidate within the Norton research group at Monash University, provided all 3D7 *Pf*AMA1$_{[104-438]}$ (DI+II) required for the completion of this study.

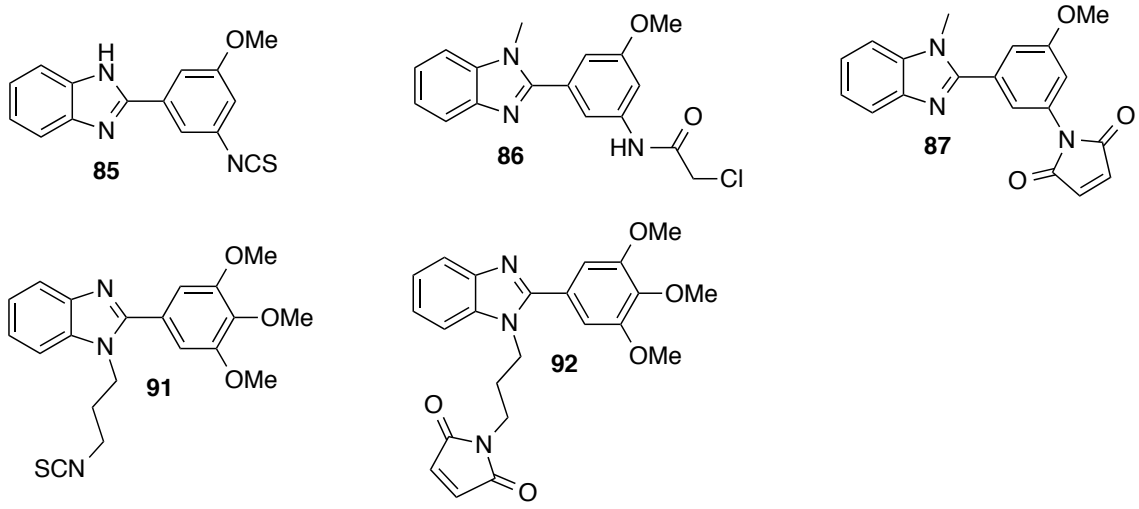


Figure 3.21 Covalently modifying probes investigated via MALDI-TOF mass spectrometry.

Unlabeled 3D7 *Pf*AMA1$_{[104-438]}$ (DI+II) (M_r = 41337.1 Da) was employed to evaluate the reactivity of synthesized covalent probes targeting the lysine residue located within the KDGG-binding region. Prior to the incubation period, an initial 3D7 AMA1 stock (M_r = 41337.1 Da) of 24 µM (1mg/mL) was prepared in PBS buffer **(**pH 7.4**)**. A single stock of the protein sample was favored to minimize concentration variability between prepared samples. All synthesized probes

were initially made up to 200 mM in 100% DMSO; stock solutions were further diluted in DMSO when required for the preparation of final reaction samples. Each ligation reaction was performed in an Eppendorf vial, made up to a final volume of 100 µL of which a 5% total volume of DMSO was observed for all final reaction samples. The lack of a well-defined target on the AMA1 surface led to a concentration-dependent investigation for selected probes to assess the amount of probe binding that may occur at different concentrations. All prepared reaction samples were mixed at rt for up to 10 h using an Eppendorf Thermomixer R Mixer, 1.5 mL Block at 1000-1200 rpm.

In preparation, for MALDI analysis, each reaction mixture required desalting to remove all DMSO and dissolved buffer salts from reaction samples. The removal of buffer salts would help improve the formation of MALDI crystals during MALDI plate preparation for analysis. The desalting column **(**PD SpinTrap G-25 containing 700 µl slurry of Sephadex G-25 medium) was pre-equilibrated, allowing for a buffer exchange using ammonium bicarbonate solution (50 mM $(NH_4)HCO_3$). Once equilibrated, the sample was added to the column and spun to elute desalted protein material, which was subsequently lyophilized in preparation for testing. All reactions were monitored by mass spectroscopy using a Shimadzu 7090 MALDI-TOF-TOF mass spectrometer (Shimadzu, UK) equipped with a 2 kHz UV laser. Protein analysis was performed in linear mode in the range *m/z* 20000-70000 and laser power was user optimised. Protein samples were prepared using sinapinic acid (Sigma, Australia) at 10 mg/mL in 1:1 acetonitrile: 0.1% trifluoroacetic acid. Bovine serum albumin (BSA) was used as the calibrant. The preparation of the MALDI plate involved the suspension of lyophilized sample in a mixture of MeCN: H_2O, a ratio of between 30:70-70:30 MeCN: H_2O was considered acceptable for MALDI experiments. The ratio of 30:70 (MeCN: H_2O) was predominately used for all samples. The dried-droplet crystallization method was used to give the best results during sample preparation.

Results and Discussion

All five electrophilic cross-linking probes synthesized were investigated against 3D7 AMA1 using MALDI-TOF MS. This study was designed to measure the extent of the change in molecular mass (m/z) resulting from the incubation of selected covalent modifying probes with 3D7 AMA1. The concentration of 3D7 AMA1 used was 1 mg/mL, which was kept constant for all experiments. The variable reactivity of each covalent probe was also investigated to determine ideal reaction conditions favourable for the formation of a 1:1 protein-ligand complex. A lack of specific experimental literature describing covalent probe reactivity for protein models lacking specific binding regions resulted in the completion of a concentration-dependent study focusing initially on

the most reactive covalent probe to determine the extent of probe binding. The haloacetamide reactive probe **86** (**Figure 3.21**) was chosen owing to the reactive class's ability to bind with a larger range of amino acid residues, including lysine, in comparison to the isothiocyanate or Michael acceptor classes. This investigation involved the development of multiple mixtures of ligand **86** incubated with 3D7 AMA1, described as 1:1, 5:1 and 10:1 respectively, to identify how effective probe reactivity would at each concentration tested.

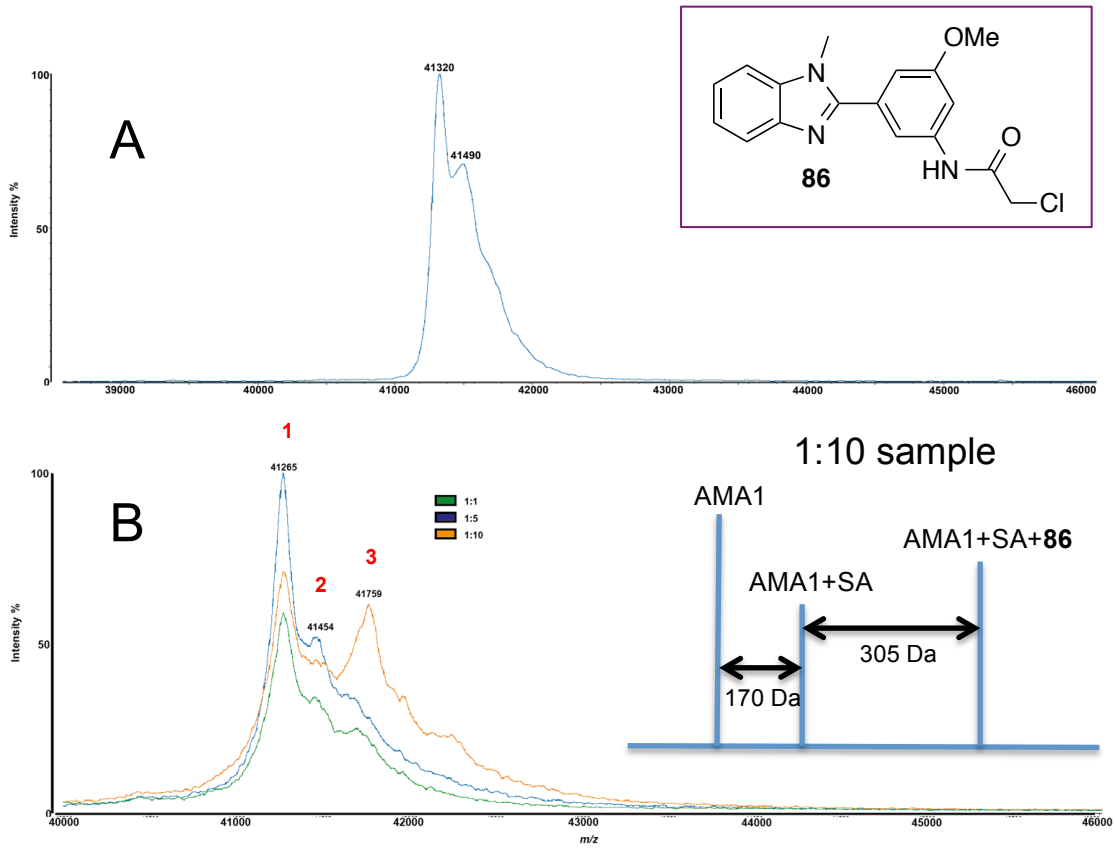


Figure 3.22 MALDI mass spectra of (A) Purified 3D7 *Pf*AMA1$_{[104-438]}$ (DI+II) (B) Covalent probe **86** tested at three different concentrations against 3D7 AMA1**.** Resultant concentration-dependent profiles were overlaid.

By completing the initial analysis of the reactive warhead **86** (**Figure 3.22**) it was possible to modify future experiments in search of 1:1 binding interactions; the identification of 1:1 binding was critical for future X-ray crystallography studies. This initial investigation of **86** also helped improve the efficiency of sample preparation whilst also limit unnecessary consumption of AMA1 protein. As shown in Figure 3.22 A, the peak cluster at 41320 Da was in close agreement with the calculated mass of 3D7 AMA1 (41337 Da). The estimated error for all *m/z* measurements was *m/z* ± 20 Da. Also shown in Figure 3.22 A was the presence of a shoulder peak cluster at 41490 Da,

suggested to result from the formation of a sinapinic acid (SA) adduct (AMA1+SA), which could account for a peak shift of 170 Da.[71] In relation to the reactivity of **86**, Figure 3.22 B represented all data acquired for the concentration-dependent investigation of **86**. The peak cluster observed at 41265 Da for the 1:1 sample (green) was shown to be in agreement with the calculated mass of 3D7 AMA1 as displayed by Figure 3.22 A. It was suspected ligand-induced dissociation (LID) during fragmentation could be responsible for the observed variation in the mass observed when compared to the AMA1-only sample.[72] For all concentration-dependent experiments performed, the identified protein peak at 41265 Da (depicted by peak 1 of **Figure 3.22 B**) was used as the calibration peak for all subsequent experiments. A review of the 1:1 (green) and 1:5 (blue) spectra involving **86** identified no obvious peak shift, suggesting that a higher ligand concentration is required to facilitate the proposed covalent modification of AMA1. For both the 1:1 and 1:5 samples, the formation of the AMA1+SA adduct was observed, depicted as peak 2 for Figure 3.22 B. The identification of the AMA1+SA adduct in the protein-only sample (**Figure 3.22 A**) implied the formation of the AMA1+SA complex was not a ligand-induced affect.

Analysis of the 1:10 (orange) spectra of Figure 3.22 B confirmed the covalent modification of 3D7 AMA1; the resultant mixture was able to form a new peak at 41759 Da (depicted as peak 3) thus exhibiting a total peak shift of 493 Da from peak 1 of Figure 3.22 B. To account for the difference between peak 1 and 3 observed, upon covalent bond formation, the probe **86** with a molecular mass of 329 Da (monomer unit) would undergo the loss of a leaving group (Cl^- atom) resulting in the mass of 294 Da added per molecule bound. Peak 3 (**Figure 3.22 B**) could represent the formation of the complex AMA1+SA+**86**, accounting for the calculated total peak shift (in relation to peak 1) that equated to 493 Da. Considering a peak shift of 305 Da was observed between peak 2 and 3, the addition of a monomer unit of **86** accounted for the total peak shift of 493 Da, suggesting the existence of a 1:1:1 (m/z ± 20 Da) binding profile. The reduced intensity of both peak 1 and 2 in the 1:10 sample was also in agreement with the formation of a new complex species within the mixture although peak 1 (free 3D7 *Pf*AMA1$_{[104\text{-}438]}$ (DI+II)) remained the dominant species. The identification of covalent modification of AMA1 for the 1:10 sample meant no further testing was required for **86**.

Considering a 10-fold excess of **86** was recognized to facilitate the covalent modification of AMA1, reactive probes **85**, **87**, **91** and **92** (**Figure 3.21**), which were determined to be considerably less reactive than **86**, were initially tested at 10-fold excess (**Figure 3.23**). As shown in Figure 3.23, both **85** (A) and **87** (B) were shown to covalently modify AMA1, causing peak shifts of 236 and 362 Da, respectively. For spectra A of Figure 3.23, the resultant total peak shift was lower than

expected for the possible formation of a 1:1 binding interaction between AMA1 and **85**. Although the expected change in molecular mass upon **85** binding was estimated to be around 282 Da, it is possible LID could be responsible for the observed smaller peak shift. In addition, the formation of an SA adduct did not represent the peak shift identified for spectra A of Figure 3.23 estimated to shift a maximum of 206 Da. Thereby probe **85** was shown to bind AMA1 in a 1:1 manner. For **87**, the resultant peak cluster at 41728 Da clearly displayed a 1:1 binding interaction of AMA1 and **87** with a total peak shift of 362 Da even though the resultant mass was slightly higher than the expected mass per molecule of **87** (334 Da). Akin to the reaction of **85**, the mass accuracy could be affected by fragmentation whilst acquiring MALDI spectra. For both separate experiments, the newly formed complexes of either isothiocyanate (**85**) or maleimide (**87**) were shown as the dominant species when comparing spectra peak intensities with free AMA1. Interestingly, both probes also displayed a similar binding profile (peak shape and intensity), which could be attributed to the similar positioning of electrophilic groups (site 1) on the benzimidazole scaffold.

An evaluation of probes **91** and **92** was also included in Figure 3.23; both **91** (C) and **92** (D) appeared to react with AMA1. For spectra C of Figure 3.23, three peak clusters of decreasing intensity were observed. In relation to the free AMA1 cluster identified as 41328 Da, the largest peak shift equated to 365 Da representative of one unit of **91** binding to AMA1 to form the peak cluster at 41693 Da. The total peak shift was considered smaller than expected which could be the result of LID. Similarly to the binding profile of **86**, the presence of a peak cluster at 41524 Da suggested the formation of an SA adduct with AMA1. However, the newly formed species at 41693 Da was observed to not include SA binding, as opposed to **86**, since the observed shift between the two peaks at 41524 and 41693 Da did not equate to the mass of **91**, calculated at 365 Da. Thus probe **91** was identified to bind in a 1:1 manner with AMA1 and not a 1:1:1 manner as displayed by **86**.

The MALDI profile of the final probe **92** screened also displayed three separate peak clusters identified as 41343 Da, 41527 Da and 41741 Da. Unlike the binding profile of **91**, the shifted peak clusters were observed to increase in intensity particularly the peak cluster at 41527 Da, which represented the dominant species. The largest peak shift corresponded to an *m/z* change of 398 Da representative of one unit of **92** binding to AMA1. Thus the peak cluster at 41741 Da was representative of an AMA1+**92** complex. Similarly to **91** binding, the presence of a 1:1:1 binding interaction as observed for **86** was not evident for **92** considering the difference in mass between peaks at 41527 Da and 41741 Da did not equate to the addition of one unit of **92**, calculated at 424 Da. In addition, the spectra for both **91** (C) and **92** (D) were able to display similar binding profiles

albeit with different intensities, which could be attributed to the specific positioning of reactive groups (site 2) on the benzimidazole scaffold.

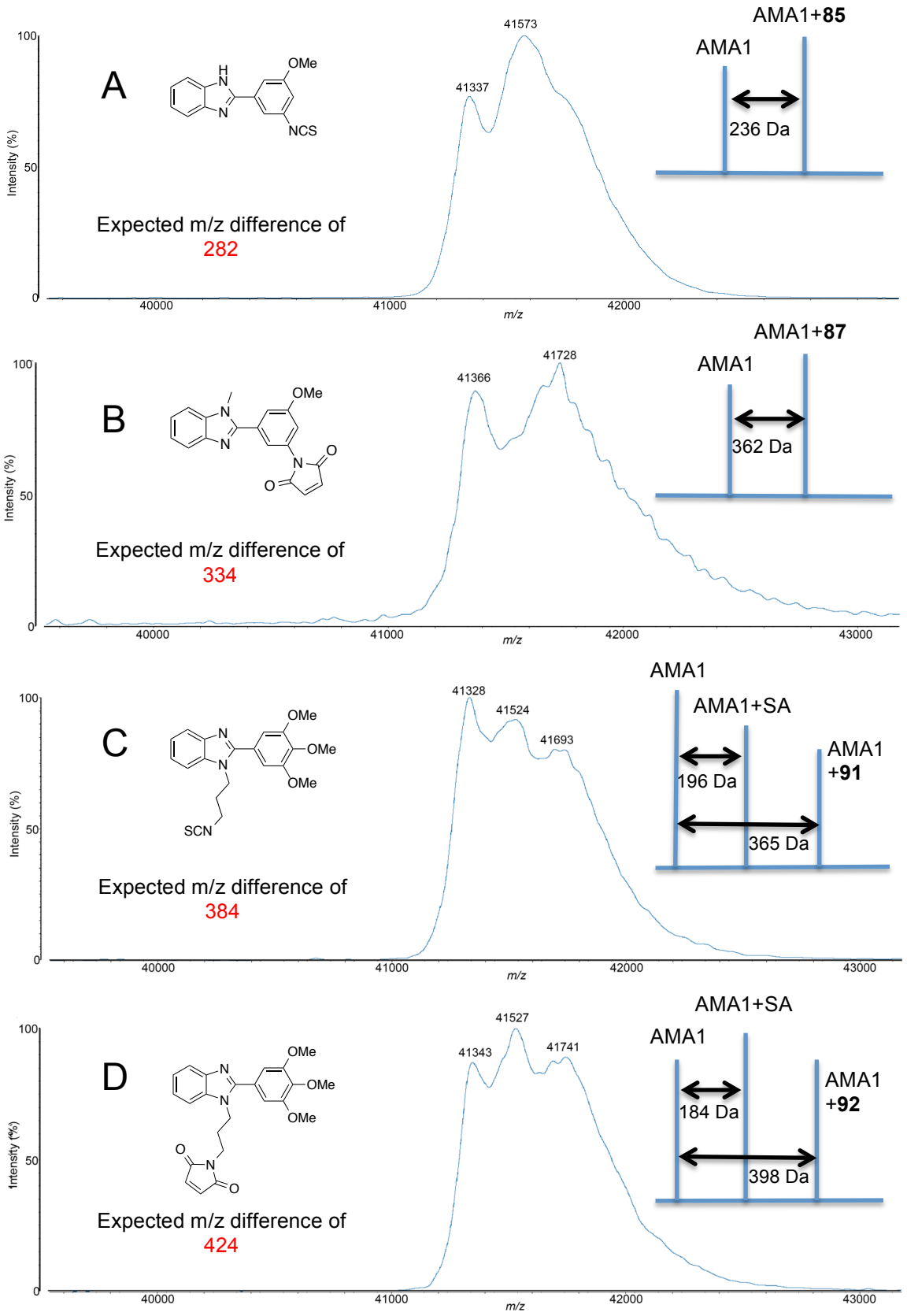


Figure 3.23 MALDI spectra of 3D7 AMA1 with (A) **85**. (B) **87**. (C) **91**. (D) **92**. A 10-fold excess of the probe was used in each experiment in search of covalent modification.

The variation in the binding profiles of **85**, **86** and **87** when compared to **91** and **92** suggested more than one binding site might exist conducive for probe binding. Thus, the identification of each synthesized probe possessing a 1:1 binding profile with 3D7 *Pf*AMA1 meant no further concentration-dependent studies were attempted. Thereby each probe was considered a viable candidate for follow-up X-ray crystallography studies to map preferred probe binding regions. This work is currently being pursued in collaboration with other members of the group.

3.4 Conclusions

Based on SPR and competition NMR data for fragment hits **51** and **52**, the design and synthesis of the 2-substituted aryl benzimidazole series was conducted in an attempt to optimize the core in search of SAR through the expansion of the 2-phenyl substituent accommodating a range of polar and non-polar substituents. Prior to the elaboration of this series, an additional benzimidazole library incorporating various functional moieties of interest to this body of work was screened to more efficiently probe the surrounding chemical space of the scaffold-binding region. Only 2 of 63 molecules screened from the Wellcome Trust library were considered viable hits, both of which were of the chemical class: 2-substituted aryl benzimidazoles. The expansion of the 2-substituted aryl benzimidazoles was conducted via two synthetic methodologies of which, addition of $LaCl_3$ to a solution of *o*-phenylenediamine and substituted aryl benzaldehyde was the preferred method. Addition of $LaCl_3$ allowed for a more efficient synthesis helping to reduce reaction times and temperatures required to afford desired product. The use of $LaCl_3$ methodology required the use of benzaldehydes as starting reagents, which fortunately aided the synthesis through the reduction in by-product formation, witnessed in the synthetic methodology employing benzoic acids under coupling conditions. SPR screening of the 2-substituted aryl benzimidazole series revealed non-stoichiometric binding for each of the analogues indicating that a one-site binding model was not appropriate for the analysis of such data. The inability to identify stoichiometric binding over the concentration range studied, coupled with the inability to test compounds at higher concentrations, meant no feasible SAR could be established for the 2-aryl substituted benzimidazole series. Additionally, inconsistencies in response and affinities observed when examining both original and resynthesized analogues led to the abandonment of SPR as the frontline assay for the validation of compounds targeting AMA1.

Our persistence with the 2-substituted aryl benzimidazoles, which were still considered interesting to this body of work, resulted from encouraging TROSY NMR data that was able to identify several analogues bound to AMA1. A flow cytometry based growth assay was also

employed to screen 26 benzimidazole-based analogues, identifying numerous analogues with potent antimalarial activity against both wild-type and transgenic *P. falciparum* lines. In particular, the better-performing analogues seemed to display functional groups possessing either strong lipophilicity or strong electron-donating ability, as recognized for **62** (3D7 IC_{50} 23 nM; and FVO IC_{50} 10 nM) and **64** (3D7 IC_{50} 10 nM; and FVO IC_{50} <31 nM)**.** The observation of nanomolar potency for analogues suggested that their inhibitory action was independent of AMA1 binding and thus the exact target region for the chemical class remains unknown. The investigation into the mode of action of these analogues of interest will be pursued in collaboration with other members of the group.

An inability to elucidate a clear structural binding profile for the analogues of interest including **53** prompted the exploration of alternative approaches in search of structural information on the binding site of hits. The identification of a four-residue sequence of AMA1, via 2D [^{1}H-^{15}N] TROSY NMR, known as KDGG, was recognized as an important region for benzimidazole analogue binding. The consistent identification of this sequence when screening benzimidazole analogues of interest led to the investigation of particular electrophilic functionalities targeted against residues of the 'binding pocket', in particular the lysine residue. The design of covalently modifying probes concentrated on chemoreactive classes, whilst avoiding the incorporation of any photoreactive functionality owing to the potential AMA1 degradation that may result from UV irradiation.

In total, five covalently modifying probes were synthesized based on compounds **53** and **73**. These two compounds were selected based on encouraging STD NMR data. The haloacetamide, maleimide and isothiocyanate functionalities represented suitable electrophiles that displayed preferential binding for the lysine residue. All three reactive functionalities were synthesized via the formation of an amino moiety in the required position, either site 1 or 2 of the benzimidazole scaffold. All three probes modified at site 1 on the aryl ring were synthesized via conventional methods in moderate to good yield (46-75%). The *N*-alkylation of the pharmacophore **73** required an unconventional strategy to access the key amino-containing intermediate **90** located at site 2 on the benzimidazole scaffold. This reaction proceeded in good yield and was adopted for all covalent probe synthesis concentrated at site 2 of the benzimidazole scaffold.

The reactivity of the covalent probes designed to bind 3D7 AMA1 was evaluated using MALDI-TOF mass spectrometry. For the 5 covalent probes examined, each probe was identified to bind AMA1 in a 1:1 manner with the chloroacetamide group considered the best performing electrophilic group, initially projected to be the most likely electrophilic group to display non-

specific binding if tested against 3D7 AMA1 at high concentration. The chloroacetamide moiety was also observed to improve probe solubility when synthesized on the benzimidazole scaffold and thus would be prioritized for investigation by X-ray crystallography. To our knowledge, this is the first application of covalently modifying molecules targeting AMA1.

3.5 Bibliography

(1) Fonseca, T.; Gigante, B.; Gilchrist, T. L. A short synthesis of phenanthro [2,3-d] imidazoles from dehydroabietic acid. Application of the methodology as a convenient route to benzimidazoles. *Tetrahedron* **2001**, *57*, 1793-1799.
(2) Peddibhotla, S.; Shi, R.; Khan, P.; Smith, L. H.; Mangravita-Novo, A.; Vicchiarelli, M.; Su, Y.; Okolotowicz, K. J.; Cashman, J. R.; Reed, J. C. Inhibition of Protein Kinase C-Driven Nuclear Factor-κB Activation: Synthesis, Structure−Activity Relationship, and Pharmacological Profiling of Pathway Specific Benzimidazole Probe Molecules. *Journal of Medicinal Chemistry* **2010**, *53*, 4793–4797.
(3) Nakano, H.; Inoue, T.; Kawasaki, N.; Miyataka, H. Synthesis of benzimidazole derivatives as antiallergic agents with 5-lipoxygenase inhibiting action. *Chemical and Pharmaceutical Bulletin* **1999**, *47*, 1573-1578.
(4) Valdez, J.; Cedillo, R.; Hernández-Campos, A.; Yépez, L.; Hernández-Luis, F.; Navarrete-Vázquez, G.; Tapia, A.; Cortés, R.; Hernández, M.; Castillo, R. Synthesis and antiparasitic activity of 1*H*-benzimidazole derivatives. *Bioorganic & Medicinal Chemistry Letters* **2002**, *12*, 2221–2224.
(5) Keurulainen, L.; Vahermo, M.; Puente-Felipe, M.; Sandoval-Izquierdo, E.; Crespo-Fernández, B.; Guijarro-López, L.; Huertas-Valentín, L.; las Heras-Dueña, de, L.; Leino, T. O.; Siiskonen, A.; Ballell-Pages, L.; Sanz, L. M.; Castañeda-Casado, P.; Jiménez-Díaz, M. B.; Martínez-Martínez, M. S.; Viera, S.; Kiuru, P.; Calderón, F.; Yli-Kauhaluoma, J. A Developability-Focused Optimization Approach Allows Identification of in Vivo Fast-Acting Antimalarials: *N*-[3-[(Benzimidazol-2-yl)amino]propyl]amides. *Journal of Medicinal Chemistry* **2015**, *58*, 4573–4580.
(6) Camacho, J.; Barazarte, A.; Gamboa, N.; Rodrigues, J.; Rojas, R.; Vaisberg, A.; Gilman, R.; Charris, J. Synthesis and biological evaluation of benzimidazole-5-carbohydrazide derivatives as antimalarial, cytotoxic and antitubercular agents. *Bioorganic & Medicinal Chemistry* **2011**, *19*, 2023–2029.
(7) Lim, S. S.; Debono, C. O.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening. *Australian Journal of Chemistry* **2013**, *66*, 1530–1536.
(8) Hajduk, P. J. Fragment-Based Drug Design: How Big Is Too Big? *Journal of Medicinal Chemsitry* **2006**, *49*, 6972–6976.
(9) Ramachandran, S.; Hameed P, S.; Srivastava, A.; Shanbhag, G.; Morayya, S.; Rautela, N.; Awasthy, D.; Kavanagh, S.; Bharath, S.; Reddy, J.; Panduga, V.; Prabhakar, K. R.; Saralaya, R.; Nanduri, R.; Raichurkar, A.; Menasinakai, S.; Achar, V.; Jiménez-Díaz, M. B.; Martínez, M. S.; Angulo-Barturen, I.; Ferrer, S.; Sanz, L. M.; Gamo, F. J.; Duffy, S.; Avery, V. M.; Waterson, D.; Lee, M. C. S.; Coburn-Flynn, O.; Fidock, D. A.; Iyer, P. S.; Narayanan, S.; Hosagrahara, V.; Sambandamurthy, V. K. *N*-aryl-2-aminobenzimidazoles: novel, efficacious, antimalarial lead compounds. *Journal of Medicinal Chemistry* **2014**, *57*, 6642–6652.
(10) Giannetti, A. M. 8 - From Experimental Design to Validated Hits: A Comprehensive Walk-Through of Fragment Lead Identification Using Surface Plasmon Resonance. *Fragment Based Drug Design Tools, Practical Approaches, and Examples* **2011**, *493*, 169–218.
(11) Harris, K. S.; Casey, J. L.; Coley, A. M.; Karas, J. A.; Sabo, J. K.; Tan, Y. Y.; Dolezal, O.; Norton, R. S.; Hughes, A. B.; Scanlon, D.; Foley, M. Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive *N*-methyl scanning. *Journal of Biological Chemistry* **2009**, *14*, 9361-9371.
(12) Devine, S. M.; Mulcair, M. D.; Debono, C. O.; Leung, E. W. W.; Nissink, J. W. M.; Lim, S. S.; Chandrashekaran, I. R.; Vazirani, M.; Mohanty, B.; Simpson, J. S.; Baell, J. B.; Scammells, P. J.; Norton, R. S.; Scanlon, M. J. Promiscuous 2-aminothiazoles (PrATs): a frequent hitting scaffold. *Journal of Medicinal Chemistry* **2015**, 58, 1205–1214.

(13) LaPlante, S. R.; Carson, R.; Gillard, J.; Aubry, N.; Coulombe, R.; Bordeleau, S.; Bonneau, P.; Little, M.; O'Meara, J.; Beaulieu, P. L. Compound Aggregation in Drug Discovery: Implementing a Practical NMR Assay for Medicinal Chemists. *Journal of Medicinal Chemistry* **2013**, *56*, 5142–5150.
(14) LaPlante, S. R.; Aubry, N.; Bolger, G.; Bonneau, P.; Carson, R.; Coulombe, R.; Sturino, C.; Beaulieu, P. L. Monitoring Drug Self-Aggregation and Potential for Promiscuity in Off-Target In Vitro Pharmacology Screens by a Practical NMR Strategy. *Journal of Medicinal Chemistry* **2013**, *56*, 7073-7083.
(15) Krishnarjuna, B.; Lim, S. S.; Devine, S. M.; Debono, C. O.; Lam, R.; Chandrashekaran, I. R.; Jaipuria, G.; Yagi, H.; Atreya, H. S.; Scanlon, M. J.; MacRaild, C. A.; Scammells, P. J.; Norton, R. S. Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials. *Journal of Molecular Recognition* **2016**, *29*, 281–291.
(16) Hisano, T.; Ichikawa, M.; Tsumoto, K.; Tasaki, M. Synthesis of benzoxazoles, benzothiazoles and benzimidazoles and evaluation of their antifungal, insecticidal and herbicidal activities. *Chemical and Pharmaceutical Bulletin* **1982**, *30*, 2996-3003.
(17) Zelenin, K. N.; Ukraintsev, I. V.; Alekseev, V. V. Intermediates in the reaction of *o*-phenylenediamine with carbonyl compounds and their subsequent conversions. *Chemistry of Heterocyclic Compounds* **1998**, 34, 329-333.
(18) Alaqeel, S. I. Synthetic approaches to benzimidazoles from *o*-phenylenediamine: A literature review. *Journal of Saudi Chemical Society* **2017**, *21*, 229–237.
(19) Azarifar, D.; Pirhayati, M.; Maleki, B.; Sanginabadi, M.; Yami, N., Razieh Acetic acid-promoted condensation of *o*-phenylenediamine with aldehydes into 2-aryl-1-(arylmethyl)-1*H*-benzimidazoles under microwave irradiation. *Journal of Serbian Chemical Society* **2010**, *75*, 1181–1189.
(20) Venkateswarlu, Y.; Kumar, S. R.; Leelavathi, P. Facile and efficient one-pot synthesis of benzimidazoles using lanthanum chloride. *Organic and Medicinal Chemistry Letters* **2013**, *3*, 1–8.
(21) Hameed P, S.; Chinnapattu, M.; Shanbag, G.; Manjrekar, P.; Koushik, K.; Raichurkar, A.; Patil, V.; Jatheendranath, S.; Rudrapatna, S. S.; Barde, S. P.; Rautela, N.; Awasthy, D.; Morayya, S.; Narayan, C.; Kavanagh, S.; Saralaya, R.; Bharath, S.; Viswanath, P.; Mukherjee, K.; Bandodkar, B.; Srivastava, A.; Panduga, V.; Reddy, J.; Prabhakar, K. R.; Sinha, A.; Jiménez-Díaz, M. B.; Martínez, M. S.; Angulo-Barturen, I.; Ferrer, S.; Sanz, L. M.; Gamo, F. J.; Duffy, S.; Avery, V. M.; Magistrado, P. A.; Lukens, A. K.; Wirth, D. F.; Waterson, D.; Balasubramanian, V.; Iyer, P. S.; Narayanan, S.; Hosagrahara, V.; Sambandamurthy, V. K.; Ramachandran, S. Aminoazabenzimidazoles, a novel class of orally active antimalarial agents. *Journal of Medicinal Chemistry* **2014**, *57*, 5702–5713.
(22) Alp, M.; Goker, H.; Brun, R.; Yildiz, S. Synthesis and antiparasitic and antifungal evaluation of 2'-arylsubstituted-1'*H*-[2,5']bisbenzimidazolyl-5-carboxamidines. *European Journal of Medicinal Chemistry* **2009**, *44*, 2002–2008.
(23) Wilson, D. W.; Crabb, B. S.; Beeson, J. G. Development of fluorescent *Plasmodium falciparum* for in vitro growth inhibition assays. *Malaria Journal* **2010**, *9*, 1-12.
(24) Drew, D. R.; Hodder, A. N.; Wilson, D. W.; Foley, M.; Mueller, I.; Siba, P. M.; Dent, A. E.; Cowman, A. F.; Beeson, J. G. Defining the Antigenic Diversity of *Plasmodium falciparum* Apical Membrane Antigen 1 and the Requirements for a Multi-Allele Vaccine against Malaria. *PLoS One* **2012**, *7*, e51023.
(25) Tipker, J.; Verloop, A. Use of STERIMOL, MTD, and MTD* Steric Parameters in Quantitative Structure—Activity Relationships. In *ACS Symposium Series*; ACS Symposium Series; American Chemical Society: Washington, D.C., 2009; Vol. 255, pp. 279–296.
(26) Verloop, A.; Tipker, J. Physical basis of STERIMOL and related steric constants. *Pharmacochemistry library* **1987**, *10*, 97–102.

(27) Krogsgaard-larsen, P.; Liljefors, T.; Madsen, U. Quantitative structure-activity relationships and experimental design. *Textbook of Drug Design and Discovery*. CRC press, **2002**, 3, 117-156.
(28) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty years on: the impact of fragments on drug discovery. *Nature Review Drug Discovery* **2016**, *15*, 605-619.
(29) Shepherd, C. A.; Hopkins, A. L.; Navratilova, I. Fragment screening by SPR and advanced application to GPCRs. *Progress in Biophysics and Molecular Biology* **2014**, *116*, 113–123.
(30) Jörg, M.; Scammells, P. J. Guidelines for the Synthesis of Small-Molecule Irreversible Probes Targeting G Protein-Coupled Receptors. *ChemMedChem* **2016**, *11*, 1488–1498.
(31) Johnson, D. S.; Weerapana, E.; Cravatt, B. F. Strategies for discovering and derisking covalent, irreversible enzyme inhibitors. *Future Medicinal Chemistry* **2010**, *2*, 949–964.
(32) Davie, B. J.; Valant, C.; White, J. M.; Sexton, P. M.; Capuano, B.; Christopoulos, A.; Scammells, P. J. Synthesis and Pharmacological Evaluation of Analogues of Benzyl Quinolone Carboxylic Acid (BQCA) Designed to Bind Irreversibly to an Allosteric Site of the M_1 Muscarinic Acetylcholine Receptor. *Journal of Medicinal Chemistry* **2014**, *57*, 5405–5418.
(33) Muranaka, H.; Momose, T.; Handa, C.; Ozawa, T. Photoaffinity Labeling of the Human A_{2A} Adenosine Receptor and Cross-link Position Analysis by Mass Spectrometry. *ACS Medicinal Chemistry Letters* **2017**, *8*, 660-665.
(34) Kulkarni, P. M.; Kulkarni, A. R.; Korde, A.; Tichkule, R. B.; Laprairie, R. B.; Denovan-Wright, E. M.; Zhou, H.; Janero, D. R.; Zvonok, N.; Makriyannis, A.; Cascio, M. G.; Pertwee, R. G.; Thakur, G. A. Novel Electrophilic and Photoaffinity Covalent Probes for Mapping the Cannabinoid 1 Receptor Allosteric Site(s). *Journal of Medicinal Chemistry* **2016**, *59*, 44–60.
(35) Jacobson, K. A. Functionalized Congener Approach to the Design of Ligands for G Protein-Coupled Receptors (GPCRs). *Bioconjugate Chemistry* **2009**, *20*, 1816–1835.
(36) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The resurgence of covalent drugs. *Nature Reviews Drug Discovery* **2011**, *10*, 307–317.
(37) Lipinski, C.; Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **2004**, *432*, 855–861.
(38) Pei, Y.; Mercier, R. W.; Anday, J. K.; Thakur, G. A.; Zvonok, A. M.; Hurst, D.; Reggio, P. H.; Janero, D. R.; Makriyannis, A. Ligand-Binding Architecture of Human CB2 Cannabinoid Receptor: Evidence for Receptor Subtype-Specific Binding Motif and Modeling GPCR Activation. *Chemistry & Biology* **2008**, *15*, 1207-1219.
(39) Chen, Z.; Jing, C.; Gallagher, S. S.; Sheetz, M. P.; Cornish, V. W. Second-generation covalent TMP-tag for live cell imaging. *Journal of American Chemical Society.* **2012**, *134*, 13692–13699.
(40) Ciruela, F.; Jacobson, K. A.; Fernandez-Duenas, V. Portraying G Protein-Coupled Receptors with Fluorescent Ligands. *ACS Chemical Biology* **2014**, *9*, 1918-1928.
(41) Weichert, D.; Kruse, A. C.; Manglik, A.; Hiller, C.; Zhang, C.; Hubner, H.; Kobilka, B. K.; Gmeiner, P. Covalent agonists for studying G protein-coupled receptor activation. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111*, 10744–10748.
(42) Weichert, D.; Gmeiner, P. Covalent Molecular Probes for Class A G Protein-Coupled Receptors: Advances and Applications. ACS Chemical Biology **2015**, *10*, 1376-1386.
(43) Rosenbaum, D. M.; Zhang, C.; Lyons, J. A.; Holl, R.; Aragao, D.; Arlow, D. H.; Rasmussen, S. G. F.; Choi, H.-J.; DeVree, B. T.; Sunahara, R. K.; Chae, P. S.; Gellman, S. H.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Caffrey, M.; Gmeiner, P.; Kobilka, B. K. Structure and function of an irreversible agonist-β_2 adrenoceptor complex. *Nature* **2010**, *469*, 236–240.
(44) Takemori, A. E.; Portoghese, P. S. Affinity labels for opioid receptors. *Annual Review of Pharmacology and Toxicology* **1985**, *25*, 193–223.
(45) Morey, T. E.; Belardinelli, L.; Dennis, D. M. Validation of Furchgott's method to determine agonist-dependent A_1-adenosine receptor reserve in guinea-pig atrium. *British Journal of Pharmacology.* **1998**, *123*, 1425–1433.

(46) Buck, E.; Bourne, H.; Wells, J. A. Site-specific Disulfide Capture of Agonist and Antagonist Peptides on the C5a Receptor. *Journal of Biological Chemistry* **2005**, *280*, 4009–4012.
(47) Flanagan, M. E.; Abramite, J. A.; Anderson, D. P.; Aulabaugh, A.; Dahal, U. P.; Gilbert, A. M.; Li, C.; Montgomery, J.; Oppenheimer, S. R.; Ryder, T.; Schuff, B. P.; Uccello, D. P.; Walker, G. S.; Wu, Y.; Brown, M. F.; Chen, J. M.; Hayward, M. M.; Noe, M. C.; Obach, R. S.; Philippe, L.; Shanmugasundaram, V.; Shapiro, M. J.; Starr, J.; Stroh, J.; Che, Y. Chemical and Computational Methods for the Characterization of Covalent Reactive Groups for the Prospective Design of Irreversible Inhibitors. *Journal of Medicinal Chemistry* **2014**, *57*, 10072–10079.
(48) Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289*, 739–745.
(49) Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F. Activity-based probes for the proteomic profiling of metalloproteases. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 10000–10005.
(50) Salisbury, C. M.; Cravatt, B. F. Activity-based probes for proteomic profiling of histone deacetylase complexes. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 1171–1176.
(51) Fedan, J. S.; Hogaboom, G. K.; O'Donnell, J. P. Photoaffinity labels as pharmacological tools. *Biochemical Pharmacology* **1984**, *33*, 1167–1180.
(52) Fleming, S. A. Chemical reagents in photoaffinity labeling. *Tetrahedron* **1995**, *51*, 12479-12520.
(53) Dorman, G.; Prestwich, G. D. Benzophenone Photophores in Biochemistry. *Biochemistry* **1994**, *33*, 5661-5673.
(54) Weber, P. J.; Beck-Sickinger, A. G. Comparison of the photochemical behavior of four different photoactivatable probes. *Jounral of Peptide Research* **1997**, *49*, 375–383.
(55) Das, J. Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chemical Reviews* **2011**, *111*, 4405–4417.
(56) Brunner, J.; Senn, H.; Richards, F. M. 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *Journal of Biological Chemistry* **1980**, *255*, 3313–3318.
(57) Lamarche, N.; Gaudreau, P.; Massie, B.; Langelier, Y. Affinity of synthetic peptides for the HSV-2 ribonucleotide reductase R1 subunit measured with an iodinated photoaffinity peptide. *Analytical Biochemistry* **1994**, *220*, 315–320.
(58) Schuster, G. B.; Platz, M. S. Photochemistry of Phenyl Azide. *Advances in Photochemistry* **1992**, *17*, 69-143.
(59) Bräse, S.; Gil, C.; Knepper, K.; Zimmermann, V. Organic Azides: An Exploding Diversity of a Unique Class of Compounds. *Angewandte Chemie International Edition.* **2005**, *44*, 5188–5240.
(60) Buck, E.; Wells, J. A. Disulfide trapping to localize small-molecule agonists and antagonists for a G protein-coupled receptor. *Proceedings of the National Academy of Sciences* **2005**, *102*, 2719–2724.
(61) Doorn, J. A.; Petersen, D. R. Covalent Modification of Amino Acid Nucleophiles by the Lipid Peroxidation Products 4-Hydroxy-2-nonenal and 4-Oxo-2-nonenal. *Chemical Research in Toxicology* **2002**, *15*, 1445-1450.
(62) Griffin, M. T.; Thomas, E. A.; Ehlert, F. J. Kinetics of Activation and *in Vivo* Muscarinic Receptor Binding of *N*-(2-Bromoethyl)-4-Piperidinyl Diphenylacetate: An Analog of 4-DAMP Mustard[1]. *Journal of Pharmacology and Experimental Therapeutics* **1993**, *266*, 301-305.

(63) Baraldi, P. G.; Cacciari, B.; Moro, S.: Romagnoli, R.; Ji, X,-D.; Jacobson, K. A.; Gessi, S.; Borea, P. A.; Spalluto, G. Fluorosulfonyl- and Bis-(*β*-chloroethyl)amino-phenylamino Functionalized Pyrazolo[4,3-*e*]1,2,4-triazolo[1,5-*c*]pyrimidine Derivatives: Irreversible Antagonists at the Human A3 Adenosine Receptor and Molecular Modeling Studies. *Journal of Medicinal Chemistry* **2001**, *44*, 2735-2742.
(64) Thurkauf, A.; Mattson, M. V.; Huguenin, P. N.; Rice, K. C.; Jacobson, A. E. Etoxadrol-*meta*-isothiocyanate: a potent, enantioselective, electrophilic affinity ligand for the phencyclidine-binding site. *FEBS Letters* **1988**, *238*, 369–374.
(63) Wong, R.; Dolman, S. J. Isothiocyanates from Tosyl Chloride Mediated Decomposition of in Situ Generated Dithiocarbamic Acid Salts. *Journal of Organic Chemistry* **2007**, *72*, 3969-3971.
(65) Wong, R.; Dolman, S. J. Isothiocyanates from Tosyl Chloride Mediated Decomposition of in Situ Generated Dithiocarbamic Acid Salts. *Journal of Organic Chemistry* **2007**, *72*, 3969-3971.
(66) Baumgold, J.; Karton, Y.; Malka, N.; Jacobson, K. A. High-affinity acylating antagonists for muscarinic receptors. *Life Sciences* **1992**, *51*, 345–351.
(67) Tsotinis, A.; Afroudakis, P. A.; Davidson, K.; Prashar, A.; Sugden, D. Design, Synthesis, and Melatoninergic Activity of New Azido- and Isothiocyanato-Substituted Indoles. *Journal of Medicinal Chemistry* **2007**, *50*, 6436–6440.
(68) Tahtaoui, C.; Balestre, M.-N.; Klotz, P.; Rognan, D.; Barberis, C.; Mouillac, B.; Hibert, M. Identification of the Binding Sites of the SR49059 Nonpeptide Antagonist into the V 1aVasopressin Receptor Using Sulfydryl-reactive Ligands and Cysteine Mutants as Chemical Sensors. *Journal of Biological Chemistry* **2003**, *278*, 40010–40019.
(69) Jöst, C.; Nitsche, C.; Scholz, T.; Roux, L. Klein, C. D. Promiscuity and Selectivity in Covalent Enzyme Inhibition: A Systematic Study of Electrphilic Fragments. *Journal of Medicinal Chemistry* **2014**, *57*, 7590-7599.
(70) Burke, T. R.; Bajwa, B. S.; Jacobson, A. E.; Rice, K. C.; Streaty, R. A.; Klee, W. A. Probes for Narcotic Receptor Mediated Phenomena. 7.[1] Synthesis and Pharmacological Properties of Irreversible Ligands Specific for *μ* or *δ* Opiate Receptors. *Journal of Medicinal Chemistry* **1984**, *27*, 1570-1574.
(71) Keller, B. O.; Sui, J.; Young, A. B.; Whittal, R. M. Interferences and contaminants encountered in modern mass spectrometry. *Analytica Chimica Acta* **2008**, *627*, 71–81.
(72) Macht, M.; Asperger, A.; Deininger, S.-O. Comparison of laser-induced dissociation and high-energy collision-induced dissociation using matrix-assisted laser desorption/ionization tandem time-of-flight(MALDI-TOF/TOF) for peptide and protein identification. *Rapid Communications in Mass Spectrometry* **2004**, *18*, 2093–2105.

Chapter 4 - Conclusions and future work

Exploration of the 2-aminothiazole and benzimidazole series as AMA1 inhibitors was conducted using competition SPR and NMR data from MIPS library fragment hits. A lack of structural data for fragments bound to AMA1 led to the expansion of each scaffold in search of higher affinity binders of AMA1. Optimization of each scaffold begun with the development of a series of analogues in search of SAR using a range of polar and non-polar substituents to probe the surrounding chemical space in and around the hydrophobic cleft of AMA1.

The 2-aminothiazole scaffold (Chapter 2) was considered the most prevalent chemical class identified from the MIPS fragment library. All 2-aminothiazoles identified from the screen also possessed ideal LE, crucial to the progression of an FBDD investigation. In total, two separate series, derived from the 2-aminothiazole scaffold, were optimized. Design and synthesis initially focused on the expansion of the aryl ring of the 4-aryl 2-aminothiazole scaffold. Screening by SPR revealed the *m*-methyl substituted aminothiazole **14** as the highest ranked analogue of the series however the SAR was considered inconsistent since the majority of analogues bound non-specifically to AMA1. However, the binding mode of **14** could never be established via X-ray crystallography. Further attempts to identify the binding region of selected analogues using 2D [^{1}H-^{15}N]-TROSY NMR exposed no definitive evidence for a single preferred binding region which made characterization difficult. Despite this the 2-aminothiazole scaffold remained critical to the identification of AMA1 inhibitors leading to the elaboration of the 4-methyl-5-aryl substituted aminothiazoles, another series derived from the screening campaign.

Expanding the 4-methyl-5-aryl substituted aminothiazole series focused on the 5-aryl ring, shown to be crucial for the preservation of AMA1 analogue binding (Chapter 2, section 2.3). Although these analogues were able to bind more strongly than the previous 2-aminothiazole series discussed, on average the increased lipophilicity owing to the presence of a 4-methyl group resulted in poor analogue solubility during the screening campaign. This meant any data obtained for the series would be difficult to interpret thus the series was suggested to display no recognizable SAR. Of the 26 analogues synthesized, only 2 analogues in the form of the *p*-hydroxy **31** and *p*-amino **34** were shown to consistently bind via SPR however **34** still showed signs of concentration-dependent aggregation. A follow-up investigation using 2D [^{1}H-^{15}N]-TROSY NMR to screen the more consistent binders of the series confirmed the existence of multiple AMA1 binding regions. Of this investigation, the binding region, entitled the 'KDGG' binding region was consistently shown to support analogue binding providing the first significant structural finding to aid development of

these and other series in the future. However, the development of this scaffold was halted owing to the continual identification of 2-aminothiazoles as positive hits against a myriad of different protein targets. A study into the promiscuity of the scaffold was completed in parallel to the AMA1 FBDD campaign. This investigation identified the scaffold as a promiscuous binding scaffold suggesting the scaffold to be a poor starting point for an FBDD investigation. In response to the outcome of the study, all work completed on the series was abandoned and the removal of all aminothiazoles from compound libraries was advised.

Similar to the 2-aminothiazoles (Chapter 2), a majority of the 2-substituted aryl benzimidazole analogues synthesized (Chapter 3, section 3.1) were shown to display non-stoichiometric binding when screened via SPR. The presence of non-stoichiometric binding inferred that a one-site binding model again was not appropriate for the analysis of such analogues. Thereby all attempts to identify any plausible SAR via SPR gave inconclusive results eventually leading to the abandonment of SPR as the frontline assay for the project. In parallel, efforts to identify the binding region of the better performed benzimidazoles, such as the *m*-hydroxy **51** and *m*-amino **52**, using 2D [^{1}H-^{15}N]-TROSY NMR were shown to be in agreement with 2-aminothiazole data suggesting preferential binding for the 'KDGG' binding region characterized by increased residue perturbations. Although we were able to acquire this helpful structural information, the expansion of the series was halted by the inconsistent measurement of compound affinity via SPR and the inability to crystallize analogues with AMA1.

During fragment development, the difficulty in acquiring structural data for bound analogues via NMR hindered the optimization of fragment hits. In an effort to access more specific higher affinity inhibitors of AMA1, the design and synthesis of covalently modifying molecules was undertaken (Chapter 3, section 3.2). Prior to the synthesis of covalent probes, the elaboration of the *m*-OMe benzimidazole **53**, the only benzimidazole to display stoichiometric binding via SPR, was performed. The elaboration of **53** led to the development of numerous analogues (**Figure 4.1**); this optimization campaign was able to identify two regions capable of accommodating the addition of reactive groups without affecting analogue affinity for AMA1. The successful synthesis of covalent reactive groups at two different positions on the binding scaffold helped to assess the reactivity and selectivity of each covalently modifying compound. Each probe was shown to successfully covalently bind to 3D7 AMA1 in a 1:1 manner following an experimentally determined incubation protocol. This was the first reported application of covalently modifying molecules designed to bind AMA1 in search of higher specificity and affinity. Although we could not determine the exact binding region of the probes using MALDI mass spectrometry, X-ray crystallography was planned

to follow up all covalent probes tested in order to characterize the binding regions of compounds. The successful investigation into the covalent modification of AMA1 meant it would be logical to attempt to obtain crystal structures of covalent probes bound to AMA1 to determine its exact location on the protein structure. The possibility of non-specific binding was drastically reduced owing to the 1:1 binding observed during MALDI analysis.

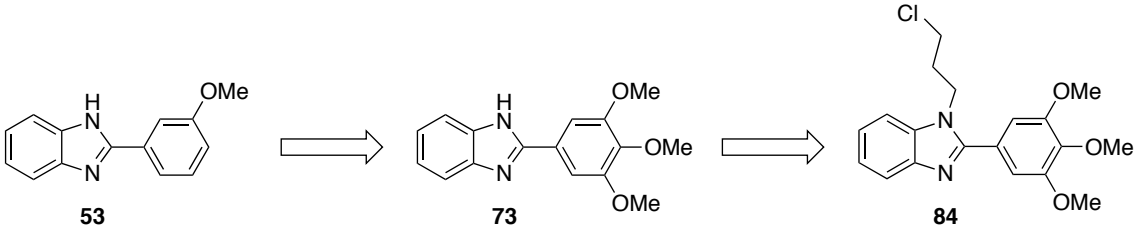


Figure 4.1 Evolution of the benzimidazole **53**.

In addition to X-ray crystallography, other techniques to identify the binding mode of AMA1 have been planned and could be completed in collaboration with other members of the group. One method was a site-directed mutagenesis study of the AMA1 to help identify the binding region(s) of compounds. This technique would include using previously acquired TROSY NMR data to mutate specific perturbed residues to determine whether compound binding would be impacted. This process of elimination would help to identify which region of AMA1 is preferred for binding and also identify other interactions with the backbone that might not be identified via NMR. One particular region targeted throughout the project was the 'KDGG' region, this region would be an ideal starting point for this investigation considering each amino acid residue was consistently perturbed by a variety of benzimidazole analogues when examined via TROSY NMR.

Another method that could be used to investigate the binding mode of each covalent probe would involve the in-solution or in-gel tryptic digestion for mass spectrometry of an AMA1-ligand complex. This technique is used to digest proteins into peptides to be analyzed by MALDI-MS or liquid chromatography-tandem mass spectrometry (LC-MS/MS). The analysis of each peptide formed would allow for the identification of the covalent probe on a specific peptide belonging to the protein.

One more technique that could be used to identify the binding region of an analogue is site-directed spin labeling. This technique had previously been used apart of the tRON2 peptide study in an effort to map the surrounding area of the peptide when bound to AMA1 via TROSY NMR. This work was completed within the Ray Norton group in parallel to the FBDD study of AMA1. This

technique would involve the attachment of a spin-label, an organic molecule containing an unpaired electron, to the peptide resulting in the signal of any residue within a certain range (typically within 20 Å) broadening. The spin-label would need to be attached to the compound via a chemical handle such as a reactive thiol or amine whilst also considering the resultant changes to the physicochemical properties of the molecule prior to analysis via TROSY NMR. Although this strategy is not as informative as other techniques previously discussed, this technique is considered the simplest to perform in comparison to the others.

The abandonment of SPR as the primary screening assay led to the employment of a *P. falciparum* growth inhibition assay to re-screen 2-substituted aryl benzimidazoles. This assay was able to identify numerous 2-substituted aryl benzimidazole analogues possessing potent nanomolar antimalarial activity against multiple transgenic *P. falciparum* strain lines. The *p*-amino **62** and *p*-trifluoromethyl **64** both displayed similar strong potency for 3D7 and FVO AMA1. Although strong evidence existed exhibiting benzimidazole analogues bound to AMA1, the data observed for the growth inhibition data did not match previously acquired binding data suggesting an alternative mode of action was plausible for the 2-substituted aryl benzimidazole scaffold. One method that could help define the mode of action of an analogue is mass spectrometry-based metabolomics in search of the metabolic phenotype of the inhibition process identified via the growth-inhibition assay. This assay could adopt both a targeted or global approach to search and quantify metabolites of interest that may result from the inhibition of particular pathways or receptors. This work would be completed in collaboration with members of another group located at Monash University.

Although the covalent probe investigation and the optimization and continual growth inhibition screening of the benzimidazole series will remain high on the agenda in the near future, another way to improve analogue affinity could involved the endogenous ligand RON2 investigated as a peptide inhibitor of AMA1. A study by Wang *et. al*[1] investigated a 13-residue β-hairpin based on the C-terminal loop of RON2 to probe a conserved binding site on *P. falciparum* AMA1. This work focused on establishing SAR around the 13-residue peptide via a series of mutations in an attempt to improve binding affinity against multiple strains of AMA1. This study was able to show markedly improvement in affinity for both FVO and 3D7 AMA1, designing to their knowledge the most potent strain-transcending peptide yet reported. Considering the 13-residue β-hairpin peptide is a truncated version of full length RON2, binding AMA1 at the hydrophobic cleft, the strategy of covalently attaching a small molecule inhibitor to the truncated peptide could help to solve the issue of poor analogue affinity. To date the tRON2 mutated peptides (unpublished work) have also been crystallized with AMA1 thereby through utilizing the potency of the peptide (strong micromolar

K_D) and a free reactive thiol present within the peptide, it could be possible to attach a small molecule to the peptide and subsequently attempt to crystallize the complex with AMA1 to obtain a crystal structure which could aid the development of high-affinity molecules. The analogue would initially require optimization in order to covalently attach to the peptide via a specific reactive handle. This work is yet to be started and could be an option moving forward to integrate peptide work with the FBDD investigation against AMA1.

4.1 Bibliography

(1) Wang, G.; Drinkwater, N.; Drew, D. R.; MacRaild, C. A.; Chalmers, D. K.; Mohanty, B.; Lim, S. S.; Anders, R. F.; Beeson, J. G.; Thompson, P. E.; McGowan, S.; Simpson, J. S.; Norton, R. S.; Scanlon, M. J. Structure–Activity Studies of β-Hairpin Peptide Inhibitors of the *Plasmodium falciparum* AMA1–RON2 Interaction. *Journal of Molecular Biology* **2016**, *428*, 3986–3998.

Chapter 5 - Experimental

5.1 General experimental

Small molecule NMR spectra were recorded at 400 MHz, for ^{1}H NMR and 101 MHz, for ^{13}C NMR on a Bruker Avance III Nanobay 400 MHz coupled to a BACS 60 automatic sample changer, NMR spectrometer all at 298K (each case is specified by MHz). The spectrometer is equipped with a 5mm PABBO BB-1H/D Z-GRD probe. Data acquisition and processing was managed using XWINNMR (Bruker) software package version 3.5 and plotting was managed using iNMR (mac compatible) v5.5.3. All chemical shifts for ^{1}H NMR were measured in parts per million (ppm) referenced to an internal standards of residual proteo-solvent, δ 4.79 for deuterium oxide, δ 7.26 for chloroform, δ 2.50 for DMSO and δ 3.31 for methanol.[1] ^{13}C NMR were measured in parts per million (ppm) referenced to an internal standard of residual proteo-solvent, δ 77.16 for chloroform, δ 39.52 for DMSO and δ 49.00 for methanol.[1] ^{19}F NMR were measured in parts per million (ppm) referenced to an internal standard of residual proteo-solvent, δ 77.16 for chloroform, δ 39.52 for DMSO and δ 49.00 for methanol.[1] Multiplicity is quoted as a (apparent), b (broad), s (singlet), d (doublet), t (triplet), q (quartet), p (pentet) and m (multiplet). All compounds were >95% purity by ^{1}H NMR prior to biological testing. High Resolution Mass Spectrometry analysis were performed on an Agilent 6224 TOF LC/MS Mass Spectrometer coupled to an Agilent 1290 Infinity (Agilent, Palo Alto, CA). All data was acquired and reference mass corrected via a dual-spray electrospray ionization (ESI) source. Each scan or data point on the Total Ion Chromatogram (TIC) is an average of 13,700 transients, producing a spectrum every second. Mass spectra were created by averaging the scans across each peak and background subtracted against the first 10 seconds of the TIC. Acquisition was performed using the Agilent Mass Hunter Data Acquisition software version B.05.00 Build 5.0.5042.2 and analysis was performed using Mass Hunter Qualitative Analysis version B.05.00 Build 5.0.519.13. Low Resolution Mass Spectrometry analyses were performed using a Micromass Platform II single quadrupole mass spectrometer equipped with an atmospheric pressure (ESI/APCI) ion source. Sample management was facilitated by an Agilent 1100 series HPLC system and the instrument was controlled using MassLynx v3.5. Microwave reactions were conducted in a Biotage Initiator TM, in 0.5-2.0 mL vials according to manufacturer's instructions. Protein LCMS were conducted on a Shimadzu LC-20 LCMS system with an SGE ProteoCol C8 HQ1003 150mm X 2.0mm ID 3μM 1000 Å column monitoring at 214 nm. Buffer A: 99.9% water, 0.1% formic acid; buffer B: 99.9% acetonitrile, 0.1% formic acid. A gradient of 0-60% buffer B over 16 mins followed by 1 min isocratic 60% buffer B, 1 min gradient to 100% buffer A and 2 min

of isocratic 100% buffer A was used as an elution profile. Both the following gradients were used during protein/peptide testing; 0-60% over 10 mins and 0-100% over 15 mins (acetonitrile). Data was analysed using LabSolutions v5.31 and mMass v5.0.1. Thin layer chromatography was performed on Merck Silica Gel 60 F_{254} plates. TLC plates were visualized under UV illumination at 254 nm and/or with the aid of phosphomolybdic acid (PMA), vanillin, *p*-anisaldahyde, iodine or potassium permanganate solutions. Column chromatography was conducted using Davisil silica gel LC60A (40-63 μm). All commercially available chemicals were purchased from Aldrich/Fluka, Merck, Alfa-Aesar, Boron Molecular, GL Biochem and Matrix scientific and stored appropriately and used as required. HPLC-grade THF, DCM and DMF were dried using an Mbraun solvent purification system (SPS-800) according to the manufacturer's instructions. All other solvents were reagent grade and used as required.

5.2 2-Aminothiazole derivatives and intermediates

5.2.1 4-Aryl-substituted aminothiazole analogues

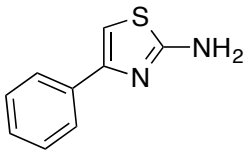
General procedure A for the synthesis of 4-Aryl-substituted-2-aminothiazole analogues

To a solution of aryl-substituted acetophenone (200 mg) dissolved in MeCN (3 mL) stirred at rt, bromine (Br_2) (1.1 equiv) was added and reacted for 2 h at 100 °C. Once the bromination reaction completed, thiourea (1.1 equiv) was added to the mixture and further reacted for 1 h at 100 °C. The resultant precipitate was collected by vacuum filtration, washed using aqueous Na_2CO_3 (1 M) (20 mL) and extracted in EtOAc (2 x 20 mL). If no precipitate forms, the extraction of the desired material was performed using EtOAc and aqueous Na_2CO_3 as previously mentioned. The combined organic layers were dried over $MgSO_4$ and reduced *in vacuo* to afford a product. If necessary, purification by column chromatography was applicable for resultant material.

General procedure B for the synthesis of 4-Aryl-substituted-2-aminothiazole analogues

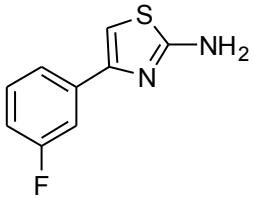
Aryl-substituted acetophenone (200 mg) and thiourea (3 equiv) were added to a microwave vial. Iodine (I_2) (1.5 equiv) was added before the vial was sealed and heated at 130 °C for 1 h. The reaction was diluted using $Na_2S_2O_3.5H_2O$ (1 M) (20 mL) and then basified using sat. Na_2CO_3. Once the mixture was basic (pH 9-10), extraction was completed using EtOAc (2 x 20 mL). The organic layer was washed using brine solution (30 mL), dried over $MgSO_4$ and reduced *in vacuo*. If necessary, purification by column chromatography was applicable for resultant material.

*Synthesis of 4-phenylthiazol-2-amine (**1**)*



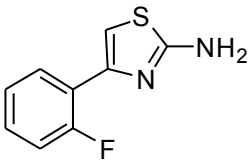
Phenacyl bromide (500 mg, 2.5 mmol) was suspended in MeCN. Thiourea (249 mg, 3.2 mmol) was added and mixture was heated for 2 h at 100 °C. Once reacted, mixture was cooled in an ice bath 0 °C to afford a white precipitate (630 mg). The resultant solid was dissolved in EtOAc (30 mL) and washed using sat Na_2CO_3 soln. (100 mL). The organic layer was collected and dried over $MgSO_4$ and reduced *in vacuo* to afford title compound (345 mg, 78 %). ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.79-7.76 (m, 2H, Ar), 7.40-7.36 (m, 1H, Ar), 7.31-7.26 (m, 2H, Ar), 6.73 (s, 1H, Ar), 5.06 (s, 2H, NH_2); ^{13}C NMR (101 MHz; DMSO-d_6): δ 168.1, 149.8, 134.9, 128.4, 127.1, 125.5, 101.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 177.0408; Observed $[M+H]^+$ = 177.0409.

*Synthesis of 4-(3-fluorophenyl)thiazol-2-amine (**10**)*



Applying the general procedure A, 3'-fluoroacetophenone (200 mg, 1.44 mmol, 178 µL). The crude material was purified via column chromatography (DCM), affording the title compound as a yellow solid (44 mg, 16 %). R_f: 0.2. m.p: 75-78 °C. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.57-7.52 (m, 1H, Ar), 7.51-7.45 (d, 1H, *J* = 8.0 Hz, Ar), 7.37-7.29 (m, 1H, Ar), 7.02-6.94 (t, 1H, *J* = 12.0 Hz, Ar), 6.75 (s, 1H, Ar), 5.20 (s, 2H, NH_2); ^{19}F NMR (400 MHz; $CDCl_3$) δ -113.3 (s, 1F, F); ^{13}C NMR (101 MHz; DMSO-d_6): δ 169.2, 162.4 (d, *J* = 243.0 Hz), 149.5 (d, *J* = 3.0 Hz), 137.3 (d, *J* = 8.0 Hz), 130.4 (d, *J* = 9.0 Hz), 121.5 (d, *J* = 3.0 Hz), 113.7 (d, *J* = 21.0 Hz), 112.0 (d, *J* = 22.0 Hz), 103.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 195.0314; Observed $[M+H]^+$ = 195.0389.

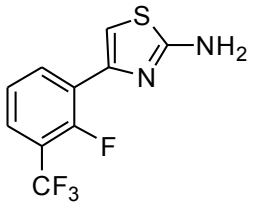
*Synthesis of 4-(2-fluorophenyl)thiazol-2-amine (**11**)*



2'-Fluoroacetophenone (200 mg, 1.45 mmol, 178 µL) was added to a solution of MeOH (5 mL) and NaBr (164 mg, 1.59 mmol). Oxone® (979 mg, 1.59 mmol) was added to the reaction mixture and stirred for 2 h at 100 °C, closely monitored by TLC. The resultant mixture was quenched in aqueous sodium thiosulphate ($Na_2S_2O_3$) and extracted in EtOAc twice. The combined organic layers was dried over $MgSO_4$, filtered and reduced *in vacuo* affording a brownish solid. Note: work-up required immediately after completion to avoid degradation. The crude product was taken up in EtOH (3 mL) and thiourea (39 mg, 0.51 mmol) was added into the reaction mixture. The mixture was heated for 2 h at 100 °C. The resultant mixture reduced *in vacuo* and the crude product was purified by

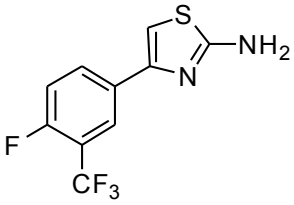
column chromatography (DCM: MeOH: NH_4OH; 96: 2: 2) to afford the title compound as a pink solid (31 mg, 11 %). R_f: 0.3. m.p: 76-80 °C. ^{1}H NMR (400 MHz; $CDCl_3$): δ 8.05-7.98 (t, 1H, *J* = 8.0 Hz, Ar), 7.29-7.21 (m, 1H, Ar), 7.21-7.15 (t, 1H, *J* = 8.0 Hz, Ar), 7.14-7.07 (m, 1H, Ar), 7.02 (s, 1H, Ar), 5.16 (s, 2H, NH_2); ^{19}F NMR (400 MHz; $CDCl_3$) δ -114.2 (s, 1F, F); ^{13}C NMR (101 MHz; $CDCl_3$): 166.6, 160.3 (d, *J* = 223.0 Hz), 144.8 (d, *J* = 4.2 Hz), 129.8 (d, *J* = 6.1 Hz), 128.9 (d, *J* = 35.3 Hz), 124.4, 122.4 (d, *J* = 21.0 Hz), 116.0 (d, *J* = 45.0 Hz), 108.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 195.0314; Observed $[M+H]^+$ = 195.0391.

Synthesis of 4-(2-fluoro-3-(trifluoromethyl)phenyl)thiazol-2-amine



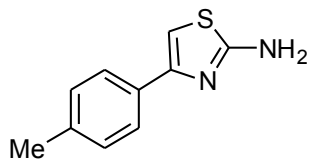
Applying the general procedure A, 2'-fluoro-3'-(trifluoromethyl)acetophenone (200 mg, 0.97 mmol, 196 µL). The resulting mixture was reduced *in vacuo*, where the resulting residue was purified via column chromatography (DCM) to afford the title compound as an olive-green powder (58 mg, 23 %). R_f: 0.4. m.p: 75-77 °C. ^{1}H NMR (400 MHz; $CDCl_3$): δ 8.29-8.21 (td, 1H, *J* = 8.0 Hz, Ar), 7.55-7.48 (td, 1H, *J* = 8.0 Hz, Ar), 7.32-7.24 (m, 1H, Ar), 7.10 (s, 1H, Ar), 5.16 (s, 2H, NH_2); ^{19}F NMR (400 MHz; $CDCl_3$) δ -61.3 (d, 3F, *J* = 12.0 Hz, F), -116.7 (q, 1F, *J* = 12.0 Hz, CF_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 166.5, 156.7 (d, *J* = 215.0 Hz), 133.8 (d, *J* = 3.0 Hz), 131.7 (d, *J* = 4.0 Hz), 125.8 (d, *J* = 16 Hz), 124.2 (d, *J* = 4.0 Hz), 122.5 (d, *J* = 226.0 Hz), 115.2, 109.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 263.0188; Observed $[M+H]^+$ = 263.0263.

Synthesis of 4-(4-fluoro-3-(trifluoromethyl)phenyl)thiazol-2-amine



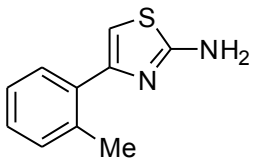
Applying the general procedure A using NBS (190 mg, 1.07 mmol), 4'-fluoro-3'-(trifluoromethyl)acetophenone (200 mg, 0.97 mmol). The resultant yellow powder was purified via column chromatography (DCM: NH_4OH; 98: 2) to afford a pale yellow powder (34 mg, 14 %). R_f: 0.4. m.p: 133-135 °C. ^{1}H NMR (400 MHz; $CDCl_3$) δ 8.05-8.00 (d, 1H, *J* = 8.0 Hz, Ar), 7.95-7.89 (m, 1H, Ar), 7.24-7.15 (t, 1H, *J* = 8.0 Hz, Ar), 6.73 (s, 1H, Ar), 5.18 (s, 2H, NH_2); ^{19}F NMR (400 MHz; $CDCl_3$) δ -61.4 (d, 1F, *J* = 12.0 Hz, F), -116.0 (q, 3F, *J* = 12.0 Hz, CF_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 167.7, 158.6 (d, *J* = 222.4 Hz), 149.8, 142.3, 131.2 (d, *J* = 8.0 Hz), 124.9 (q, *J* = 4.0 Hz), 121.4 (d, *J* = 231.1 Hz), 117.2 (d, *J* = 21.0 Hz), 103.8. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 263.0188; Observed $[M+H]^+$ = 263.0263.

*Synthesis of 4-(p-tolyl)thiazol-2-amine (**12**)*



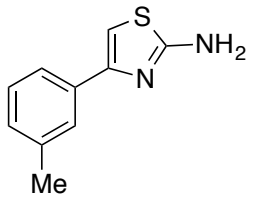
Applying the general procedure A, 4'-methylacetophenone (200 mg, 1.49 mmol, 200 µL). A green solid precipitates out of solution once thiourea has dissolved. The green solid was collected by vacuum filtration and was washed using aqueous Na_2CO_3 (1 M) and extracted in EtOAc. The organic portion was dried over $MgSO_4$ and reduced *in vacuo* to afford a light yellow powder (40 mg, 14 %). m.p: 100-102 °C. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.70-7.60 (d, 2H, *J* = 8 Hz, Ar), 7.21-7.16 (d, 2H**,** *J* = 8 Hz, Ar), 6.67 (s, 1H, Ar), 5.03 (s, 2H, NH_2), 2.36 (s, 3H, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): 167.2, 151.5, 137.7, 132.1, 129.4, 126.0, 102.2, 21.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 191.0565; Observed $[M+H]^+$ = 191.0639.

*Synthesis of 4-(o-tolyl)thiazol-2-amine (**13**)*



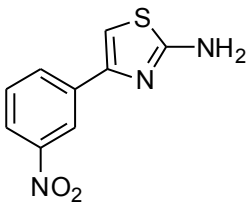
Applying the general procedure A, 2'-methylacetophenone (200 mg, 1.49 mmol, 197 µL). The combined organic layer was dried over $MgSO_4$ and reduced *in vacuo* to afford a pinkish solid (84 mg, 30 %). m.p: 73-75 °C. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.55-7.50 (m, 1H, Ar), 7.26-7.17 (m, 3H, Ar), 6.45 (s, 1H, Ar), 5.28 (s, 2H, NH_2), 2.44 (s, 3H, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 166.7, 151.4, 136.2, 135.1, 130.8, 129.7, 127.9, 125.9, 105.8, 21.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 191.0565; Observed $[M+H]^+$ = 191.0632.

*Synthesis of 4-(m-tolyl)thiazol-2-amine (**14**)*



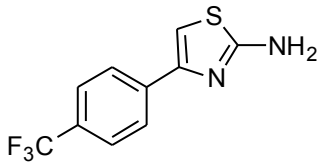
Applying the general procedure A, 3'-methylacetophenone (200 mg, 1.49 mmol, 203 µL). The precipitate was collected by vacuum filtration, washed using aqueous Na_2CO_3 (1 M) (20 mL) and extracted in EtOAc (2 x 20 mL). The combined organic layer was dried over $MgSO_4$ and reduced *in vacuo* to afford a neon yellow powder (63 mg, 22 %). m.p: 56-58 °C. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.26 (s, 1H, Ar), 7.60-7.55 (d, 1H, *J* = 8.0 Hz, Ar), 7.30-7.20 (t, 1H, *J* = 8.0 Hz, Ar), 7.09-7.04 (d, 1H, *J* = 8.0 Hz, Ar), 7.02 (s, 2H, NH_2), 6.96 (s, 1H, Ar), 2.32 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 169.1, 149.9, 137.4, 134.8, 128.3, 127.8, 126.2, 122.7, 101.3, 21.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 191.0565; Observed $[M+H]^+$ = 191.0637.

*Synthesis of 4-(3-nitrophenyl)thiazol-2-amine (**15**)*



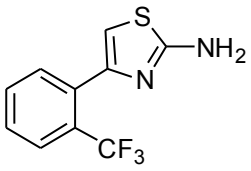
Applying the general procedure A, 3'-nitroacetophenone (200 mg, 1.21 mmol). The organic layer was dried over $MgSO_4$ and reduced to dryness *in vacuo* to afford title compound as a bright yellow powder (85 mg, 34 %). m.p: 188-190 °C. ^{1}H NMR (400 MHz; DMSO-d_6): δ 8.61 (s, 1H, Ar), 8.27-8.21 (m, 1H, Ar), 8.13-8.07 (m, 1H, Ar), 7.70-7.62 (t, 1H, *J* = 8.0 Hz, Ar), 7.34 (s, 1H, Ar), 7.22 (s, 2H, NH_2); ^{13}C NMR (101 MHz; DMSO-d_6): δ 169.6, 149.2, 147.4, 136.4, 131.5, 130.1, 121.7, 119.9, 104.3. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 222.0259; Observed $[M+H]^+$ = 222.0333.

*Synthesis of 4-(4-(trifluoromethyl)phenyl)thiazol-2-amine (**16**)*



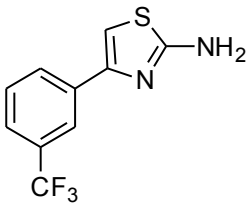
Applying the general procedure A, 4'-(trifluoromethyl)acetophenone (200 mg, 1.06 mmol). The organic layer was dried over $MgSO_4$ and reduced to dryness *in vacuo* to afford title compound as a light orange powder (39 mg, 15 %). m.p: 158-160 °C. ^{1}H NMR (400 MHz; $CDCl_3$): 8.00-7.80 (d, 2H, *J* = 8.0 Hz, Ar), 7.70-7.50 (d, 2H, *J* = 8.0 Hz, Ar), 6.84 (s, 1H, Ar), 5.1 (s, 2H, NH_2); ^{19}F NMR (400 MHz; $CDCl_3$) δ -62.5 (s, 3F, CF_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 169.5, 149.2, 139.5, 127.4 (q, *J* = 6.0 Hz), 126.0, 125.4 (q, *J* = 3.0 Hz), 124.5 (q, *J* = 272.0 Hz), 104.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 245.0282; Observed $[M+H]^+$ = 245.0353.

*Synthesis of 4-(2-(trifluoromethyl)phenyl)thiazol-2-amine (**17**)*



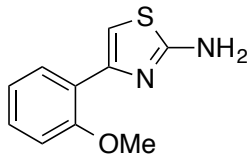
Applying the general procedure A, 2'-(trifluoromethyl)acetophenone (200 mg, 1.06 mmol, 159 µL). The organic layer was collected and dried over $MgSO_4$ and reduced to dryness *in vacuo* to yield a crystalline white title compound (48 mg, 20 %). m.p: 116-118 °C. ^{1}H NMR (400 MHz; MeOD): δ 7.80-7.70 (m, 1H, Ar), 7.70-7.60 (m, 1H, Ar), 7.60-7.50 (m, 2H, Ar), 6.52 (s, 1H, Ar); ^{19}F NMR (400 MHz; MeOD) δ -59.2 (s, 3F, CF_3); ^{13}C NMR (101 MHz; MeOD): δ 170.8, 149.4, 136.2 (q, *J* = 2.0 Hz), 133.2, 132.8 (q, *J* = 1.0 Hz), 129.8 (q, *J* = 1.0 Hz), 129.4 (d, *J* = 21.0 Hz), 127.1 (d, *J* = 6.0 Hz), 125.5 (q, *J* = 223.1 Hz), 106.6 (q, *J* = 3.0 Hz). *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 245.2436; Observed $[M+H]^+$ = 245.0353.

*Synthesis of 4-(3-(trifluoromethyl)phenyl)thiazol-2-amine (**18**)*



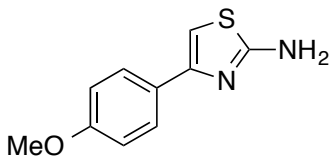
Applying the general procedure A, 3'-(trifluoromethyl)acetophenone (200 mg, 1.06 mmol, 162 µL). The reaction mixture was then reduced to dryness *in vacuo* and purified via column chromatography using a step gradient from Hexane: EtOAc (2: 1) to EtOAc: MeOH (9: 1) affording the title compound as a golden yellow solid (59 mg, 23 %). m.p: 78-80 °C. ^{1}H NMR (400 MHz; MeOD): δ 8.04 (s, 1H, Ar), 8.00-7.90 (d, 1H, *J* = 4.0 Hz, Ar), 7.60-7.40 (m, 2H, Ar), 6.94 (s, 1H, Ar); ^{19}F NMR (400 MHz; MeOD) δ -64.1 (s, 3F, CF_3); ^{13}C NMR (101 MHz; MeOD): δ 171.4, 149.6, 136.8 (d, *J* = 1.0 Hz), 131.9, 130.3 (d, *J* = 31.0 Hz), 130.2 (d, *J* = .0 Hz), 126.0 (d, *J* = 3.0 Hz), 124.9, 123.5 (q, *J* = 222.0 Hz), 109.8. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 245.0282; Observed $[M+H]^+$ = 245.0355.

*Synthesis of 4-(2-methoxyphenyl)thiazol-2-amine (**19**)*



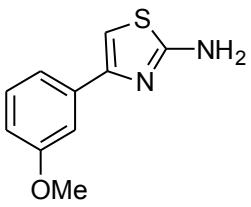
2'-Methoxyacetophenone (500 mg, 3.3 mmol) was added to $CHCl_3$ at 0 °C. Br_2 (629 mg, 3.5 mmol) was added drop-wise to the solution over a 5 min period. The solution was allowed to cool to rt upon completion of the addition. The vial was sealed and heated at 50 °C for 2 h. The mixture was neutralized using sat. Na_2CO_3 (20 mL), extracted in $CHCl_3$ (20 mL) and washed using brine solution. The organic layer was dried over $MgSO_4$ and reduced *in vacuo* to afford crude material (800 mg). The crude compound (800 mg) was added to MeCN in a microwave vial. Once dissolved, Thiourea (279 mg, 3.7 mmol) was added and the vial was sealed and heated at 100 °C for 2 h. The completion of the reaction afforded a white precipitate that was collected via vacuum filtration. The solid was washed using sat. Na_2CO_3 (20 mL), extracted in EtOAc (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant compound was purified via column chromatography using (EtOAc: petroleum spirits = 1:5) to afford title compound (170 mg, 25 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 8.03-8.00 (dd, 1H, *J* = 7.7 Hz, 1.8 Hz, Ar), 7.25-7.21 (m, 1H, Ar), 7.11 (s, 1H, Ar), 7.06-7.04 (dd, 1H, *J* = 8.2 Hz, 1.1 Hz, Ar), 6.98-6.94 (m, 1H, Ar), 6.95 (s, 2H, NH_2), 3.87 (s, 3H, Me). ^{13}C NMR (101 MHz; DMSO-d_6): δ 166.2, 156.5, 145.7, 129.2, 128.0, 123.0, 120.2, 111.4, 105.7, 55.3. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 207.0514; Observed $[M+H]^+$ = 207.0587.

*Synthesis of 4-(4-methoxyphenyl)thiazol-2-amine (**20**)*



4'-Methoxyacetophenone (500 mg, 3.3 mmol) was added to $CHCl_3$ (3 mL) at 0 °C. Br_2 (629 mg, 3.5 mmol) was added drop-wise to the solution over a period of 5 min. The solution was allowed to cool to rt upon completion of the addition. The vial was sealed and heated at 50 °C for 2 h. The mixture was neutralized using sat. Na_2CO_3 (30 mL) extracted in $CHCl_3$ (20 mL) and washed using brine solution (20 mL). The organic layer was dried over $MgSO_4$ and reduced *in vacuo* to yield crude material (744 mg). The crude compound (744 mg) was added to MeCN (3 mL) in a microwave vial. Once dissolved, thiourea (260 mg, 3.4 mmol) was added and the vial was sealed and heated at 100 °C for 2 h. A white precipitate was collected via vacuum filtration once reaction was completed. The solid was washed using sat. Na_2CO_3 (30 mL), extracted in EtOAc (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. Once dry, compound was purified via column chromatography using (DCM: EtOAc = 6: 1) to afford a yellow amorphous solid (92 mg, 13 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.72-7.70 (d, 2H, J = 8.9 Hz, Ar), 6.98 (s, 2H, NH_2), 6.92-6.90 (d, 2H, J = 8.9 Hz, Ar), 6.81 (s, 1H, ArH), 3.76 (s, 3H, Me). ^{13}C NMR (101 MHz; DMSO-d_6): δ 168.0, 158.4, 149.6, 127.8, 126.8, 113.7, 99.3, 55.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 207.0514; Observed $[M+H]^+$ = 207.0589.

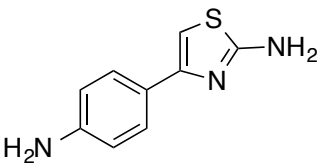
*Synthesis of 4-(3-methoxyphenyl)thiazol-2-amine (**21**)*



3'-Methoxyacetophenone (500 mg, 3.33 mmol, 457 µL) was dissolved in $CHCl_3$ (3 mL). To the mixture, Br_2 (629 mg, 3.50 mmol, 203 µL) was added drop-wise over 10 min whilst stirred at 0 °C. The mixture was then heated at 50 °C for 2 h. Once elapsed, the mixture was cooled to rt and then quenched using aqueous Na_2CO_3 (1 M) and then extracted using $CHCl_3$. The organic layer was collected, washed using brine solution, and dried over $MgSO_4$. The organic phase was reduced *in vacuo*, affording a yellow-viscous oil. This oil was dissolved in MeCN and stirred at rt. Thiourea (146 mg, 1.92 mmol) was then added and the mixture heated at 100 °C for 2 h. The white precipitate was collected by vacuum filtration, washed using aqueous Na_2CO_3 (1 M) (40 mL) and dissolved in EtOAc (20 mL). The organic layer was dried over $MgSO_4$ and reduced to dryness *in vacuo* to yield yellow rod-like crystals (95 mg, 14 %). m.p: 98-100 °C. ^{1}H NMR (400 MHz; MeOD): δ 7.35-7.30 (m, 2H, Ar), 7.30-7.20 (m, 1H, J = 8 Hz, Ar), 6.90-6.82 (m, 1H, Ar), 6.81 (s, 1H, Ar), 3.83 (s, 3H,

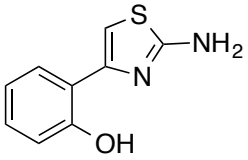
CH_3); ^{13}C NMR (101 MHz; MeOD): δ 171.2, 161.4, 142.1, 137.6, 130.0, 120.3, 114.8, 114.5, 109.8, 56.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 207.0514; Observed $[M+H]^+$ = 207.0585.

Synthesis of 4-(4-aminophenyl)thiazol-2-amine ***(22)***



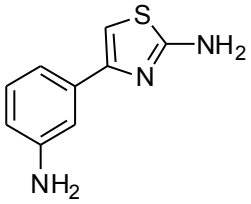
Applying the general procedure B, 4'-aminoacetophenone (200 mg, 1.5 mmol). The crude material was purified via column chromatography using ($CHCl_3$: EtOAc = 1: 1) to yield title compound (60 mg, 21 %). R_f: 0.2. ^1H NMR (400 MHz; DMSO-d_6): δ 7.45 (d, 2H, *J* = 8.6 Hz, Ar), 6.86 (s, 2H, NH_2), 6.56 (s, 1H, Ar), 6.52 (d, 2H, *J* = 8.6 Hz, Ar), 5.13 (s, 2H, NH_2). ^{13}C NMR (101 MHz; DMSO-d_6): δ 167.6, 150.8, 148.0, 126.5, 123.2, 113.5, 96.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 192.0517; Observed $[M+H]^+$ = 192.0590.

Synthesis of 2-(2-aminothiazol-4-yl)phenol ***(23)***



Applying the general procedure B, 2'-hydroxyacetophenone (179 µL, 200 mg, 1.5 mmol). The crude material was purified via column chromatography using (DCM: MeOH: NH_4OH = 99: 1: (2 drops)) to yield title compound (67 mg, 23 %). R_f: 0.3. ^1H NMR (400 MHz; DMSO-d_6): δ 11.96 (s, 1H, OH), 7.69 (dd, 1H, *J* = 7.8 Hz, 1.4 Hz, Ar), 7.47 (s, 2H, NH_2), 7.12 (m, 1H, Ar), 7.07 (s, 1H, Ar), 6.83-6.77 (m, 2H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 168.3, 155.4, 147.1, 128.9, 126.2, 118.8, 118.2, 116.9, 100.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 193.0357; Observed $[M+H]^+$ = 193.0431.

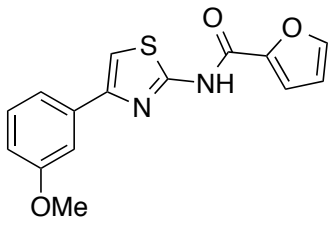
Synthesis of 4-(3-aminophenyl)thiazol-2-amine ***(24)***



3'-Aminoacetophenone (500 mg, 3.70 mmol) was dissolved in $CHCl_3$ (3 mL). To the reaction mixture, Br_2 (698 mg, 3.88 mmol, 225 µL) was added drop-wise over 10 min whilst stirred at 0 °C. The mixture was then heated at 50 °C for 2 h. Once elapsed, mixture was cooled to rt and then quenched using aqueous Na_2CO_3 (1 M) and then extracted using $CHCl_3$. The organic layer was collected and washed using brine solution (10 mL). The organic portion was collected and dried over $MgSO_4$ and reduced to dryness *in vacuo* to afford a cream-coloured powder. The crude powder was dissolved in MeCN (4 mL) at rt. Once completely solubilized, thiourea (295 mg, 3.88 mmol) was added and mixture heated at 100 °C for 2 h. Upon heating the mixture, a creamy white solid precipitated and

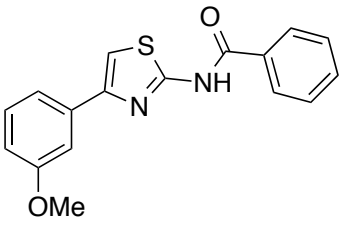
filtered via vacuum filtration. The precipitate was washed using aqueous Na_2CO_3 (1 M) (30 mL), extracted in EtOAc (15 mL), dried over $MgSO_4$ and then reduced to dryness *in vacuo* to afford a white powder. The compound was purified using column chromatography using a step gradient from Hexane: EtOAc (1: 1) to EtOAc to yield a light brown solid (36 mg, 5 %). m.p: 160-163 °C. ^{1}H NMR (400 MHz; MeOD): δ 7.15-7.05 (m, 3H, Ar), 6.71 (s, 1H, Ar), 6.69-6.65 (m, 1H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 167.8, 150.7, 149.6, 135.5, 128.8, 113.4, 113.0, 111.4, 100.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 192.0517; Observed $[M+H]^+$ = 192.0432.

Synthesis of N-(4-(3-methoxyphenyl)thiazol-2-yl)furan-2-carboxamide



2-Furoic acid (109 mg, 1 mmol), EDC.HCl (280 mg, 2 mmol) and HOBt (197 mg, 2 mmol) were added to a microwave vial. DMF (3 mL) and Et_3N (270 µL, 196 mg, 2 mmol) was added to the mixture followed by 4-(3-methoxyphenyl)thiazol-2-amine (100 mg, 0.5 mmol). The vial was sealed and heated at 100 °C for 3 h. The mixture was diluted with H_2O (2 x 20 mL) and extracted in EtOAc (2 x 20 mL). Organic layer was washed using brine solution, dried over $MgSO_4$ and reduced *in vacuo*. Resultant crude was purified via column chromatography using (Petroleum: EtOAc = 4: 1) to afford a brown oil (51 mg, 34 %). R_f: 0.2. ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.89 (s, 1H, NH), 7.56-7.55 (m, 1H, Ar), 7.43-7.41 (m, 2H, Ar), 7.36-7.30 (m, 2H, Ar), 7.18 (s, 1H, Ar), 6.90-6.87 (m, 1H, Ar), 6.59 (dt, 1H, *J* = 3.5 Hz, 1.7 Hz, Ar), 3.87 (s, 3H, Me); ^{13}C NMR (101 MHz; $CDCl_3$): δ 160.1, 157.2, 155.2, 149.8, 145.9, 145.6, 135.4, 129.9, 118.6, 117.4, 114.2, 113.1, 111.5, 108.4, 55.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 301.0569; Observed $[M+H]^+$ = 301.0642.

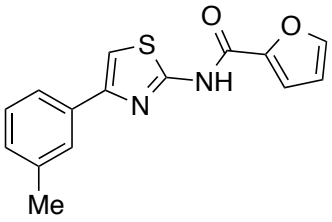
Synthesis of N-(4-(3-methoxyphenyl)thiazol-2-yl)benzamide



Benzoic acid (118 mg, 1 mmol), EDC.HCl (280 mg, 2 mmol) and HOBt (197 mg, 2 mmol) were added to a microwave vial. DMF (3 mL) and Et_3N (270 µL, 196 mg, 2 mmol) was added to the mixture followed by 4-(3-methoxyphenyl)thiazol-2-amine (100 mg, 0.5 mmol). The vial was sealed and heated at 100 °C for 3 h. The resultant mixture was diluted using H_2O (2 x 20 mL) and extracted in EtOAc (2 x 20 mL). Organic layer was washed using brine solution, dried over $MgSO_4$ and reduced *in vacuo*. Resultant crude was purified via column chromatography using (Petroleum: EtOAc = 4: 1 (plus a few drops of AcOH)) affording a

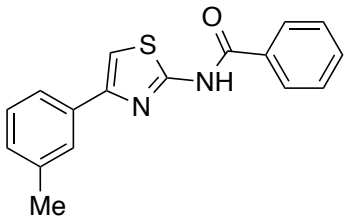
yellow oil (72 mg, 46 %). R_f: 0.2. ^{1}H NMR (400 MHz; $CDCl_3$): δ 10.5 (s, 1H, NH), 8.01-7.99 (m, 2H, Ar), 7.63-7.59 (m, 1H, Ar), 7.54-7.50 (m, 2H, Ar), 7.39-7.36 (m, 2H, Ar), 7.35-7.30 (m, 1H, Ar), 7.17 (s, 1H, Ar), 6.90-6.81 (m, 1H, Ar), 3.87 (s, 3H, Me); ^{13}C NMR (101 MHz; $CDCl_3$): δ 164.6, 160.0, 158.9, 149.0, 134.7, 133.1, 131.5, 129.9, 129.0, 127.6, 118.5, 114.3, 111.5, 108.2, 55.3. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 311.0776; Observed $[M+H]^+$ = 311.0848.

Synthesis of N-(4-(m-tolyl)thiazol-2-yl)furan-2-carboxamide



2-Furoic acid (118 mg, 1.1 mmol), EDC.HCl (404 mg, 2.1 mmol) and HOBt (284 mg, 2.1 mmol) were added to a microwave vial. DMF (3 mL) and Et_3N (293 µL, 213 mg, 2.1 mmol) was added to the mixture followed by 4-(*m*-tolyl)thiazol-2-amine (100 mg, 0.5 mmol). The vial was sealed and heated at 100 °C for 3 h. The mixture was diluted with H_2O (2 x 20 mL) and extracted in EtOAc (2 x 20 mL). Organic layer was washed using brine solution, dried over $MgSO_4$ and reduced *in vacuo*. Resultant crude was purified via column chromatography using (Toluene: EtOAc: AcOH = 75: 20: 5) to afford a brown oil (74 mg, 52 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 10.35 (s, 1H, NH), 7.65 (m, 1H, Ar), 7.62-7.60 (m, 1H, Ar), 7.54 (m, 1H, Ar), 7.37-7.35 (dd, 1H, *J* = 0.8 Hz, 0.8 Hz, Ar), 7.30 (t, 1H, *J* = 7.6 Hz, Ar), 7.16-7.14 (m, 2H, Ar), 6.58 (dd, 1H, *J* = 1.6 Hz, 3.6 Hz, Ar), 2.4 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 163.6, 157.5, 153.3, 150.3, 145.6, 138.5, 129.0, 128.8, 127.2, 126.9, 123.3, 117.2, 113.0, 108.1, 21.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 285.0619; Observed $[M+H]^+$ = 285.0691.

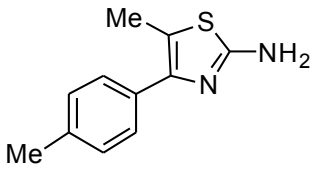
Synthesis of N-(4-(m-tolyl)thiazol-2-yl)benzamide



Benzoic acid (128 mg, 1 mmol), EDC.HCl (404 mg, 2.1 mmol) and HOBt (284 mg, 2.1 mmol) were added to a microwave vial. DMF (3 mL) and Et_3N (293 µL, 213 mg, 2.1 mmol) was added to the mixture followed by 4-(3-methoxyphenyl)thiazol-2-amine (100 mg, 0.5 mmol). The vial was sealed and heated at 100 °C for 3 h. The mixture was diluted with H_2O (2 x 20 mL) and extracted in EtOAc (2 x 20 mL). Organic layer was washed using brine solution, dried over $MgSO_4$ and reduced *in vacuo*. Resultant crude oil was purified via column chromatography using (Toluene: EtOAc: AcOH = 75: 20: 5) to afford a yellow oil (86 mg, 58 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 11.64 (s, 1H, NH), 8.03-8.00 (m, 2H, Ar), 7.60-7.48

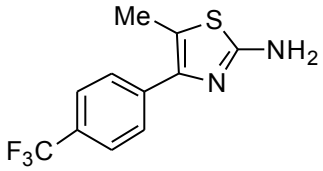
(m, 5H, Ar), 7.30 (t, 1H, *J* = 7.6 Hz, Ar), 7.16-7.13 (m, 1H, Ar), 7.13 (s, 1H, Ar), 2.4 (s, 3H, Me); ^{13}C NMR (101 MHz; $CDCl_3$): δ 176.8, 165.1, 159.5, 149.9, 138.5, 133.8, 133.0, 129.1, 129.0, 128.8, 127.8, 127.1, 123.5, 108.2, 21.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 295.0827; Observed $[M+H]^+$ = 295.0901.

*Synthesis of 5-methyl-4-(p-tolyl)thiazol-2-amine (**25**)*



Applying the general procedure A, 4'-methylpropiophenone (200 mg, 1.35 mmol, 202 μL). The resultant white powder collected was collected by vacuum filtration and washed using aqueous Na_2CO_3 (1 M) (30 mL). The organic portion was extracted in EtOAc (15 mL), dried over $MgSO_4$ and then reduced to dryness *in vacuo*. The procedure affords the title compound as a white powder (61 mg, 22 %). m.p: 144-146 °C. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.48-7.42 (d, 2H, *J* = 8.0 Hz, Ar), 7.22-7.16 (d, 2H, *J* = 8.0 Hz, Ar), 6.74 (s, 2H, NH_2), 2.32 (s, 3H, CH_3), 2.31 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 164.1, 145.0, 135.8, 132.7, 128.7, 127.8, 113.9, 20.8, 12.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 205.0721; Observed $[M+H]^+$ = 205.0794.

*Synthesis of 5-methyl-4-(4-(trifluoromethyl)phenyl)thiazol-2-amine (**26**)*



Applying the general procedure A, 4'-(trifluoromethyl)propiophenone (200 mg, 0.99 mmol). The light yellow solid was purified via column chromatography using a step gradient from Hexane: EtOAc (1: 1) to EtOAc affording the title compound as a pale yellow solid (80 mg, 33%). m.p: 122-124 °C. R_f: 0.5. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.82-7.76 (d, 2H, *J* = 8.0 Hz, Ar), 7.76-7.70 (d, 2H, *J* = 8.0 Hz, Ar), 6.87 (s, 2H, NH_2), 2.37 (s, 3H, CH_3); ^{19}F NMR (400 MHz; DMSO-d_6) δ -60.9 (s, CF_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 164.5, 143.4, 139.3, 128.3, 126.7 (q, *J* = 32.0 Hz), 125.0 (q, *J* = 4.0 Hz), 124.5 (q, *J* = 272.0 Hz), 117.2, 12.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 259.0439; Observed $[M+H]^+$ = 259.0512.

5.2.2 4-Methyl-5-aryl substituted aminothiazole analogues

General procedure for 4-methyl-5-aryl substituted 2-aminothiazole series

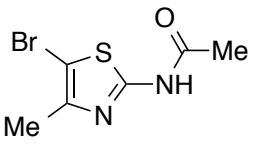
A solution of *N*-(5-iodo-4-methylthiazol-2-yl)acetamide (**28**) (200 mg, 0.7 mmol) and substituted phenylboronic acid in THF (3 mL) was made up in a microwave vial. The mixture was stirred at rt

and degasified with N_2 for approximately 10 min. Aqueous Na_2CO_3 (1 M) was added to the reaction mixture during this period. Before the de-gas period has elapsed, $Pd(Cl)_2(PPh_3)_2$ was added and the vial was sealed in order to mimic microwave conditions. The reaction mixture was heated for 60 min at 100 °C (temp. used varies depending on boronic acid reactivity). Once complete, the mixture was cooled to rt and the organic phase was passed through an improvised pipette plug of cotton wool and $MgSO_4$ to collect any Pd waste and dry the organic layer. EtOAc was used to complete organic extraction. The organic layer was reduced *in vacuo* and purified via column chromatography. Shane Devine of the Scammells Research group synthesized two compounds included in study using this adopted general procedure.

General procedure for deprotection of 4-methyl-5-aryl substituted 2-aminothiazole series

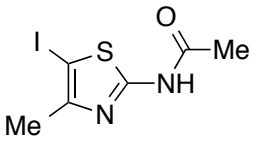
Resultant acetyl protected 2-aminothiazole compound was dissolved in a solution of EtOH: H_2O (1: 1). Once dissolved, NaOH (1 M) solution (3 mL) was added to vial, vial sealed and heated at 100 °C overnight. Mixture was reduced *in vacuo* and purified via column purification.

Synthesis of N-(5-bromo-4-methylthiazol-2-yl)acetamide



2-acetamido-4-methylthiazole (**27** was commercially available) (600 mg, 3.8 mmol) was dissolved in AcOH. NBS (684 mg, 3.8 mmol) was slowly added to mixture and stirred at rt for 1.5 h. After 1.5 h, the resultant white solid was filtered and dried. Once dry, the solid was dissolved in DCM (50 mL) and partitioned between sat. $NaHCO_3$ solution (2 x 100 mL). The organic layer was collected and finally washed using brine (100 mL), dried over $MgSO_4$ and reduced *in vacuo*. The crude solid (416 mg) was purified via column chromatography using (Petroleum: EtOAc = 60: 40) to afford title compound (300 mg, 33 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.24 (s, 1H, NH), 2.22 (s, 3H, Me), 2.13 (s, 3H, Me).

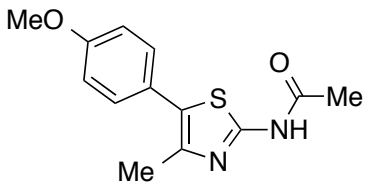
*Synthesis of N-(5-iodo-4-methylthiazol-2-yl)acetamide (**28**)*



2-Acetamido-4-methylthiazole (**25** was commercially available) (1 g, 6.4 mmol) was added to MeCN (3 mL). Once dissolved, NIS (3 g, 13.4 mmol) was slowly added to the reaction mixture and stirred for 15 min at rt. A white precipitate resulted after 10 min of stirring. The solid was collected via vacuum filtration to afford a white fluffy solid (958 mg). The resultant mother liquor was reduced *in vacuo*. The afforded solid

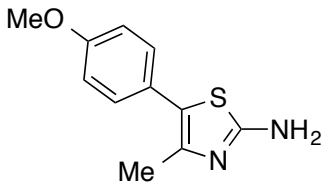
was taken up in EtOAc and washed using 1M $Na_2S_2O_3$ solution, organic layer was dried over $MgSO_4$ and reduced *in vacuo*. Resultant solid was suspended in MeCN and filtered to afford title compound (575 mg, 85 % combined). ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.20 (s, 1H, NH), 2.25 (s, 3H, Me), 2.11 (s, 3H, Me).

Synthesis of N-(5-(4-methoxyphenyl)-4-methylthiazol-2-yl)acetamide



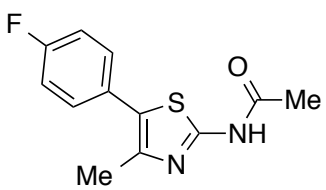
Applying the general method for Suzuki coupling above, 4-methoxyphenylboronic acid (113 mg, 0.74 mmol). The reaction mixture was heated for 2 h at 100 °C. This compound was purified via column chromatography using (Hexane: EtOAc = 2: 1) to afford title compound (115 mg, 62 %). R_f: 0.2. ^{1}H NMR (400 MHz; $CDCl_3$): δ 11.6 (s, 1H, NH), 7.38-7.30 (d, 2H, *J* = 8.0 Hz, Ar), 6.98-6.90 (d, 2H, *J* = 8 Hz, Ar), 3.8 (s, 3H, OCH_3), 2.4 (s, 3H, CH_3), 2.2 (s, 3H, CH_3).

*Synthesis of 5-(4-methoxyphenyl)-4-methylthiazol-2-amine (**29**)*



Applying the general deprotection method above, *N*-(5-(4-methoxyphenyl)-4-methylthiazol-2-yl)acetamide (115 mg, 0.52 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (Petroleum: EtOAc = 1: 1) to afford a light reddish solid (56 mg, 58 %). m.p: 164-166 °C. R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.32-7.24 (d, 2H, *J* = 8.0 Hz, Ar), 6.95-6.87 (d, 2H, *J* = 8.0 Hz, Ar), 4.98 (s, 2H, NH_2), 3.83 (s, 3H, CH_3), 2.26 (s, 3H, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 164.8, 158.8, 142.7, 130.3, 125.1, 121.3, 114.2, 55.4, 16.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 221.0670; Observed $[M+H]^+$ = 221.0739.

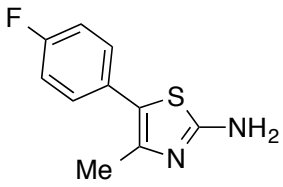
Synthesis of N-(5-(4-fluorophenyl)-4-methylthiazol-2-yl)acetamide



Applying the general method for Suzuki coupling above, 4-fluorophenylboronic acid (104 mg, 0.74 mmol). This compound was purified via column chromatography using (Hexane: EtOAc = 1: 1) to afford an orange solid (128 mg, 72 %). R_f: 0.2. ^{1}H NMR (400 MHz;

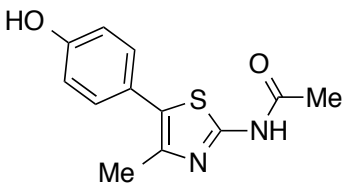
$CDCl_3$): δ 11.2 (s, 1H, NH), 7.35-7.30 (m, 2H, Ar), 7.15-7.05 (m, 2H, Ar), 2.3 (s, 3H, CH_3), 2.2 (s, 3H, CH_3).

*Synthesis of 5-(4-fluorophenyl)-4-methylthiazol-2-amine (**30**)*



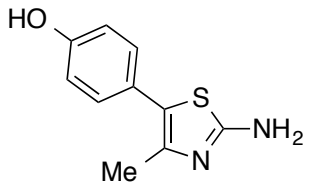
Applying the general deprotection method above, *N*-(5-(2-methoxyphenyl)-4-methylthiazol-2-yl)acetamide (128 mg, 0.61 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (Hexane: EtOAc = 1: 1) to afford a light reddish solid (40 mg, 45 %). R_f: 0.2. m.p: 132-134 °C . ^{1}H NMR (400 MHz; MeOD): δ 7.38-7.30 (m, 2H, Ar), 7.14-7.06 (m, 2H, Ar), 2.18 (s, 3H, CH_3); ^{19}F NMR (400 MHz; MeOD) δ -114.9 (s, 1F, F); ^{13}C NMR (101 MHz; MeOD): δ 169.1, 163.1 (d, J = 246.0 Hz), 143.5, 131.6 (d, J = 8.0 Hz), 130.4 (d, J = 3.0 Hz), 119.2, 116.4 (d, J = 22.0 Hz), 15.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 209.0470; Observed $[M+H]^+$ = 209.0543.

Synthesis of N-(5-(4-hydroxyphenyl)-4-methylthiazol-2-yl)acetamide



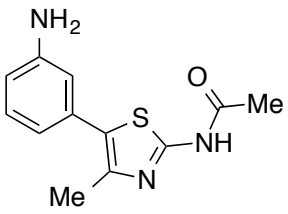
Applying the general method for Suzuki coupling above, 4-hydroxyphenylboronic acid (103 mg, 0.74 mmol). This compound was purified via column chromatography using (Hexane: EtOAc = 1: 1) to afford title compound (68 mg, 39 %). R_f: 0.5. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.10 (s, 1H, OH), 9.62 (s, 1H, NH), 7.31-7.22 (d, 2H, J = 8.0 Hz, Ar), 6.90-6.80 (d, 2H, J = 8.0 Hz, Ar), 2.30 (s, 3H, CH_3), 2.13 (s, 3H, CH_3).

*Synthesis of 4-(2-amino-4-methylthiazol-5-yl)phenol (**31**)*



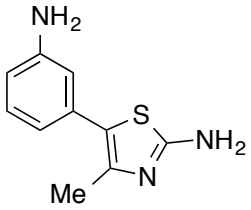
Applying the general deprotection method above, *N*-(5-(4-hydroxyphenyl)-4-methylthiazol-2-yl)acetamide (68 mg, 0.32 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (Petroleum: EtOAc = 1: 1) to afford a light reddish solid (45 mg, 80 %). R_f: 0.3. m.p: 232-234 °C. ^{1}H NMR (400 MHz; MeOD): δ 7.21-7.14 (d, 2H, J = 8.0 Hz, Ar), 6.85-6.75 (d, 2H, J = 8.0 Hz, Ar), 2.17 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 164.9, 155.9, 141.5, 129.4, 123.6, 117.8, 115.4, 15.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 207.0514; Observed $[M+H]^+$ = 207.0589.

Synthesis of N-(5-(3-aminophenyl)-4-methylthiazol-2-yl)acetamide



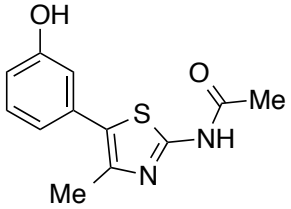
Compound synthesized by Shane Devine applying the general method for Suzuki coupling above, 3-aminophenylboronic acid (108 mg). (Unpublished data)

*Synthesis of 5-(3-aminophenyl)-4-methylthiazol-2-amine (**32**)*



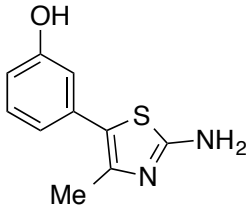
Compound was synthesized by Shane Devine applying the general deprotection method above, *N*-(5-(3-aminophenyl)-4-methylthiazol-2-yl)acetamide. (Unpublished data)

Synthesis of N-(5-(3-hydroxyphenyl)-4-methylthiazol-2-yl)acetamide



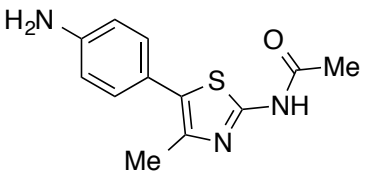
Applying the general method for Suzuki coupling above, 3-hydroxyphenylboronic acid (108 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 1: 1) to afford a brown amorphous solid (30 mg, 18 %). R_f: 0.1. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.25 (t, 1H, *J* = 16.0 Hz, Ar), 6.96-6.92 (m, 2H, Ar), 6.86-6.83 (m, 1H, Ar), 2.40 (s, 3H, Me), 2.25 (s, 3H, Me).

*Synthesis of 3-(2-amino-4-methylthiazol-5-yl)phenol (**33**)*



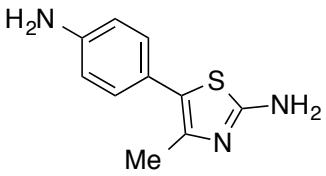
Applying the general deprotection method above, *N*-(5-(3-hydroxyphenyl)-4-methylthiazol-2-yl)acetamide (30 mg, 0.1 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (DCM: MeOH = 95: 5) to afford an orange amorphous solid (19 mg, 92 %). R_f: 0.2. ^{1}H NMR (400 MHz; MeOD): δ 7.19-7.15 (t, 1H, *J* = 15.2 Hz, Ar), 6.81-6.78 (m, 2H, Ar), 6.70-6.68 (m, 1H, Ar), 2.22 (s, 3H, Me); ^{13}C NMR (101 MHz; MeOD): δ 168.8, 158.5, 142.3, 134.9, 130.6, 120.8, 120.4, 116.3, 114.9, 15.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 207.0514; Observed $[M+H]^+$ = 207.0587.

Synthesis of N-(5-(4-aminophenyl)-4-methylthiazol-2-yl)acetamide



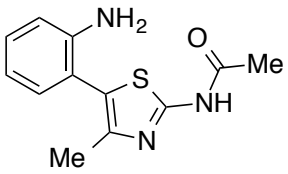
Applying the general method for Suzuki coupling above, 4-aminophenylboronic acid pinacol ester (171 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 1: 1) to afford title compound (79 mg, 46 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.93 (s, 1H, NH), 7.08 (d, 2H, *J* = 8.4 Hz, Ar), 6.60 (d, 2H, *J* = 8.4 Hz, Ar), 2.26 (s, 3H, Me), 2.10 (s, 3H, Me).

*Synthesis of 5-(4-aminophenyl)-4-methylthiazol-2-amine (**34**)*



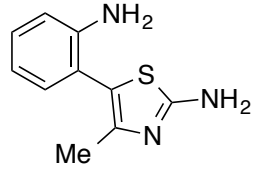
Applying the general deprotection method above, *N*-(5-(4-aminophenyl)-4-methylthiazol-2-yl)acetamide (79 mg, 0.3 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (DCM: MeOH: NH_4OH = 95: 5: (2-3 drops)) to afford a orange/brown amorphous solid (17 mg, 28 %). R_f: 0.2. ^{1}H NMR (400 MHz; MeOD): δ 7.07 (d, 2H, *J* = 8.8 Hz, Ar), 6.71 (d, 2H, *J* = 8.8 Hz, Ar), 2.15 (s, 3H, Me); ^{13}C NMR (101 MHz; MeOD): δ 168.2, 148.1, 141.1, 130.7, 123.0, 121.2, 116.3, 15.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 206.0674; Observed $[M+H]^+$ = 206.0747.

Synthesis of N-(5-(2-aminophenyl)-4-methylthiazol-2-yl)acetamide



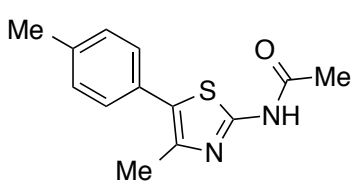
Compound synthesized by Shane Devine applying the general method for Suzuki coupling above, 2-aminophenylboronic acid (108mg). (Unpublished data)

*Synthesis of 5-(2-aminophenyl)-4-methylthiazol-2-amine (**35**)*



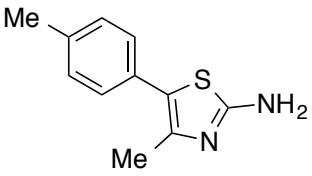
Compound was synthesized by Shane Devine applying the general deprotection method above, *N*-(5-(2-aminophenyl)-4-methylthiazol-2-yl)acetamide. (Unpublished data)

Synthesis of N-(4-methyl-5-(p-tolyl)thiazol-2-yl)acetamide



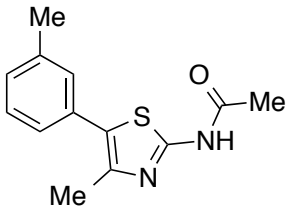
Applying the general method for Suzuki coupling above, 4-tolyboronic acid (106 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford title compound (115 mg, 67 %). R_f: 0.3. ^{1}H NMR (400 MHz; MeOD): δ 7.32-7.30 (d, 2H, *J* = 8.4 Hz, Ar), 7.25-7.23 (d, 2H, *J* = 8.0 Hz, Ar), 2.37 (s, 3H, Me), 2.34 (s, 3H, Me), 2.21 (s, 3H, Me).

*Synthesis of 4-methyl-5-(p-tolyl)thiazol-2-amine (**36**)*



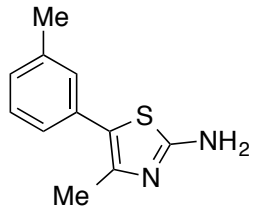
Applying the general deprotection method above, *N*-(4-methyl-5-(*p*-tolyl)thiazol-2-yl)acetamide (115 mg, 0.5 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using ($CHCl_3$: MeOH: NH_4OH = 97: 3: 2 drops) to afford title compound (35 mg, 34 %). R_f: 0.3. ^{1}H NMR (400 MHz; MeOD): δ 7.24-7.17 (m, 4H, Ar), 2.34 (s, 3H, Me), 2.19 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 165.3, 142.6, 135.3, 130.1, 129.2, 127.8, 117.4, 20.6, 16.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 205.0723; Observed $[M+H]^+$ = 205.0796.

Synthesis of N-(4-methyl-5-(m-tolyl)thiazol-2-yl)acetamide



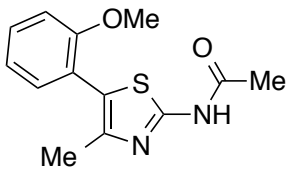
Applying the general method for Suzuki coupling above, 3-methylphenylboronic acid (106 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford title compound (98 mg, 57 %). R_f: 0.3. ^{1}H NMR (400 MHz; MeOD): δ 7.30 (t, 1H, *J* = 15.2 Hz, Ar), 7.24-7.20 (m, 2H, Ar), 7.16-7.14 (m, 1H, Ar), 2.37 (s, 3H, Me), 2.34 (s, 3H, Me), 2.21 (s, 3H, Me).

*Synthesis of 4-methyl-5-(m-tolyl)thiazol-2-amine (**37**)*



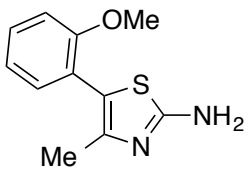
Applying the general deprotection method above, *N*-(4-methyl-5-(*m*-tolyl)thiazol-2-yl)acetamide (98 mg, 0.4 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (DCM: MeOH: NH_4OH = 97: 3: (2 drops)) to afford title compound (36 mg, 44 %). R_f: 0.2. ^{1}H NMR (400 MHz; MeOD): δ 7.24 (t, 1H, *J* = 15.2 Hz, Ar), 7.16-7.12 (m, 2H, Ar), 7.09-7.06 (m, 1H, Ar), 2.35 (s, 3H, Me), 2.21 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 165.5, 143.0, 137.8, 132.9, 128.5, 128.4, 126.7, 125.0, 117.5, 21.0, 16.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 205.0724; Observed $[M+H]^+$ = 205.0797.

Synthesis of N-(5-(2-methoxyphenyl)-4-methylthiazol-2-yl)acetamide



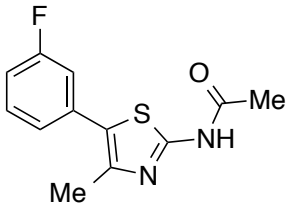
Applying the general method for Suzuki coupling above, 2-methoxyphenylboronic acid (119 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford title compound (95 mg, 51 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.98 (s, 1H, NH), 7.36-7.32 (m, 1H, Ar), 7.31-7.29 (dd, 1H, *J* = 0.8 Hz, 0.8 Hz, Ar), 7.02-6.96 (m, 2H, Ar), 3.57 (s, 3H, Me), 2.27 (s, 3H, Me), 2.24 (s, 3H, Me).

*Synthesis of 5-(2-methoxyphenyl)-4-methylthiazol-2-amine (**38**)*



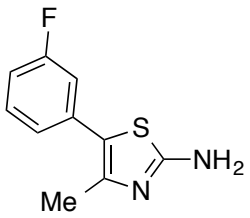
Applying the general deprotection method above, *N*-(5-(2-methoxyphenyl)-4-methylthiazol-2-yl)acetamide (95 mg, 0.3 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford a light reddish amorphous solid (61 mg, 77 %). R_f: 0.1. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.32-7.28 (m, 2H, Ar), 7.00-6.96 (m, 1H, Ar), 6.95-6.93 (m, 1H, Ar), 4.79 (s, 2H, NH_2), 3.83 (s, 3H, Me), 2.18 (s, 3H, Me); ^{13}C NMR (101 MHz; $CDCl_3$): δ 166.2, 157.0, 145.3, 132.1, 129.1, 121.4, 120.6, 116.8, 111.2, 55.6, 16.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 221.0671; Observed $[M+H]^+$ = 221.0745.

Synthesis of N-(5-(3-fluorophenyl)-4-methylthiazol-2-yl)acetamide



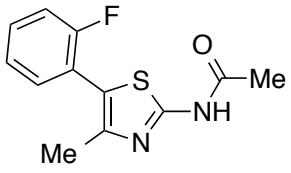
Applying the general method for Suzuki coupling above, 3-fluorophenylboronic acid (109 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford an orange amorphous solid (112 mg, 64 %). R_f: 0.3. ^{1}H NMR (400 MHz; MeOD): δ 7.47-7.41 (m, 1H, Ar), 7.25 (m, 1H, Ar), 7.20-7.16 (m, 1H, Ar), 7.10-7.05 (m, 1H, Ar), 2.37 (s, 3H, Me), 2.21 (s, 3H, Me).

*Synthesis of 5-(3-fluorophenyl)-4-methylthiazol-2-amine (**39**)*



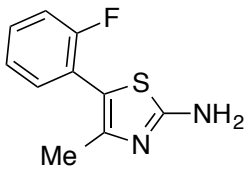
Applying the general deprotection method above, *N*-(5-(3-fluorophenyl)-4-methylthiazol-2-yl)acetamide (50 mg, 0.2 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford title compound (15 mg, 37 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.35-7.30 (m, 1H, Ar), 7.14-7.11 (ddd, 1H, Ar), 7.13-7.10 (m, 1H, Ar), 6.98-6.94 (m, 1H, Ar), 4.89 (s, 2H, NH_2), 2.31 (s, 3H, Me); ^{19}F NMR (400 MHz, $CDCl_3$): δ -112.82; ^{13}C NMR (101 MHz; $CDCl_3$): δ 165.2, 164.0 (d, J = 227.0 Hz), 161.6, 144.4, 135.0, 130.1 (d, J = 8.7 Hz), 124.5 (d, J = 2.7 Hz), 115.5 (d, J = 22.3 Hz), 113.7 (d, J = 21.1 Hz), 16.3. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 209.0470; Observed $[M+H]^+$ = 209.0543.

Synthesis of N-(5-(2-fluorophenyl)-4-methylthiazol-2-yl)acetamide



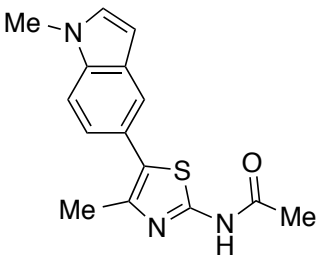
Applying the general method for Suzuki coupling above, 2-fluorophenylboronic acid (109 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford a yellow oil (50 mg, 29 %). R_f: 0.4. ^{1}H NMR (400 MHz; $CDCl_3$): δ 10.10 (s, 1H, NH), 7.39-7.32 (m, 2H, Ar), 7.21-7.19 (dd, 1H, J = 1.2 Hz, 1.2 Hz, Ar), 7.18-7.13 (m, 1H, Ar), 2.29 (s, 3H, Me), 2.25 (s, 3H, Me).

*Synthesis of 5-(2-fluorophenyl)-4-methylthiazol-2-amine (**40**)*



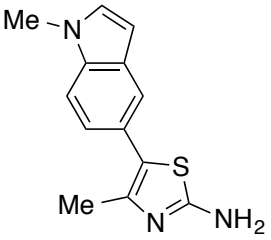
Applying the general deprotection method above, *N*-(5-(2-fluorophenyl)-4-methylthiazol-2-yl)acetamide (95 mg, 0.3 mmol). The reaction was cooled to rt and reduced *in vacuo*. The crude material was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford title compound (61 mg, 77 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.36-7.31 (m, 1H, Ar), 7.30-7.26 (m, 1H, Ar), 7.17-7.13 (m, 1H, Ar), 7.14-7.09 (m, 1H, Ar), 4.84 (s, 2H, NH_2), 2.20 (s, 3H, Me); ^{19}F NMR (400 MHz; $CDCl_3$): δ -113.32; ^{13}C NMR (101 MHz; DMSO-d_6): δ 167.2, 160.0 (d, *J* = 231.0 Hz), 157.6, 145.5, 131.7 (d, *J* = 3.3 Hz), 124.5 (d, *J* = 3.3 Hz), 116.0 (d, *J* = 22.1 Hz), 115.7 (d, *J* = 24.0 Hz), 109.6 (d, *J* = 11.1 Hz), 15.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 209.0471; Observed $[M+H]^+$ = 209.0545

Synthesis of N-(4-methyl-5-(1-methyl-1H-indol-5-yl)thiazol-2-yl)acetamide



Applying the general method for Suzuki coupling above, 1-methylindole-5-boronic acid pinacol ester (200 mg, 0.8 mmol). This compound was purified via column chromatography using (Toluene: EtOAc = 80: 20) to afford title compound (117 mg, 59 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.00 (s, 1H, NH), 7.59 (m, 1H, Ar), 7.49 (d, 1H, *J* = 8.4 Hz, Ar), 7.37 (d, 1H, *J* = 2.8 Hz, Ar), 7.22-7.19 (dd, 1H, *J* = 8.4 Hz, 1.7 Hz, Ar), 6.70 (s, 1H, Ar), 6.46 (dd, 1H, *J* = 3.0 Hz, 0.8 Hz, Ar), 3.80 (s, 3H, Me), 2.33 (s, 3H, Me), 2.32 (s, 3H, Me).

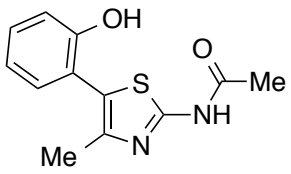
*Synthesis of 4-methyl-5-(1-methyl-1H-indol-5-yl)thiazol-2-amine (**41**)*



Applying the general deprotection method above, *N*-(4-methyl-5-(1-methyl-1*H*-indol-5-yl)thiazol-2-yl)acetamide (110 mg, 0.4 mmol). The reaction was cooled to rt where a precipitate will form. Precipitate was collected via vacuum filtration to afford title compound (16 mg, 16 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.49 (d, 1H, *J* = 1.1 Hz, Ar), 7.42 (d, 1H, *J* = 8.5 Hz, Ar), 7.33 (d, 1H, *J* = 3.0 Hz, Ar), 7.11 (dd, 1H, *J* = 8.4 Hz, 1.6 Hz, Ar), 6.77 (s, 2H, NH_2), 6.41 (dd, 1H, *J* = 3.0 Hz, 0.7 Hz, Ar), 3.78 (s, 3H, Me), 2.16 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ

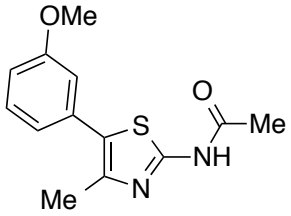
165.06, 141.30, 135.14, 130.36, 128.21, 123.63, 122.12, 120.02, 119.10, 109.79, 100.38, 32.52, 16.02. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 244.3280; Observed $[M+H]^+$ = 244.3364.

Synthesis of N-(5-(2-hydroxyphenyl)-4-methylthiazol-2-yl)acetamide



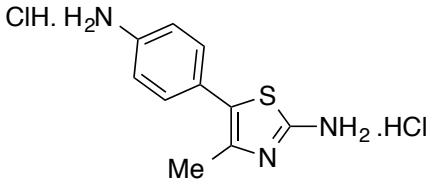
Applying the general method for Suzuki coupling above, 2-hydroxyphenylboronic acid (107 mg, 0.8 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 3: 1) to afford title compound (75 mg, 43 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.96 (s, 1H, OH), 7.22-7.15 (m, 2H, Ar), 6.93-6.91 (dd, 1H, *J* = 0.8 Hz, 0.8 Hz, Ar), 6.86- 6.82 (td, 1H, *J* = 1.2 Hz, 1.2 Hz, 1.2 Hz, Ar), 2.18 (s, 3H, Me), 2.11 (s, 3H, Me).

Synthesis of N-(5-(3-methoxyphenyl)-4-methylthiazol-2-yl)acetamide



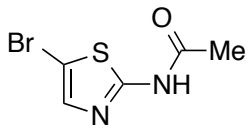
Applying the general method for Suzuki coupling above, 3-methoxyphenylboronic acid (162 mg, 1.1 mmol). This compound was purified via column chromatography using (Petroleum: EtOAc = 2: 1) to afford a brown oil (123 mg, 67 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 10.07 (s, 1H, NH), 7.34-7.30 (m, 1H, Ar), 7.04-7.01 (ddd, 1H, *J* = 7.6 Hz, 1.6 Hz, 0.94 Hz, Ar), 6.97-6.96 (dd, 1H, *J* = 2.4 Hz, 1.69 Hz, Ar), 6.89-6.86 (ddd, 1H, *J* = 8.3 Hz, 2.5 Hz, 0.9 Hz, Ar), 3.62 (s, 3H, Me), 2.42 (s, 3H, Me), 2.24 (s, 3H, Me).

Synthesis of 5-(4-aminophenyl)-4-methyloxazol-2-amine dihydrochloride



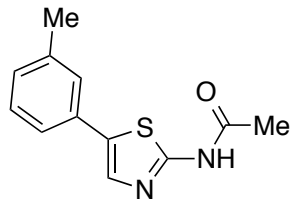
5-(4-Aminophenyl)-4-methylthiazol-2-amine (37 mg, 0.2 mmol) was dissolved in EtOAc (2 mL) and stirred at 0 °C. Once soluble, mixture was allowed to warm to rt, HCl in ether (2 mL) was then added drop-wise to reaction mixture and stirred for 15 min. A white precipitate formed after the addition of HCl in ether. Once reacted, mixture was reduced *in vacuo* and dried under the house vacuum for 1 h. The resultant solid was filtered via vacuum filtration and washed using ether to afford title compound (50 mg, 79 %). ^{1}H NMR (400 MHz; D_2O): δ 7.57-7.55 (d, 2H, *J* = 8.8 Hz, Ar), 7.45-7.42 (d, 2H, *J* = 8.8 Hz, Ar), 2.27 (s, 3H, Me).

Synthesis of N-(5-bromothiazol-2-yl)acetamide



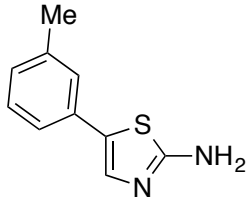
2-Amino-5-bromothiazole hydrobromide (500 mg, 1.9 mmol) was added to a sealed vial under N_2 gas. Ac_2O (1.1 mL, 11.5 mmol) was added along with pyridine (310 µL, 3.8 mmol) and stirred at 90 °C for 15 min. Once reacted, mixture was allowed to cool and was diluted using H_2O (10 mL). A pale pink amorphous solid was filtered and dried under the house vacuum overnight (358 mg, 84 %). ^{1}H NMR (400 MHz; $CDCl_3$): δ 12.35 (s, 1H, NH), 7.53 (s, 1H, Ar), 2.14 (s, 3H, Me).

Synthesis of N-(5-(m-tolyl)thiazol-2-yl)acetamide



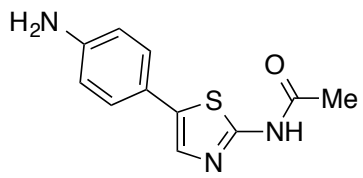
N-(5-Bromothiazol-2-yl)acetamide (200 mg, 0.9 mmol) and 3-methylphenyl boronic acid (185 mg, 1.3 mmol) were added to a mixture of dioxane (3 mL) and H_2O (1 mL). CsF (345 mg, 2.2 mmol) was also added to reaction mixture. Mixture was degassed using N_2 for a period of 5 min before $Pd(dppf)Cl_2$.DCM (37 mg, 0.04 mmol) was added and heated at 100 °C for 2 h in a sealed vial. Once reacted, organic layer was collected and pushed through an improvised cotton wool plug to remove any visible Pd by-product from resultant mother liquor. The resultant mother liquor was reduced *in vacuo* and purified via coloum chromatography using (Petroleum: EtOAc = 3: 1) to afford title compound (62 mg, 30 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 11.87 (s, 1H, NH), 7.58 (s, 1H, Ar), 7.39-7.36 (m, 2H, Ar), 7.31-7.27 (m, 1H, Ar), 7.15-7.12 (m, 1H, Ar), 2.39 (s, 3H, Me), 2.37 (s, 3H, Me).

Synthesis of 5-(m-tolyl)thiazol-2-amine



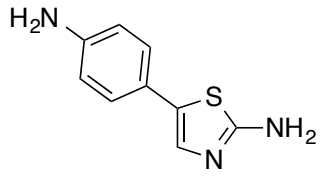
N-(5-(*m*-Tolyl)thiazol-2-yl)acetamide (62 mg , 0.2 mmol) was dissolved in EtOH (2 mL). NaOH (1 M) (2 mL) was added to the reaction mixture which was stirred at 100 °C overnight. Once reacted, the mixture was reduced *in vacuo* and purfied via column chromatography using (DCM: MeOH = 98: 2) to afford title compound (23 mg, 46 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.29 (s, 1H, Ar), 7.23-7.21 (m, 3H, Ar), 7.06-7.04 (m, 1H, Ar), 4.89 (s, 2H, NH_2), 2.36 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 167.9, 138.0, 135.0, 132.4, 128.7, 126.8, 125.1, 125.0, 121.8, 20.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 191.0566; Observed $[M+H]^+$ = 191.0639.

Synthesis of N-(5-(4-aminophenyl)thiazol-2-yl)acetamide



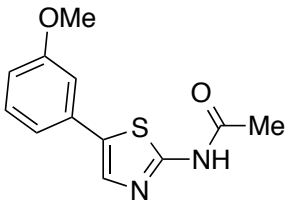
N-(5-Bromothiazol-2-yl)acetamide (200 mg, 0.9 mmol) and 4-aminophenyl boronic acid pinacol ester (299 mg, 1.3 mmol) were added to a mixture of dioxane (3 mL) and H_2O (1 mL). CsF (345 mg, 2.2 mmol) was also added to reaction mixture. Mixture was degassed using N_2 for a period of 5 min before Pd(dppf)Cl_2.DCM (37 mg, 0.04 mmol) was added and heated at 100 °C for 2 h in a sealed vial. Once reacted, organic layer was collected and pushed through an improvised cotton wool plug to remove any visible Pd by-product from resultant mother liquor. The mother liquor was reduced *in vacuo* and purified via coloum chromatography using (Petroleum: EtOAc = 1: 1) to afford title compound (23 mg, 21 %). R_f: 0.2. ^{1}H NMR (400 MHz; MeOD): δ 7.43 (s, 1H, NH), 7.3 (d, 2H, *J* = 8.7 Hz, Ar), 6.7 (d, 2H, *J* = 8.7 Hz, Ar), 2.2 (s, 3H, Me).

Synthesis of 5-(4-aminophenyl)thiazol-2-amine



N-(5-(4-Aminophenyl)thiazol-2-yl)acetamide (23 mg, 0.1 mmol) was dissolved in EtOH (2 mL). NaOH (1 M) (2 mL) was added to the reaction mixture which was stirred at 100 °C overnight. Once reacted, the mixture was reduced *in vacuo* and purified via column chromatography using (DCM: MeOH: NH_4OH = 93: 6: 1) to afford title compound (7 mg, 40 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.06 (d, 2H, *J* = 8.5 Hz, Ar), 7.03 (s, 1H, Ar), 6.81 (s, 2H, NH_2), 6.52 (d, 2H, *J* = 8.6 Hz, Ar), 5.13 (s, 2H, NH_2); ^{13}C NMR (101 MHz; $CDCl_3$): δ 166.1, 147.4, 131.5, 126.5, 125.8, 120.0, 114.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 192.2520; Observed $[M+H]^+$ = 192.2843.

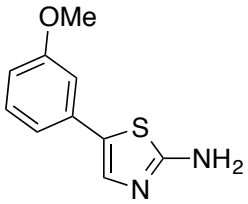
Synthesis of N-(5-(3-methoxyphenyl)thiazol-2-yl)acetamide



N-(5-Bromothiazol-2-yl)acetamide (200 mg, 0.9 mmol) and 3-methoxyphenyl boronic acid (207 mg, 1.3 mmol) were added to a mixture of dioxane (3 mL) and H_2O (1 mL). CsF (345 mg, 2.2 mmol) was also added to reaction mixture. Mixture was degassed using N_2 for a period of 5 min before Pd(dppf)Cl_2.DCM (37 mg, 0.04 mmol) was added and heated at 100 °C for 2 h in a sealed vial. Once reacted, organic layer was collected and pushed through an improvised cotton wool plug to remove any visible Pd by-product from resultant mother liquor. The

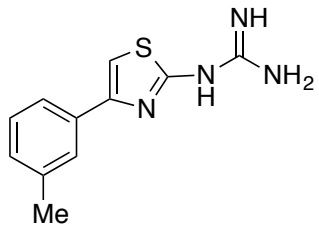
mother liquor was reduced *in vacuo* and purified via column chromatography using (Petroleum: EtOAc = 1: 1) to afford title compound (79 mg, 22 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 10.94 (s, 1H, NH), 7.59 (s, 1H, Ar), 7.31 (t, 1H, *J* = 7.9 Hz, Ar), 7.15 (ddd, 1H, *J* = 7.6 Hz, 1.6 Hz, 0.9 Hz, Ar), 7.07 (t, 1H, *J* = 2.0 Hz, Ar), 6.86 (ddd, 1H, *J* = 8.2 Hz, 2.5 Hz, 0.8 Hz, Ar), 3.79 (s, 3H, Me), 2.34 (s, 3H, Me).

Synthesis of 5-(3-methoxyphenyl)thiazol-2-amine



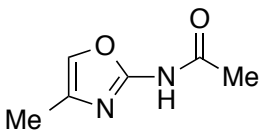
N-(5-(3-Methoxyphenyl)thiazol-2-yl)acetamide (79 mg, 0.3 mmol) was dissolved in EtOH (2 mL). NaOH (1 M) (2 mL) was added to the reaction mixture which was stirred at 100 °C overnight. Once reacted, the mixture was reduced *in vacuo* and purfied via column chromatography using (Petroleum: EtOAc = 2: 1) to afford title compound (47 mg, 72 %). R_f: 0.2. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.3 (s, 1H, Ar), 7.25 (t, 2H, *J* = 16.0 Hz, Ar), 7.01 (ddd, 1H, *J* = 8.2 Hz, 2.5 Hz, 0.7 Hz, Ar), 6.97 (m, 1H, Ar), 4.95 (s, 2H, NH_2), 3.83 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 167.5, 161.2, 137.9, 134.6, 131.8, 122.2, 119.0, 114.4, 110.8, 55.8. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 207.0511; Observed $[M+H]^+$ = 207.0585.

Synthesis of 1-(4-(m-tolyl)thiazol-2-yl)guanidine



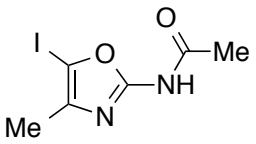
3'-Methylacetophenone (200 mg, 1.5 mmol) was added to MeCN. Br_2 (131 mg, 1.6 mmol) was added and the vial was sealed and allowed to stir for 2 h at 100 °C. After 2 h, 2-imino-4-thiobiuret (352 mg, 3.0 mmol) was added to mixture and the vial sealed once more and allowed to stir for 1 h at 100 °C. The reaction was cooled in ice once reaction was complete. A white precipitate forms and was collected via vacuum filtration. The resultant solid was dissolved in sat. Na_2CO_3 solution (20 mL) and extracted in EtOAc (2 x 10 mL). The organic layer was washed using brine solution and dried over $MgSO_4$ and reduced *in vacuo*. The title compound (203 mg, 58 %) was afforded after work-up. ^{1}H NMR (400 MHz; DMSO-d_6): δ 8.96 (s, 1H, NH), 7.63-7.60 (m, 2H, Ar), 7.26 (t, 1H, *J* = 15.2 Hz, Ar), 7.12 (s, 1H, Ar), 7.08 (d, 1H, *J* = 7.2 Hz, Ar), 6.90 (s, 4H, NH_2), 2.34 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 175.0, 156.9, 149.3, 137.5, 134.7, 128.4, 127.9, 125.9, 122.7, 102.9, 21.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 233.0780; Observed $[M+H]^+$ = 233.0853.

Synthesis of N-(4-methyloxazol-2-yl)acetamide



2-Amino-4-methyl oxazole (1 g, 10.2 mmol, 856 µl) was added to Ac_2O (3.12 g, 30.5 mmol, 2.89 mL). The mixture was heated at 130 °C for 4 h. Once reacted, the mixture was diluted in EtOAc and washed using brine solution (2 x 100 mL). The combined organic layers were collected and dried over $MgSO_4$ and reduced *in vacuo*. The resultant solid was purified via column chromatography using (Petroleum: EtOAc = 1: 1) to afford title compound (277 mg, 20 %). R_f: 0.2. ^{1}H NMR (400 MHz; $CDCl_3$): δ 10.76 (s, 1H, NH), 7.16 (d, 1H, *J* = 1.2 Hz, Ar), 2.28 (m, 3H, Me), 2.12 (d, 3H, *J* = 1.6 Hz, Me).

Synthesis of N-(5-iodo-4-methyloxazol-2-yl)acetamide



N-(4-Methyloxazol-2-yl)acetamide (262 mg, 1.8 mmol) was dissolved in MeCN. Once in solution, NIS (463 mg, 2.1 mmol) was added portionwise to reaction mixture and stirred at rt overnight. Diluted using EtOAc (50 mL), the reaction mixture was washed using 1 M sodium thiosulfate solution (2 x 100 mL). The combined organic layers were dried over $MgSO_4$ and reduced *in vacuo* to afford a golden oil (312 mg). The compound was purified via column chromatography using (Petroleum: EtOAc = 1: 1) to produce title compound (162 mg, 32 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 8.84 (bs, 1H, NH), 2.31 (bs, 3H, Me), 2.12 (s, 3H, Me).

5.3 Benzimidazole derivatives and intermediates

5.3.1 2-Substituted aryl benzimidazole analogues

General procedure A for 2-substituted aryl benzimidazoles via substituted benzoic acids

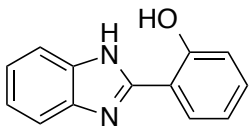
A solution of *o*-phenylenediamine (200 mg, 1.85 mmol), substituted benzoic acid (1.85 mmol) and *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (EDC.HCl) (390 mg, 2.03 mmol) was made up in DMF (3 mL) in a microwave vial. Once the mixture was dissolved, *N*,*N*-Diisopropylethylamine (DIPEA) (387 µL, 2.22 mmol) was added drop-wise and the reaction mixture was allowed to react for 16 h at rt. Once complete, the reaction mixture was diluted through the addition of sat. $NaHCO_3$ solution (a precipitate may form although it may be insignificant) and EtOAc (2 x 20 mL). The organic layers were collected, washed using brine soln., dried over $MgSO_4$ and reduced *in vacuo*. No purification via column chromatography was required once reduced *in vacuo*. The resultant coupled crude was re-dissolved in glacial AcOH (3 mL) and heated for 16 h at 100 °C in a sealed microwave vial. The resultant reaction mixture was allowed to cool to rt before

being added drop-wise to cold NaOH (1 M) solution. Once added, the mixture was diluted using H_2O (20 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine, dried over $MgSO_4$ and reduced *in vacuo*. The resultant crude material was purified via column chromatography.

General procedure B for 2-substituted aryl benzimidazoles using lanthanum chloride

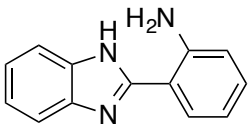
A mixture of *o*-phenylenediamine (200 mg, 1.85 mmol) and substituted benzaldehyde (1.1 mmol) was made up in MeCN (3 mL) in a microwave vial. Lanthanum chloride heptahydrate ($LaCl_3.7H_2O$) (69 mg, 0.18 mmol) was added once all solid was in solution. The reaction mixture was allowed to stir for 2 h at rt. A solid may precipitate which was collected via vacuum filtration. The solid was dissolved in a solution of H_2O and EtOAc to remove any traces of La metal that might be chelating to final product. The organic layer was extracted and then washed using brine, dried over $MgSO_4$, and reduced *in vacuo*. Resultant solid was purified via column chromatography. If no precipitate forms during reaction, the standard extraction procedure using H_2O and EtOAc would apply in which the resultant crude material afforded was purified via column chromatography.

*Synthesis of 2-(1H-benzo[d]imidazol-2-yl)phenol (**42**)*



Applying the general procedure B utilizing 4-hydroxybenzaldehyde (249 mg, 217 µL, 2.03 mmol). The resultant crude orange powder (69 mg) was purified via column chromatography using (Petroleum spirits: EtOAc= 9: 1) to afford title compound (26 mg, 7 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 13.16 (s, 1H, NH), 13.16 (s, 1H, OH), 8.05 (dd, 1H, *J* =7.8 Hz, 1.5 Hz, Ar), 7.68-7.64 (m, 2H, Ar), 7.38 (dd, 1H, *J* = 8.3 Hz, 7.0 Hz, Ar), 7.29 (s, 2H, Ar), 7.05-7.00 (m, 2H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 157.9, 151.6, 145.8, 139.8, 137.0, 131.7, 126.1, 117.1, 115.1, 112.5, 112.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 211.0794; Observed $[M+H]^+$ = 211.0867.

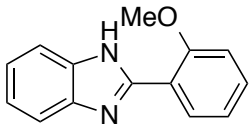
*Synthesis of 2-(1H-benzo[d]imidazol-2-yl)aniline (**43**)*



To a stirred solution of **50** (166 mg, 0.69 mmol) in methanol (15 mL) was added 5% Pd/C (7 mg). If the solid does not dissolve, HCl (1 M) (1 mL) was added. Once dissolved, the reaction mixture was stirred for 18 h under a hydrogen atmosphere. Once complete, reaction was filtered through a short pad of celite to remove

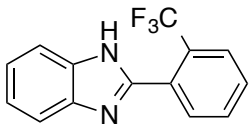
Pd waste. The resultant filtrate was reduced *in vacuo* to afford an amorphous brown salt (125 mg) and was purified via column chromatography using DCM to yield title compound (44 mg, 30 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.64 (s, 1H, NH), 7.82 (dd, 1H, *J* = 7.9 Hz, 1.4 Hz, Ar), 7.56 (d, 2H, *J* = 44.1 Hz, Ar), 7.23 (s, 2H, NH_2), 7.18 (d, 2H, *J* = 3.9 Hz, Ar), 7.14 (ddd, 1H, *J* = 8.3 Hz, 6.9 Hz, 1.4 Hz, Ar), 6.82 (dd, 1H, *J* = 8.2 Hz, 1.1 Hz, Ar), 6.64 (ddd, 1H, *J* = 7.9 Hz, 7.0 Hz, 1.0 Hz, Ar); ^{13}C NMR (101 MHz; $CDCl_3$): δ 150.5, 147.6, 144.9, 131.0, 126.5, 123.3, 120.3, 117.1, 116.8, 114.9, 108.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 210.0952; Observed $[M+H]^+$ = 210.1025.

*Synthesis of 2-(2-methoxyphenyl)-1H-benzo[d]imidazole (**44**)*



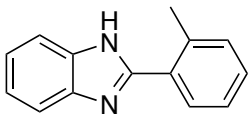
Applying the general procedure B utilizing 2-methoxybenzaldehyde (277 mg, 2.03 mmol). The resultant crude beige material (149 mg) was washed using EtOAc and H_2O to remove any $LaCl_3$ that may remain once collected via vacuum filtration. The organic layer was reduced to reveal title compound (84 mg, 20 %) requiring no further purification. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.11 (s, 1H, NH), 8.32 (dd, 1H, *J* = 7.70 Hz, 1.7 Hz, Ar), 7.65-7.59 (m, 2H, Ar), 7.48 (ddd, 1H, *J* = 8.3 Hz, 7.2 Hz, 1.7 Hz, Ar), 7.25-7.23 (m, 1H, Ar), 7.22-7.17 (m, 2H, Ar), 7.11 (td, 1H, *J* = 7.5 Hz, 0.8 Hz, Ar), 4.02 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 156.7, 148.9, 131.1, 129.7, 121.9, 121.4, 121.9, 121.4, 120.8, 118.1, 112.0, 55.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 225.0951; Observed $[M+H]^+$ = 225.1021.

*Synthesis of 2-(2-(trifluoromethyl)phenyl)-1H-benzo[d]imidazole (**45**)*



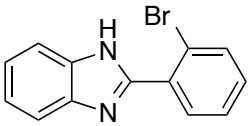
Applying the general procedure A utilizing *o*-phenylenediamine (300 mg, 2.77 mmol) and 2-(trifluoromethyl)benzoic acid (527 mg, 2.77 mmol). The afforded brown solid was purified via column chromatography using (DCM: MeOH = 99: 1) to afford title compound (140 mg, 19 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.75 (s, 1H, NH), 7.94 (d, 1H, *J* = 7.9 Hz, Ar), 7.85-7.83 (m, 1H, Ar), 7.80-7.77 (m, 2H, Ar), 7.73-7.52 (m, 2H, Ar), 7.24 (dd, 2H, *J* = 5.9 Hz, 2.8 Hz, Ar); ^{19}F NMR (400 MHz, DMSO-d_6): δ -62.5; ^{13}C NMR (101 MHz; DMSO-d_6): δ 151.6, 140.2, 134.7, 133.9, 130.2, 129.6 (d, *J* = 3.2 Hz), 128.3, 124.2 (d, *J* = 2.1 Hz), 123.9 (q, *J* = 5.4 Hz), 122.0, 114.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 263.0718; Observed $[M+H]^+$ = 263.0791.

*Synthesis of 2-(o-tolyl)-1H-benzo[d]imidazole (**46**)*



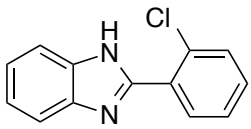
Applying the general procedure A utilizing *o*-toluic acid (252 mg, 1.85 mmol). This compound precipitated out after work-up to afford title compound (109 mg, 28 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.73 (d, 1H, *J* = 7.0 Hz, Ar), 7.68-7.65 (m, 1H, Ar), 7.55-7.52 (m, 1H, Ar), 7.41-7.35 (m, 3H, Ar), 7.20 (d, 2H, *J* = 4.3 Hz, Ar), 2.51 (s, 3H, Me)**;** ^{13}C NMR (101 MHz; DMSO-d_6): δ 151.9, 142.1, 141.1, 136.9, 131.2, 130.0, 129.6, 129.4, 129.2, 125.9, 21.0**.** *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 209.1003; Observed $[M+H]^+$ = 209.1076.

*Synthesis of 2-(2-bromophenyl)-1H-benzo[d]imidazole (**47**)*



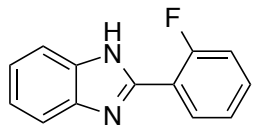
Applying the general procedure A utilizing 2-bromobenzoic acid (477 mg, 1.85 mmol). This compound precipitated out after work up to afford title compound (629 mg, 97 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.75 (s, 1H, Ar), 7.83-7.81 (m, 1H, Ar), 7.77-7.74 (dd, 1H, *J* = 7.6 Hz, 1.5 Hz, Ar), 7.62-7.54 (m, 3H, Ar), 7.49-7.44 (td, 1H, *J* = 7.9 Hz, 7.4 Hz, 1.8 Hz, Ar), 7.26-7.21 (m, 2H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 150.4, 142.3, 138.5, 133.4, 132.4, 132.2, 131.3, 127.8, 122.1, 121.5**.** *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 274.9950; Observed $[M+H]^+$ = 274.9993.

*Synthesis of 2-(2-chlorophenyl)-1H-benzo[d]imidazole (**48**)*



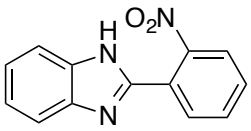
Applying the general procedure B utilizing 2-chlorobenzaldehyde (286 mg, 230 µL, 2.03 mmol). A yellow precipitate was afforded and collected via vacuum filtration (57 mg, 13 %) requiring no further purification. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.71 (s, 1H, NH), 7.91-7.89 (m, 1H, Ar), 7.70 (dd, 1H, *J* = 8.1 Hz, 1.0 Hz, Ar), 7.67-7.64 (m, 1H, Ar), 7.58-7.50 (m, 3H, Ar), 7.28-7.20 (m, 2H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 149.0, 143.6, 134.6, 132.0, 131.6, 129.9, 127.4, 122.7, 121.6, 111.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 229.0454; Observed $[M+H]^+$ = 229.0527.

*Synthesis of 2-(2-fluorophenyl)-1H-benzo[d]imidazole (**49**)*



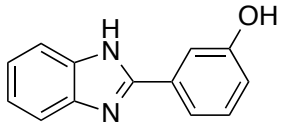
Applying the general procedure B utilizing 2-fluorobenzaldehyde (253 mg, 218 µL, 2.03 mmol). The resultant crude pinkish material (132 mg) was purified via column chromatography using (Petroleum spirits: EtOAc= 4: 1) to afford title compound (44 mg, 11 %). R_f: 0.4. ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.94 (s, 1H, NH), 8.53 (td, 1H, *J* = 7.9 Hz, 1.8 Hz, Ar), 7.73-7.65 (m, 2H, Ar), 7.45 (d, 1H, *J* = 8.2 Hz, Ar), 7.36-7.29 (m, 3H, Ar), 7.27-7.23 (m, 1H, Ar); ^{19}F NMR (400 MHz; DMSO-d_6): δ -109.9; ^{13}C NMR (101 MHz; DMSO-d_6): δ 160.6 (d, *J* = 2.1 Hz), 158.2, 146.3, 131.8 (d, *J* = 8.3 Hz), 130.2, 125.0 (d, *J* = 3.0 Hz), 118.1, 118.0, 116.6, 116.3. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 213.0749; Observed $[M+H]^+$ = 213.0822.

*Synthesis of 2-(2-nitrophenyl)-1H-benzo[d]imidazole (**50**)*



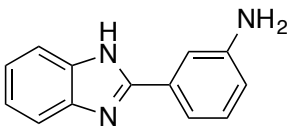
Applying the general procedure B utilizing *o*-phenylenediamine (300 mg, 2.77 mmol) and 2-nitrobenzaldehyde (461 mg, 3.05 mmol). A yellow precipitate was afforded and collected via vacuum filtration requiring no further purification (185 mg, 28 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 13.04 (s, 1H, NH), 8.02 (dd, 1H, *J* = 8.0 Hz, 1.1 Hz, Ar), 7.97 (dd, 1H, *J* = 7.7 Hz, 1.3 Hz, Ar), 7.86 (td, 1H, *J* = 7.6 Hz, 1.2 Hz, Ar), 7.75 (td, 1H, *J* = 7.7 Hz, 1.4 Hz, Ar), 7.65 (d, 1H, *J* = 7.0 Hz, Ar), 7.57 (d, 1H, *J* = 6.3 Hz, Ar), 7.24 (dd, 2H, *J* = 13.0 Hz, 7.4 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 148.9, 147.2, 137.7, 132.6, 130.8, 130.7, 129.0, 127.2, 124.2, 124.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 240.0695; Observed $[M+H]^+$ = 240.0769.

*Synthesis of 3-(1H-benzo[d]imidazol-2-yl)phenol (**51**)*



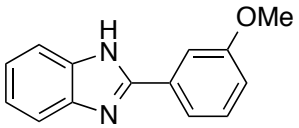
Applying the general procedure A utilizing 3-hydroxybenzoic acid (372 mg, 1.85 mmol). This compound was purified via column chromatography using (petroleum spirits: EtOAc = 2: 1) to afford title compound (88 mg, 23 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 9.83 (s, 1H, OH), 7.64-7.60 (m, 4H, Ar), 7.39-7.35 (m, 1H, Ar), 7.28-7.23 (m, 2H, Ar), 6.94 (ddd, 1H, *J* = 8.2 Hz, 2.3 Hz, 1.0 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 157.8, 150.9, 130.1, 122.6, 120.5, 117.5, 117.4, 113.5, 107.2, 99.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 211.0794; Observed $[M+H]^+$ = 211.0867.

*Synthesis of 3-(1H-benzo[d]imidazol-2-yl)aniline (**52**)*



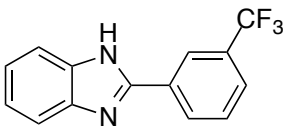
Compound synthesized by Shane Devine applying the general procedure A utilizing 3-aminobenzoic acid. (Unpublished data)

*Synthesis of 2-(3-methoxyphenyl)-1H-benzo[d]imidazole (**53**)*



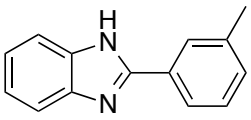
Applying the general procedure A utilizing 3-methoxybenzoic acid (704 mg, 1.85 mmol). This compound was purified via column chromatography using (DCM: MeOH = 98: 2) to afford title compound (399 mg, 38 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.89 (s, 1H, NH), 7.77-7.75 (m, 1H, Ar), 7.64-7.54 (bd, 2H, *J* = 40.4 Hz, Ar), 7.48-7.44 (m, 1H, Ar), 7.21-7.20 (d, 3H, *J* = 4.2 Hz, Ar), 7.07-7.05 (m, 1H, Ar), 3.34 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 163.7, 159.6, 151.0, 131.4, 130.0, 123.3, 118.7, 115.8, 113.2, 111.3, 55.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 225.0951; Observed $[M+H]^+$ = 225.1023.

*Synthesis of 2-(3-(trifluoromethyl)phenyl)-1H-benzo[d]imidazole (**54**)*



Applying the general procedure A utilizing 3-trifluoromethylbenzoic acid (351 mg, 1.85 mmol). This compound was purified via column chromatography using (petroleum spirits: EtOAc = 9: 1) to afford title compound (294 mg, 61 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 13.16 (s, 1H, NH), 8.52 (s, 1H, Ar), 8.49-8.47 (d, 1H, *J* = 5.9 Hz, Ar), 7.87-7.85 (m, 1H, Ar), 7.82-7.78 (t, 1H, *J* = 7.7 Hz, Ar), 7.71-7.69 (d, 1H, *J* = 7.4 Hz, Ar), 7.58-7.56 (d, 1H, *J* = 7.4 Hz, Ar), 7.28-7.20 (m, 2H, Ar); ^{19}F NMR (400 MHz; DMSO-d_6): δ -61.2; ^{13}C NMR (101 MHz; DMSO-d_6): δ 150.1, 131.6, 130.72 (d, *J* = 5.4 Hz), 130.4, 130.1, 126.7 (d, *J* = 3.6 Hz), 123.4, 123.2 (d, *J* = 3.9 Hz), 122.51, 119.60 (d, *J* = 5.3 Hz), 112.07. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 263.0719; Observed $[M+H]^+$ = 263.0791.

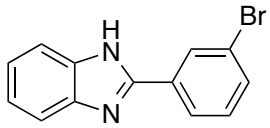
*Synthesis of 2-(m-tolyl)-1H-benzo[d]imidazole (**55**)*



Applying the general procedure A utilizing *m*-toluic acid (252 mg, 1.85 mmol). This compound was purified via column chromatography using (petroleum spirits: EtOAc = 9: 1) to afford title compound (58 mg, 15 %). R_f:

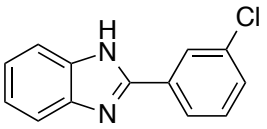
0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.50 (s, 1H, NH), 7.94-7.92 (m, 1H, Ar), 7.84-7.82 (b, 1H, Ar), 7.80-7.78 (bd, 1H, Ar), 7.49-7.47 (b, 1H, Ar), 7.41-7.37 (t, 1H, *J* = 15.2 Hz, Ar), 7.30-7.27 (m, 3H, Ar), 2.44 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 151.2, 138.1, 130.4, 130.0, 128.8, 126.9, 122.4, 121.5, 118.7, 111.2, 21.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 209.1001; Observed $[M+H]^+$ = 209.1074.

*Synthesis of 2-(3-bromophenyl)-1H-benzo[d]imidazole (**56**)*



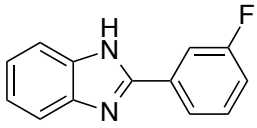
Applying the general procedure A utilizing 3-bromobenzoic acid (372 mg, 1.85 mmol). This compound was purified via column chromatography using (DCM) to afford title compound (39 mg, 8 %). R_f: 0.3. ^{1}H NMR (400 MHz; MeOD): δ 8.30-8.29 (m, 1H, Ar), 8.06 (ddd, 1H, *J* = 7.8 Hz, 1.7 Hz, 1.0 Hz, Ar), 7.68-7.61 (m, 4H, Ar), 7.49-7.45 (m, 1H, Ar), 7.28 (dt, 2H, *J* = 6.3 Hz, 3.3 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 137.4, 136.0, 134.1, 133.1, 131.9, 131.7, 131.5, 127.4, 127.1, 124.0**.** *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 274.9950; Observed $[M+H]^+$ = 274.9985.

*Synthesis of 2-(3-chlorophenyl)-1H-benzo[d]imidazole (**57**)*



Applying the general procedure B utilizing 3-chlorobenzaldehyde (286 mg, 230 μL, 2.03 mmol). A white precipitate was afforded and collected via vacuum filtration (122 mg, 29 %) requiring no further purification. ^{1}H NMR (400 MHz; DMSO-d_6): δ 13.03 (s, 1H, NH), 8.22 (t, 1H, *J* = 1.7 Hz, Ar), 8.14 (ddd, 1H, *J* = 7.1 Hz, 2.1 Hz, 1.0 Hz, Ar), 7.68 (d, 1H, *J* = 7.5 Hz, Ar), 7.61-7.54 (m, 3H, Ar), 7.23 (dq, 2H, *J* = 11.6 Hz, 5.9 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 149.6, 133.7, 132.1, 130.9, 130.3, 130.1, 129.5, 125.9, 124.9, 116.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 229.0453; Observed $[M+H]^+$ = 229.0526.

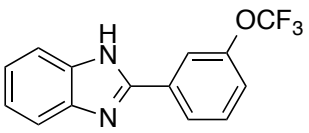
*Synthesis of 2-(3-fluorophenyl)-1H-benzo[d]imidazole (**58**)*



Applying the general procedure A utilizing 3-fluorobenzoic acid (259 mg, 1.85 mmol). The resultant crude material was recrystallized using DCM and MeOH to afford title compound **(**16 mg, 4 %). ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.26 (s, 1H, NH), 7.74 (dd, 2H, *J* = 13.0 Hz, 8.5 Hz, Ar), 7.49 (d, 2H, *J* = 5.5 Hz, Ar), 7.43-7.38 (m, 1H, Ar), 6.99-6.94 (m, 3H, Ar); ^{19}F NMR (400 MHz; $CDCl_3$): δ -111.3; ^{13}C NMR (101 MHz; $CDCl_3$): δ 165.36 (d, *J* = 2.5 Hz) 164.2, 161.7, 135.8 (d, *J* = 6.8 Hz), 130.6, 130.5 (d, *J* = 7.7 Hz),

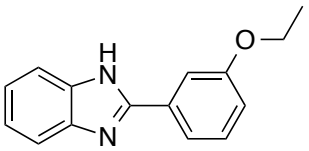
126.6, 126.0, 123.1 (d, J = 2.8 Hz), 115.1 (d, J = 23.4 Hz). *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 213.0749; Observed $[M+H]^+$ = 213.0822.

*Synthesis of 2-(3-(trifluoromethoxy)phenyl)-1H-benzo[d]imidazole (**59**)*



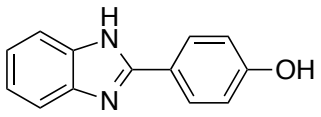
Applying the general procedure A utilizing 3-trifluoromethoxybenzoic acid (704 mg, 1.85 mmol). This compound precipitated out after work-up to afford title compound (254 mg, 49 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 8.24-8.21 (d, 2H, J = 8.0 Hz, Ar), 8.15 (s, 1H, Ar), 7.71-7.67 (t, 1H, J = 8.0 Hz, Ar), 7.63-7.61 (dd, 2H, J = 6.0 Hz, 3.2 Hz, Ar), 7.50-7.47 (m, 1H, Ar), 7.23-7.21 (dd, 2H, J = 6.0 Hz, 3.0 Hz, Ar); ^{19}F NMR (400 MHz; DMSO-d_6): δ -56.68; ^{13}C NMR (101 MHz; DMSO-d_6): δ 150.0, 148.8, 136.6, 133.5, 132.9, 132.8, 131.0, 125.3, 122.1, 121.9, 118.6**.** *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 279.0669; Observed $[M+H]^+$ = 279.0742.

*Synthesis of 2-(3-ethoxyphenyl)-1H-benzo[d]imidazole (**60**)*



Applying the general procedure A (50 mg, 0.46 mmol) utilizing 3-ethoxybenzoic acid (77 mg, 0.46 mmol). This compound precipitated out after work-up to afford title compound (28 mg, 26 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.86 (s, 1H, Ar), 7.76-7.73 (m, 2H, Ar), 7.63 (bs, 1H, Ar), 7.53 (bs, 1H, Ar), 7.46-7.42 (m, 1H, Ar), 7.23-7.18 (m, 2H, Ar), 7.06-7.03 (m, 1H, Ar), 4.13 (q, 2H, J = 6.9 Hz, CH_2), 1.38 (t, 3H, J = 6.9 Hz, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 158.8, 151.0, 131.4, 130.0, 130.0, 118.6, 116.2, 116.2, 113.8, 111.8, 63.2, 14.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 239.1099; Observed $[M+H]^+$ = 239.1173.

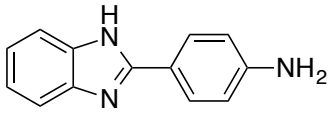
*Synthesis of 4-(1H-benzo[d]imidazol-2-yl)phenol (**61**)*



Applying the general procedure B utilizing 4-hydroxybenzaldehyde (249 mg, 2.03 mmol). The resultant crude beige powder (426 mg) was purified via column chromatography using (Petroleum spirits: EtOAc= 1: 1) to afford title compound (25 mg, 7 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.65 (s, 1H, NH), 9.94 (s, 1H, OH), 8.01-7.98 (m, 2H, Ar), 7.52 (dd, 2H, J = 5.4 Hz, 3.1 Hz, Ar), 7.15 (dd, 2H, J = 5.9 Hz, 3.1 Hz, Ar), 6.92-6.90 (m, 2H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 159.0,

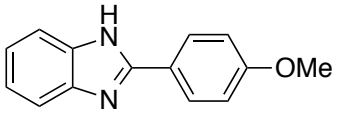
151.7, 140.2, 128.0, 121.8, 121.2, 121.1, 118.3, 115.6, 110.8. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 211.0792; Observed $[M+H]^+$ = 211.0866.

*Synthesis of 4-(1H-benzo[d]imidazol-2-yl)aniline (**62**)*



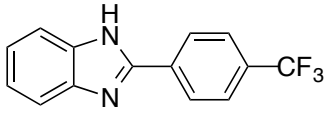
Applying the general procedure B utilizing 2-aminobenzaldehyde (247 mg, 2.03 mmol). The resultant crude material was purified via column chromatography using (Petroleum spirits: EtOAc= 1: 1) to afford title compound (86 mg, 22 %). R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.40 (s, 1H, NH), 7.83 (d, 2H, *J* = 8.5 Hz, Ar), 7.53 (t, 1H, *J* = 4.0 Hz, Ar), 7.41 (t, 1H, *J* = 4.1 Hz, Ar), 7.11-7.09 (m, 2H, Ar), 6.66 (d, 2H, *J* = 8.5 Hz, Ar), 5.58 (s, 2H, NH_2); ^{13}C NMR (101 MHz; DMSO-d_6): δ 152.5, 150.5, 144.0, 134.8, 127.7, 121.3, 121.0, 117.2, 113.5, 110.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 210.0954; Observed $[M+H]^+$ = 210.1027.

*Synthesis of 2-(4-methoxyphenyl)-1H-benzo[d]imidazole (**63**)*



Applying the general procedure B utilizing *p*-anisaldehyde (277 mg, 248 µL, 2.03 mmol). The resultant golden oil (403 mg) after workup was purified via column chromatography using (Petroleum spirits: EtOAc = 2: 1) followed by another column purification using $CHCl_3$ to yield a white solid. The afforded white solid was washed with H_2O (5 mL) to remove any $LaCl_3$ that may have chelated to the product yielding title compound (53 mg, 13 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.72 (s, 1H, NH), 8.13-8.09 (m, 2H, Ar), 7.59 (s, 1H, Ar), 7.50 (s, 1H, Ar), 7.16 (dd, 2H, *J* = 6.1 Hz, 2.9 Hz, Ar), 7.13-7.10 (m, 2H, Ar), 3.84 (s, 3H, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 160.8, 155.5, 141.2, 128.1, 128.1, 114.6, 111.7, 106.0, 55.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 225.0951; Observed $[M+H]^+$ = 225.1023.

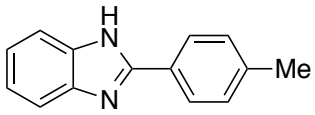
*Synthesis of 2-(4-(trifluoromethyl)phenyl)-1H-benzo[d]imidazole (**64**)*



Applying the general procedure A utilizing 4-trifluoromethylbenzoic acid (351 mg, 1.85 mmol). This compound precipitated out after work up to afford title compound (247 mg, 51 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 13.17 (s, 1H, NH), 8.40-8.37 (d, 2H, *J* = 8.4 Hz, Ar), 7.94-7.92 (d, 2H, *J* = 8.4 Hz, Ar), 7.65-7.62 (dd, 2H, *J* = 5.9 Hz, 3.2 Hz, Ar), 7.25-7.22 (dd, 2H, *J* = 6.0 Hz, 3.1 Hz, Ar); ^{19}F

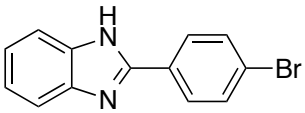
NMR (400 MHz; DMSO-d_6): δ -61.1; ^{13}C NMR (101 MHz; DMSO-d_6): δ 149.7, 141.1, 134.0 (d, *J* = 5.1 Hz), 129.7, 127.0, 125.9 (d, *J* = 3.8 Hz), 125.4 (q, *J* = 222.0 Hz), 122.7, 122.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 263.0716; Observed $[M+H]^+$ = 263.0790.

*Synthesis of 2-(p-tolyl)-1H-benzo[d]imidazole (**65**)*



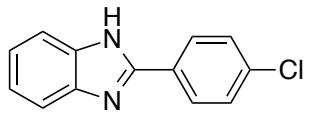
Applying the general procedure A utilizing *o*-phenylenediamine (300 mg, 2.77 mmol) and 4-toluic acid (378 mg, 2.77 mmol). The evolved greenish powder was purified via column chromatography using (DCM: MeOH = 99: 1) to afford title compound (36 mg, 6 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.81 (s, 1H, NH), 8.07 (dd, 2H, *J* = 8.5 Hz, 6.9 Hz, Ar), 7.65-7.62 (m, 1H, Ar), 7.50 (dd, 1H, *J* = 6.7 Hz, 1.5 Hz, Ar), 7.35 (d, 2H, *J* = 7.9 Hz, Ar), 7.22-7.15 (m, 2H, Ar), 2.38 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 151.3, 141.9, 130.7, 128.2, 127.7, 122.1, 116.3, 21.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 209.1001; Observed $[M+H]^+$ = 209.1075.

*Synthesis of 2-(4-bromophenyl)-1H-benzo[d]imidazole (**66**)*



Applying the general procedure A utilizing 4-bromobenzoic acid (372 mg, 1.85 mmol). This compound was purified via column chromatography using (petroleum spirits: EtOAc = 9: 1) to afford title compound (17 mg, 4 %). R_f: 0.2. ^{1}H NMR (400 MHz; MeOD): δ 8.01-7.99 (d, 2H, *J* = 8.8 Hz, Ar), 7.73-7.71 (d, 2H, *J* = 8.8 Hz, Ar), 7.61 (s, 2H, Ar), 7.28-7.26 (m, 2H, Ar); ^{13}C NMR (101 MHz; MeOD): δ 153.8, 133.3, 130.1, 129.4, 125.5, 119.6, 116.1, 101.3**.** *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 274.9950; Observed $[M+H]^+$ = 274.9997.

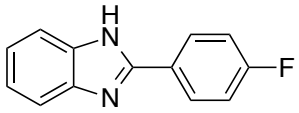
*Synthesis of 2-(4-chlorophenyl)-1H-benzo[d]imidazole (**67**)*



Applying the general procedure B utilizing 4-chlorobenzaldehyde (286 mg, 2.03 mmol). A yellow precipitate was afforded and collected via vacuum filtration (167 mg) and was purified via column chromatography using (Petroleum spirits: EtOAc = 9: 1) to yield title compound (17 mg, 4 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.99 (s, 1H, NH), 8.20-8.17 (m, 2H, Ar), 7.65-7.61 (m, 4H, Ar), 7.21 (dd, 2H, *J* = 6.0 Hz, 3.0 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 150.0, 142.5, 131.1, 130.3,

129.2, 127.8, 122.5, 114.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 229.0453; Observed $[M+H]^+$ = 229.0527.

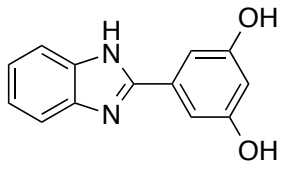
*Synthesis of 2-(4-fluorophenyl)-1H-benzo[d]imidazole (**68**)*



Applying the general procedure B utilizing 4-fluorobenzaldehyde (253 mg, 218 µL, 2.03 mmol). The resultant crude yellow compound (210 mg) was purified via column chromatography using (Petroleum spirits: EtOAc= 4: 1) to afford title compound (83 mg, 21 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.91 (s, 1H, NH), 8.23-8.20 (m, 2H, Ar), 7.58 (s, 2H, Ar), 7.43-7.38 (m, 2H, Ar), 7.20 (dd, 2H, *J* = 6.0 Hz, 3.0 Hz, Ar); ^{19}F NMR (400 MHz; DMSO-d_6): δ -111.1; ^{13}C NMR (101 MHz; DMSO-d_6): δ 164.2, 161.8, 150.3, 141.5, 132.1, 128.6 (d, *J* = 8.6 Hz), 126.8 (d, *J* = 3.0 Hz), 116.1 (d, *J* = 22.2 Hz). *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 213.0749; Observed $[M+H=]^+$ = 213.0822.

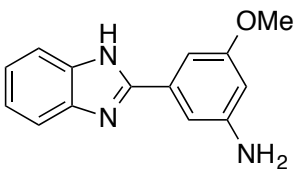
5.3.2 Investigation of site 1 and 2 of benzimidazole scaffold

*Synthesis of 5-(1H-benzo[d]imidazol-2-yl)benzene-1,3-diol **(69)***



Applying the general procedure A utilizing 3,5-dihydroxybenzoic acid (372 mg, 1.85 mmol). This compound was purified via column chromatography using (DCM: MeOH = 98: 2) to afford title compound (88.9 mg, 21 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.68 (s, 1H, NH), 9.50 (s, 2H, OH), 7.55 (s, 2H, Ar), 7.17 (dd, 2H, *J* = 7.2 Hz, 3.1 Hz, Ar), 7.03 (d, 2H, *J* = 7.2 Hz, Ar), 6.35 (s, 1H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 158.7, 151.5, 143.1, 131.7, 128.5, 121.9, 104.8, 104.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 227.0742; Observed $[M+H]^+$ = 227.0813.

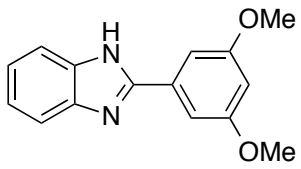
*Synthesis of 3-(1H-benzo[d]imidazol-2-yl)-5-methoxyaniline (**70**)*



Applying the general procedure A (500 mg, 4.62 mmol) utilizing 3-amino-5-methoxybenzoic acid (773 mg, 4.62 mmol). The resultant compound, once worked up, was dissolved in a mixture of EtOH: NaOH (1M) (1:1). The formation of an acetamide-protecting group results owing to the use of AcOH in reaction thus requiring deprotection. The mixture was heated for 16 h at 100 °C. Once complete, the reaction was reduced *in vacuo* and extracted using EtOAc (2 x 25 mL) and

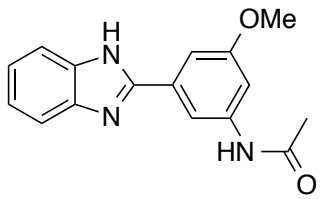
H_2O (30 mL). The organic layers were combined and washed using brine (20 mL), dried over $MgSO_4$ and reduced once again *in vacuo*. The afforded golden amorphous solid (229 mg) was purified via column chromatography using (petroleum spirits: EtOAc = 4: 6 + Et_3N) to afford title compound (67 mg, 23 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.66 (s, 1H, NH), 7.49-7.47 (m, 1H, Ar), 7.18-7.15 (m, 2H, Ar), 7.03 (t, 2H, *J* = 1.6 Hz, Ar), 6.90 (dd, 1H, *J* = 2.2 Hz, 1.4 Hz, Ar), 6.26 (t, 1H, *J* = 2.1 Hz, Ar), 5.29 (s, 2H, NH_2), 3.76 (s, 3H, Me); ^{13}C NMR (101 MHz; MeOD): δ 162.7, 151.0, 148.2, 140.5, 132.4, 124.7, 115.2, 107.3, 103.8, 102.8, 55.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 240.1059; Observed $[M+H]^+$ = 240.1132.

*Synthesis of 2-(3,5-dimethoxyphenyl)-1H-benzo[d]imidazole (**71**)*



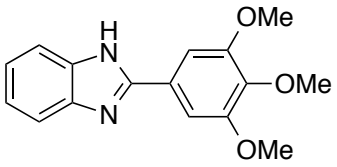
Applying the general procedure A utilizing 3,5-dimethoxybenzoic acid (704 mg, 1.85 mmol). In this case either general methods could be utilized to make this compound. This compound was purified via column chromatography using (DCM: MeOH = 98: 2 to 95: 5) to afford title compound (44.8 mg, 10 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.87 (s, 1H, NH), 7.68-7.66 (m, 1H, Ar), 7.54-7.52 (m, 1H, Ar), 7.36 (d, 2H, *J* = 2.3 Hz, Ar), 7.23-7.19 (m, 2H, Ar), 6.62-6.61 (t, 1H, *J* = 2.2 Hz, Ar), 3.84 (s, 6H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 160.8, 151.0, 143.6, 134.8, 122.6, 111.2, 104.2, 102.0, 55.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 255.1057; Observed $[M+H]^+$ = 255.1130.

*Synthesis of N-(3-(1H-benzo[d]imidazol-2-yl)-5-methoxyphenyl)acetamide **(72)***



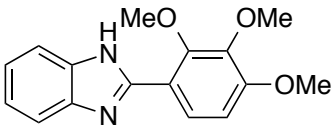
To a stirred suspension of **70** (30 mg, 0.13 mmol), was added acetic anhydride (neat) and reaction mixture was heated for 2 h at 100 °C. Once complete, the reaction was allowed to cool to rt and the resultant precipitate was collected via vacuum filtration and washed using ice cold H_2O to afford a white powder (10 mg). The mother liquor was extracted using HCl (1 M) (10 mL) to remove any unreacted starting material. The remaining compound was extracted using EtOAc (2 x 10 mL) to give a light yellow amorphous solid (21 mg, 62 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.78 (t, 1H, *J* = 1.6 Hz, Ar), 7.61 (dd, 2H, *J* = 5.9 Hz, 3.2 Hz, Ar), 7.41 (dd, 2H, *J* = 1.4 Hz, Ar), 7.27 (dd, 2H, *J* = 6.0 Hz, 3.1 Hz, Ar), 3.9 (s, 3H, Me), 2.1 (s, 3H, Me); ^{13}C NMR (101 MHz; MeOD): δ 168.1, 162.0, 153.7, 141.7, 139.2, 132.3, 124.0, 115.5, 111.8, 108.9, 108.7, 56.0, 23.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 282.1158; Observed $[M+H]^+$ = 282.1231.

*Synthesis of 2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazole (**73**)*



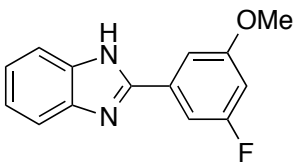
Applying the general procedure B (100 mg, 0.92 mmol scale) utilizing 3,4,5-trimethoxybenzaldehdye (200 mg, 1.01 mmol). A precipitate formed during the reaction, the precipitate was collected via vacuum filtration and subsequently washed with H_2O (5 mL) to remove any residual La metal that may chelate to desired compound. Title compound was afforded (40 mg, 16 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.83 (s, 1H, NH), 7.66-7.64 (d, 1H, *J* = 7.1 Hz, Ar), 7.54-7.52 (m, 3H, Ar), 7.21-7.17 (td, 2H, *J* = 7.7 Hz, 1.5 Hz, Ar), 3.90 (s, 6H, Me), 3.73 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 153.2, 151.2, 143.7, 138.8, 125.4, 121.6, 111.1, 103.7, 60.1, 56.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 285.1173; Observed $[M+H]^+$ = 285.1245.

*Synthesis of 2-(2,3,4-trimethoxyphenyl)-1H-benzo[d]imidazole (**74**)*



Applying the general procedure B (100 mg, 0.92 mmol scale) utilizing 2,3,4-trimethoxybenzaldehdye (200 mg, 1.01 mmol). No solid evolved during reaction time. The reaction mixture was diluted with H_2O and extracted using EtOAc (2 x 20 mL). The organic layer was combined and further washed using brine (20 mL), dried using $MgSO_4$ and reduced *in vacuo*. The resultant brown oil (258 mg) was purified via column chromatography using (petroleum spirits: EtOAc = 3: 1) to afford title compound (49 mg, 19 %). R_f: 0.2. ^{1}H NMR (400 MHz; MeOD): δ 7.03 (dt, 1H, *J* = 8.5 Hz, 0.6 Hz, Ar), 6.73-6.70 (m, 2H, Ar), 6.65-6.62 (m, 1H, Ar), 6.60-6.54 (m, 2H, Ar), 3.91 (s, 3H, CH_3), 3.83 (s, 3H, CH_3), 3.81 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 154.2, 152.6, 149.2, 143.3, 141.0, 123.7, 122.5, 115.8, 110.4, 101.1, 61.8, 61.5, 55.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 285.3150; Observed $[M+H]^+$ = 285.3378.

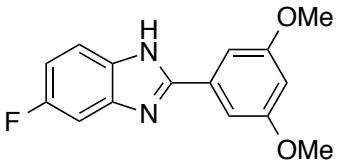
*Synthesis of 2-(3-fluoro-5-methoxyphenyl)-1H-benzo[d]imidazole (**75**)*



Applying the general procedure A (100 mg, 0.93 mmol scale) utilizing 3-fluoro-5-methoxybenzoic acid (157 mg, 0.93 mmol). The resultant solid was purified via column chromatography using (petroleum spirits: EtOAc = 4: 1) to afford title compound (62 mg, 28 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.98 (s, 1H, NH), 7.62-7.55 (m, 4H, Ar), 7.22 (dd, 2H, *J* = 6.0 Hz, 3.0 Hz, Ar). 6.97 (d, 1H, *J* = 10.8 Hz, Ar), 3.88 (s, 3H, CH_3); ^{19}F NMR (400 MHz; DMSO-d_6): δ -110.9;

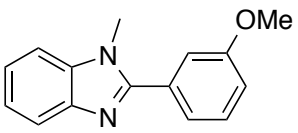
^{13}C NMR (101 MHz; DMSO-d_6): δ 164.4 (d, *J* = 210.1 Hz), 162.0, 161.1, 161.0, 149.9 (d, *J* = 3.8 Hz), 132.7, 132.6, 108.0 (d, *J* = 2.2 Hz), 105.2 (d, *J* = 24.0 Hz), 103.0 (d, *J* = 25.3 Hz), 55.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 243.0863; Observed $[M+H]^+$ = 243.0936.

*Synthesis of 2-(3,5-dimethoxyphenyl)-5-fluoro-1H-benzo[d]imidazole (**76**)*



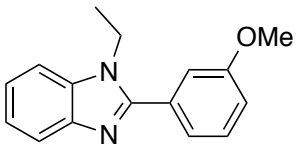
Applying the general procedure B utilizing 4-fluoro-1,2-phenylenediamine (200 mg, 1.58 mmol) and 3,5-dimethoxybenzaldehdye (316 mg, 1.90 mmol). A precipitate was afforded during the reaction; the precipitate was collected via vacuum filtration and washed to remove any residual La metal that may chelate to desired compound. Title compound was afforded (120 mg, 28 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 13.0 (s, 1H, NH), 7.65-7.34 (m, 4H, Ar), 7.09-7.04 (m, 1H, Ar), 6.62 (t, 1H, *J* = 2.2 Hz, Ar), 3.84 (s, 6H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 158.1, 157.3 (d, *J* = 45.4 Hz), 152.4, 141.5 (d, *J* = 17.4 Hz), 138.7, 132.9 (d, *J* = 14.4 Hz), 115.1 (d, *J* = 4.1 Hz), 109.8 (d, *J* = 25.5 Hz), 104.1, 102.9 (d, *J* = 11.4 Hz), 97.6 (d, *J* = 26.9 Hz), 56.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 273.0963; Observed $[M+H]^+$ = 273.1034.

Synthesis of 2-(3-methoxyphenyl)-1-methyl-1H-benzo[d]imidazole (77)



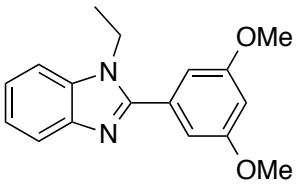
A solution of **53** (100 mg, 0.45 mmol) and potassium carbonate (K_2CO_3) (154 mg, 1.11 mmol) was made up in DMF (3 mL). Once fully dissolved, methyl iodide (MeI) (69 mg, 31 µL, 0.49 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h at rt. After 16 h, the reaction mixture was diluted using H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. Title compound was afforded once dry (90 mg, 85 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.69-7.67 (ddd, 1H, *J* = 7.8 Hz, 1.2 Hz, 0.7 Hz, Ar), 7.63-7.60 (dd, 1H, *J* = 1.2 Hz, 0.6 Hz, Ar), 7.52-7.47 (m, 1H, Ar), 7.42-7.37 (m, 2H, Ar), 7.32-7.22 (m, 2H, Ar), 7.14-7.11 (ddd, 1H, *J* = 8.2 Hz, 2.6 Hz, 1.0 Hz, Ar), 3.88 (s, 3H, Me), 3.84 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 159.2, 152.8, 142.3, 136.5, 131.4, 129.7, 122.3, 121.9, 121.4, 118.9, 115.5, 112.7, 110.5, 55.3, 31.6**.** *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 239.1106; Observed $[M+H]^+$ = 239.1178.

*Synthesis of 1-ethyl-2-(3-methoxyphenyl)-1H-benzo[d]imidazole (**78**)*



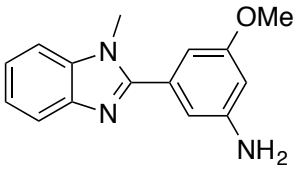
A solution of **53** (100 mg, 0.45 mmol) and potassium carbonate (K_2CO_3) (154 mg, 1.11 mmol) was made up in DMF (3 mL). Once fully dissolved, ethyl iodide (EtI) (77 mg, 39 µL, 0.49 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h at rt. After 16 h, the reaction mixture was diluted using H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. Title compound was afforded once dry (108 mg, 97 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.70-7.64 (m, 2H, Ar), 7.50 (t, 1H, *J* = 7.9 Hz, Ar), 7.33-7.23 (m, 4H, Ar), 7.14 (ddd, 1H, *J* = 8.3 Hz, 2.6 Hz, 0.9 Hz, Ar), 4.31 (q, 2H, *J* = 7.2 Hz, CH_2), 3.84 (s, 3H, CH_3), 1.33 (t, 3H, *J* = 7.1 Hz, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 159.2, 152.4, 142.6, 135.3, 131.7, 129.9, 122.4, 121.9, 121.1, 119.1, 115.5, 114.3, 110.7, 55.2, 39.2, 15.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 253.1259; Observed $[M+H]^+$ = 253.1331.

*Synthesis of 2-(3,5-dimethoxyphenyl)-1-ethyl-1H-benzo[d]imidazole (**79**)*



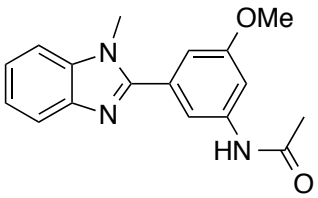
A solution of **71** (50 mg, 0.20 mmol) and potassium carbonate (K_2CO_3) (68 mg, 0.49 mmol) was made up in DMF (3 mL). Once fully dissolved, EtI (34 mg, 17 µL, 0.22 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h at rt. After 16 h, the reaction mixture was diluted using H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant crude was purified via column chromatography using (petroleum spirits: EtOAc = 4: 1) to afford title compound (45 mg, 81 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.85-7.82 (m, 1H, Ar), 7.44-7.42 (m, 1H, Ar), 7.34-7.29 (m, 2H, Ar), 6.85 (d, 2H, *J* = 2.3 Hz, Ar), 6.60 (t, 1H, *J* = 2.3 Hz, Ar), 4.31 (q, 2H, *J* = 7.2 Hz, CH_2), 3.78 (s, 6H, Me), 1.47 (t, 3H, *J* = 7.2 Hz, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 161.0, 151.9, 143.4, 135.4, 132.3, 122.9, 122.5, 120.1, 110.0, 107.4, 102.2, 55.7, 39.8, 29.8, 15.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 283.1376; Observed $[M+H]^+$ = 283.1451.

*Synthesis of 3-methoxy-5-(1-methyl-1H-benzo[d]imidazol-2-yl)aniline (**80**)*



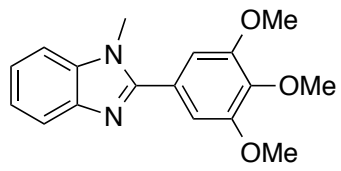
81 (115.1 mg, 0.39 mmol) was dissolved in EtOH: NaOH (5 M) (1:1). The resultant mixture was heated for 3 h at 100 °C. Once complete, the reaction was reduced *in vacuo* and extracted using DCM (2 x 25 mL) and H_2O (30 mL). The combined organic layer was washed using brine (20 mL), dried over $MgSO_4$ and reduced once again *in vacuo*. The afforded solid was purified via column chromatography using (petroleum spirits: EtOAc = 1: 1) to afford title compound (69 mg, 70 %**)**. R_f: 0.2. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.64 (ddd, 1H, *J* = 7.7 Hz, 1.3 Hz, 0.7 Hz, Ar), 7.57 (ddd, 1H, *J* = 7.9 Hz, 1.3 Hz, 0.7 Hz, Ar), 7.29-7.20 (m, 2H, Ar), 6.61 (dd, 1H, *J* = 1.9 Hz, 1.4 Hz, Ar), 6.49 (dd, 1H, *J* = 2.3 Hz, 1.4 Hz, Ar), 6.30 (t, 1H, *J* = 2.1 Hz, NH_2), 3.96 (s, 3H, Me), 3.73 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 160.1, 153.6, 150.1, 142.3, 136.4, 131.3, 122.1, 121.7, 118.8, 110.4, 107.8, 102.4, 100.4, 54.8, 31.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 254.1214; Observed $[M+H]^+$ = 254.1287.

*Synthesis of N-(3-methoxy-5-(1-methyl-1H-benzo[d]imidazol-2-yl)phenyl)acetamide (**81**)*



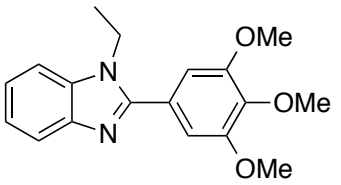
A solution of **72** (20.4 mg, 0.07 mmol) and potassium carbonate (K_2CO_3) (25 mg, 0.18 mmol) was made up in DMF (2 mL). Once fully dissolved, MeI (11 mg, 5 µL, 0.07 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h at rt. After 16 h, the reaction mixture was diluted using H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant crude (43 mg) was purified via column chromatography using (petroleum spirits: EtOAc = 1: 1) to afford title compound (14 mg, 67 %). R_f: 0.4. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.82-7.79 (m, 1H, NH), 7.46 (s, 2H, Ar), 7.42-7.40 (m, 2H, Ar), 7.36-7.31 (m, 2H, Ar), 7.07 (dd, 1H, *J* = 1.6 Hz, 1.3 Hz, Ar), 3.90 (s, 3H, Me), 3.88 (s, 3H, Me), 2.18 (s, 3H, Me); ^{13}C NMR (101 MHz; $CDCl_3$): δ 169.1, 160.3, 153.5, 142.6, 140.1, 136.5, 131.2, 123.1, 122.7, 119.6, 113.4, 110.7, 109.9, 106.9, 55.7, 31.8, 24.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 296.1319; Observed $[M+H]^+$ = 296.1393.

*Synthesis of 1-methyl-2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazole (**82**)*



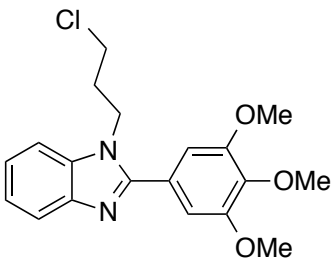
A solution of **73** (50 mg, 0.18 mmol) and potassium carbonate (K_2CO_3) (61 mg, 0.44 mmol) was made up in DMF (3 mL). Once fully dissolved, MeI (27 mg, 12 µL, 0.19 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h at rt. After 16 h, the reaction mixture was diluted using H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant crude was purified via column chromatography using (petroleum spirits: EtOAc = 1: 1) to afford title compound (40 mg, 76 %). R_f: 0.3. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.84-7.81 (m, 1H, Ar), 7.41-7.39 (m, 1H, Ar), 7.33 (dq, 2H, *J* = 6.5 Hz, 3.3 Hz, Ar), 6.96 (s, 2H, Ar), 3.93 (s, 6H, Me), 3.93 (s, 3H, Me), 3.89 (s, 3H, Me); ^{13}C NMR (101 MHz; $CDCl_3$): δ 153.5, 152.8, 152.6, 139.8, 136.7, 123.0, 122.7, 119.9, 109.7, 107.1, 106.9, 61.1, 56.5, 31.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 299.1328; Observed $[M+H]^+$ = 299.1402.

*Synthesis of 1-ethyl-2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazole (**83**)*



A solution of **73** (50 mg, 0.18 mmol) and potassium carbonate (K_2CO_3) (61 mg, 0.44 mmol) was made up in DMF (3 mL). Once dissolved, EtI (30 mg, 16 µL, 0.19 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h at rt. After 16 h, the reaction mixture was diluted using H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The afforded crude was purified via column chromatography using (petroleum spirits: EtOAc = 2: 1) to give title compound (49 mg, 89 %). R_f: 0.4. ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.84-7.81 (m, 1H, Ar), 7.45-7.42 (m, 1H, Ar), 7.31 (dt, 2H, *J* = 6.3 Hz, 2.7 Hz, Ar), 6.92 (s, 2H, Ar), 4.31 (q, 2H, *J* = 7.2 Hz, CH_2), 3.92 (s, 3H, Me), 3.91 (s, 6H, Me), 1.5 (t, 3H, *J* = 7.2 Hz, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 153.5, 153.2, 143.2, 139.7, 135.5, 125.9, 122.9, 122.6, 120.0, 110.0, 106.8, 61.1, 56.4, 39.8, 15.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 313.1486; Observed $[M+H]^+$ = 313.1560.

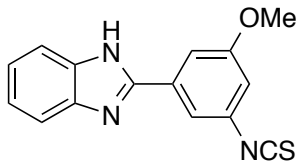
*Synthesis of 1-(3-chloropropyl)-2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazole (**84**)*



A solution of **73** (50 mg, 0.18 mmol) and potassium carbonate (K_2CO_3) (61 mg, 0.44 mmol) was made up in DMF (3 mL). Once fully dissolved, 1-chloro-3-iodopropane (40 mg, 21 µL, 0.19 mmol) was added drop-wise and the reaction mixture was allowed to stir for 48 h at rt. After 48 h, the reaction mixture was diluted in ice-cold H_2O drop-wise. Upon stirring, a precipitate was formed, collected via vacuum filtration and washed using H_2O affording title compound (21 mg, 33 %). ^{1}H NMR (400 MHz; $CDCl_3$): δ 7.67 (td, 2H, *J* = 7.7 Hz, 1.1 Hz, Ar), 7.32-7.24 (m, 2H, Ar), 7.03 (s, 2H, Ar), 4.48-4.44 (m, 2H, Ar), 3.86 (s, 6H, Me), 3.75 (s, 3H, Me), 3.65 (t, 2H, *J* = 6.0 Hz, CH_3), 2.21-2.17 (m, 2H, CH_2); ^{13}C NMR (101 MHz; DMSO-d_6): δ 152.9, 152.8, 142.4, 138.6, 135.6, 125.5, 122.4, 122.0, 119.1, 110.5, 106.6, 60.1, 56.1, 42.6, 41.8, 31.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 361.1248; Observed $[M+H]^+$ = 361.1322.

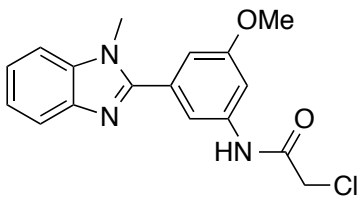
5.3.3 Covalent benzimidazole probes

Synthesis of 2-(3-isothiocyanato-5-methoxyphenyl)-1*H*-benzo[*d*]imidazole (**85**)



Compound **70** (47.3 mg, 0.20 mmol) and *N,N*-thiocarbonyldiimidazole (35 mg, 0.20 mmol) in DCM (3 mL) was allowed to stir at rt for 2 h. Once complete, the mixture was concentrated *in vacuo* to afford a light yellow solid (47 mg) which was purified via column chromatography using (petroleum spirits: EtOAc = 1: 1) to afford title compound (25 mg, 46 %). R_f: 0.4. ^{1}H NMR (400 MHz; $CDCl_3$): δ 9.69-9.50 (s, 1H, NH), 7.75-7.59 (m, 2H, Ar), 7.53 (td, 1H, *J* = 2.9 Hz, 1.4 Hz, Ar), 7.45 (t, 1H, *J* = 1.6 Hz, Ar), 7.34-7.29 (m, 2H, Ar), 6.82-6.81 (m, 1H, Ar), 3.86 (s, 3H, Me); ^{13}C NMR (101 MHz; MeOD): δ 161.2, 155.2, 142.4, 135.8, 133.5, 122.0, 117.1, 116.1, 114.1, 113.6, 112.5, 56.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 282.0619; Observed $[M+H]^+$ = 282.0691.

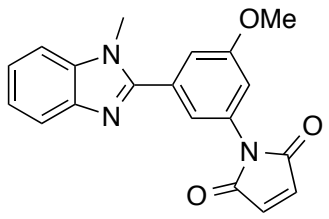
*Synthesis of 2-chloro-N-(3-methoxy-5-(1-methyl-1H-benzo[d]imidazol-2-yl)phenyl)acetamide (**86**)*



A solution of **80** (50 mg, 0.19 mmol) and potassium carbonate (K_2CO_3) (68 mg, 0.49 mmol) was made up in DMF (3 mL). Once fully dissolved, chloroacetyl chloride (24 mg, 17 µL, 0.21 mmol) was added drop-wise and the reaction mixture was allowed to stir for 16 h

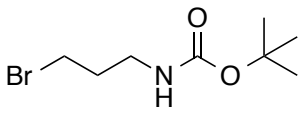
at rt. After 16 h, the reaction mixture was diluted with H_2O (20 mL) and extracted with EtOAc (2 x 10 mL). The organic layers were combined, washed with brine, dried over $MgSO_4$ and reduced *in vacuo*. The resultant clear oil was purified via column chromatography using (petroleum: EtOAc = 1: 1) to afford title compound (39 mg, 60 %). R_f: 0.4. ^{1}H NMR (400 MHz; $CDCl_3$): δ 8.44 (t, 1H, *J* = 0.4 Hz, Ar), 7.83-7.81 (m, 1H, Ar), 7.51 (t, 1H, *J* = 1.6 Hz, Ar), 7.45 (t, 1H, *J* = 2.1 Hz, Ar), 7.43-7.40 (m, 1H, Ar), 7.36-7.30 (m, 2H, Ar), 7.13 (dd, 1H, *J* = 2.3 Hz, 1.4 Hz, Ar), 4.21 (s, 2H, CH_2), 3.92 (s, 3H, CH_3), 3.90 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 164.9, 161.3, 159.5, 149.2, 139.7, 136.5, 131.4, 122.5, 122.0, 118.9, 112.6, 110.6, 109.9, 106.3, 55.4, 42.6, 31.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 330.0926; Observed $[M+H]^+$ = 330.1003.

*Synthesis of 1-(3-methoxy-5-(1-methyl-1H-benzo[d]imidazol-2-yl)phenyl)-1H-pyrrole-2,5-dione (**87**)*



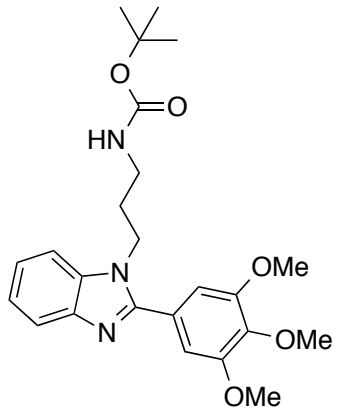
To a solution of maleic anhydride (16 mg, 0.16 mmol) in DCM (3 mL) was added **80** (40 mg, 0.16 mmol). The solution was left to stir at rt for 1 h and monitored via LCMS. Upon completion, the solution was reduced *in vacuo* and the crude was dissolved in a solution of sodium acetate (5.2 mg, 0.06 mmol) in acetic anhydride (3 mL). The reaction mixture was stirred at 100 °C for 1 h. The resultant mixture was poured onto a suspension of ice-cold sat. $NaHCO_3$ (15 mL) to neutralize the reaction mixture. Once neutralized, the aqueous layer was extracted using DCM (2 x 20 mL). The organic layers were combined, washed with brine (20 mL) and dried over $MgSO_4$ before being reduced *in vacuo*. A yellow solid was formed (52 mg) and was purified via column chromatography using (petroleum spirits: EtOAc = 1:1) to afford title compound (39 mg, 75 **%)**. R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.69 (ddd, 1H, *J* = 7.8 Hz, 1.2 Hz, 0.7 Hz, Ar), 7.63 (dt, 1H, *J* = 8.1 Hz, 0.8 Hz, Ar), 7.43 (dq, 2H, *J* = 2.9 Hz, 1.4 Hz, Ar), 7.34-7.24 (m, 2H, Ar), 7.24-7.23 (m, 2H, Ar), 7.13 (dd, 1H, *J* = 2.4 Hz, 1.8 Hz, Ar), 3.91 (s, 3H, Me), 3.87 (s, 3H, Me); ^{13}C NMR (101 MHz; DMSO-d_6): δ 169.7, 159.4, 151.9, 142.3, 136.6, 134.7, 132.8, 131.5, 122.6, 122.1, 119.8, 119.1, 114.0, 113.7, 110.6, 55.7, 31.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 334.1115; Observed $[M+H]^+$ = 334.1191.

*Synthesis of tert-butyl (3-bromopropyl)carbamate (**88**)*



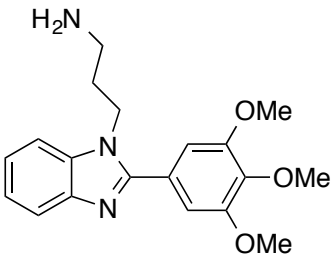
Et_3N was added drop-wise to a stirred mixture of 3-bromopropylamine hydrobromide (2 g, 9.13 mmol) and di-*tert*-butyl dicarbamate (1.8 g, 8.22 mmol) in DCM (20 mL). Once all Et_3N was added, the mixture was stirred at rt for 1 h. Once 1 h has elapsed, the mixture was diluted with DCM (20 mL), washed with $KHSO_4$ (1 M) (2 x 20 mL), brine (20 mL) and dried over $MgSO_4$ and reduced *in vacuo* to give the title compound (1.8 g, 83 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 6.87 (s, 1H, NH), 3.49 (t, 2H, J = 6.6 Hz, CH_2), 3.02 (td, 2H, J = 6.7 Hz, 5.7 Hz, CH_2), 1.90 (quint, 2H, J = 6.7 Hz, CH_2), 1.37 (s, 9H, CH_3).

*Synthesis of tert-butyl (3-(2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazol-1-yl)propyl)carbamate (**89**)*



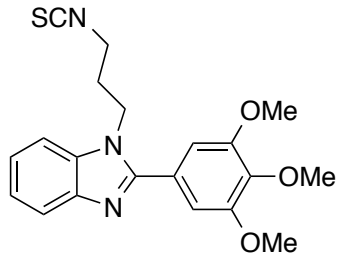
A solution of **73** (175 mg, 0.62 mmol) and potassium carbonate (K_2CO_3) (213 mg, 1.54 mmol) was made up in DMF (3 mL). Once fully dissolved, **88** (177 mg, 0.74 mmol) was added drop-wise and the reaction mixture was sealed and heated to 150 °C for 4-6 h. Once complete, the reaction mixture was added drop-wise to ice-cold H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. A brown solid was afforded once dry (111 mg). The resultant brown solid was purified via column chromatography using a combination of (chloroform: petroleum spirits = 1:1) to afford title compound (56 mg, 21 %). R_f: 0.4 for EtOAc solvent system. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.69-7.64 (m, 2H, Ar), 7.26 (ddd, 2H, J = 9.4 Hz, 7.7 Hz, 1.5 Hz, Ar), 6.99 (s, 2H, Ar), 6.88 (t, 1H, J = 2.6 Hz, NH), 4.32-4.28 (m, 2H, CH_2), 3.86 (s, 6H, CH_3), 3.76 (s, 3H, CH_3), 2.95-2.94 (m, 2H, CH_2), 1.90-1.86 (m, 2H, CH_2), 1.34 (s, 9H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 155.5, 152.9, 152.7, 142.4, 135.6, 125.6, 122.3, 121.9, 119.0, 110.6, 106.6, 77.5, 60.0, 56.0, 42.4, 41.0, 37.4, 30.8, 28.1. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 442.5280; Observed $[M+H]^+$ = 442.5432.

*Synthesis of 3-(2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazol-1-yl)propan-1-amine (**90**)*



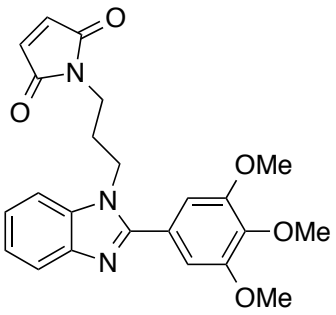
TFA (5 mL) was added to a mixture of **89** (212 mg, 0.48 mmol) in DCM (5 mL). The mixture was stirred at rt for 2 h. Once complete, the reaction was concentrated *in vacuo*. The crude was dissolved in EtOAc (2 x 20 mL) and washed using sat. $NaHCO_3$ (20 mL). The organic layers were combined, washed using brine, dried over $MgSO_4$ and reduced *in vacuo*. A white solid was afforded and purified via column chromatography using (chloroform: EtOAc = 1: 1) to yield title compound (37 mg, 23 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.70-7.67 (m, 2H, Ar), 7.32-7.23 (m, 2H, Ar), 7.03 (s, 2H, Ar), 4.39 (t, 2H, *J* = 7.5 Hz, CH_2), 3.87 (s, 6H, CH_3), 3.75 (s, 3H, CH_3), 2.62 (t, 2H, *J* = 7.1 Hz, CH_2), 1.90-1.86 (m, 2H, CH_2), ^{13}C NMR (101 MHz; DMSO-d_6): δ 152.9, 152.7, 142.4, 138.6, 135.6, 125.6, 122.3, 121.9, 119.0, 110.7, 106.6, 60.1, 56.1, 42.0, 37.7, 20.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 342.4110; Observed $[M+H]^+$ = 342.4223.

*Synthesis of 1-(3-isothiocyanatopropyl)-2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazole (**91**)*



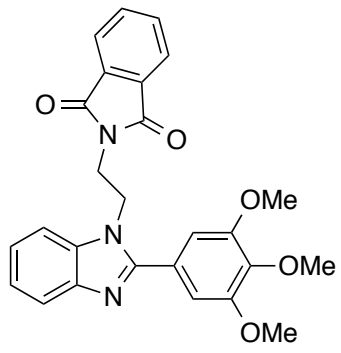
90 (36 mg, 0.11 mmol) and *N,N*-thiocarbonyldiimidazole (19 mg, 0.11 mmol) in DCM (3 mL) was allowed to stir at rt for 2 h. Once complete, the mixture was concentrated *in vacuo* to afford a light yellow solid (51 mg) which was purified via column chromatography using (petroleum spirits: EtOAc = 1: 1) to afford title compound (31 mg, 76 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.85-7.82 (m, 1H, Ar), 7.44-7.41 (m, 1H, Ar), 7.36-7.33 (m, 2H, Ar), 6.9 (s, 2H, Ar), 4.49 (t, 2H, *J* = 6.9 Hz, CH_2), 3.94 (s, 6H, CH_3), 3.65 (s, 3H, Me), 3.40 (dd, 2H, *J* = 6.6 Hz, 5.6 Hz, CH_2), 2.10 (dd, 2H, *J* = 6.4 Hz, 5.1 Hz, CH_2), ^{13}C NMR (101 MHz; MeOD): δ 155.1, 155.0, 150.9, 143.2, 136.3, 126.3, 124.5, 124.1, 119.8, 114.6, 111.7, 108.0, 61.2, 56.9, 43.1, 42.9, 30.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 384.1306; Observed $[M+H]^+$ = 384.1379.

*Synthesis of 1-(3-(2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazol-1-yl)propyl)-1H-pyrrole-2,5-dione (**92**)*



To a solution of maleic anhydride (10 mg, 0.10 mmol) in DCM (4 mL) was added **90** (35 mg, 0.10 mmol). The solution was left to stir at rt for 1 h. Upon completion, the solution was reduced *in vacuo* and dissolved along with sodium acetate (3 mg, 0.04 mmol) in acetic anhydride (2 mL). The reaction mixture was stirred at 100 °C for 1 h. The resultant mixture was poured onto a suspension of ice-cold sat. $NaHCO_3$ (15 mL). Once neutralized, the aqueous layer was extracted using DCM (2 x 20 mL). The organic layers were combined and dried over $MgSO_4$ before being reduced *in vacuo*. A greyish solid formed (26 mg) and was purified via column chromatography using (petroleum spirits: EtOAc = 1:1) to afford title compound (13 mg, 32 %). R_f: 0.3. ^{1}H NMR (400 MHz; DMSO-d_6): δ 7.38-7.36 (m, 2H, CH), 7.34-7.30 (m, 2H, Ar), 6.85 (s, 2H, Ar), 6.63 (s, 2H, Ar), 4.24 (t, 2H, *J* = 8.0 Hz, CH_2), 3.92 (s, 9H, CH_3), 3.53 (t, 2H, *J* = 6.9 Hz, CH_2), 2.14-2.07 (m, 2H, CH_2); ^{13}C NMR (101 MHz; DMSO-d_6): δ 152.9, 135.5, 134.3, 125.6, 122.4, 121.9, 119.0, 110.6, 106.4, 99.4, 60.1, 56.0, 41.8, 39.5, 34.7, 28.3. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 422.1640; Observed $[M+H]^+$ = 422.1715.

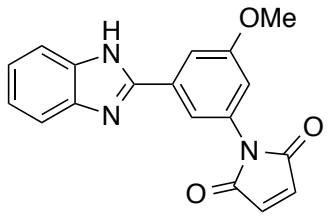
Synthesis of 2-(2-(2-(3,4,5-trimethoxyphenyl)-1H-benzo[d]imidazol-1-yl)ethyl)isoindoline-1,3-dione



A stirred mixture of K_2CO_3 (639 mg, 4.62 mmol) in DMF, *o*-phenylenediamine (200 mg, 1.85 mmol) was added. *N*-(2-bromoethyl)phthalimide (dissolved in DMF) (705 mg, 2.77 mmol) was added drop-wise to the stirred mixture that was allowed to stir at rt for 2 h. Once complete, the reaction mixture was diluted in H_2O (30 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined, washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. A red oil was afforded once dry (732 mg) and was reacted with (3,4,5-trimethoxybenzaldehdye (558 mg, 2.84 mmol) via the application of the general procedure B. The yellow precipitate formed was collected via vacuum filtration and washed to remove any residual La metal that may chelate to desired compound. After the wash step, a dark orange amorphous solid was afforded (941 mg). The afforded solid was purified via column chromatography using (petroleum spirits: EtOAc = 2: 1) yielding title compound (98 mg, 8 %). R_f: 0.3. ^{1}H NMR (400

MHz; $CDCl_3$): δ 7.81-7.79 (m, 1H, Ar), 7.75-7.72 (m, 2H, Ar), 7.72-7.68 (m, 2H, Ar), 7.57-7.55 (m, 1H, Ar), 7.29-7.27 (m, 2H, Ar), 6.87 (s, 2H, Ar), 4.61 (t, 2H, J = 6.5 Hz, CH_2), 4.06 (t, 2H, J = 6.5 Hz, CH_2), 3.89 (s, 3H, CH_3), 3.88 (s, 6H, CH_3); ^{13}C NMR (101 MHz; $CDCl_3$): δ 167.6, 153.5, 143.0, 139.5, 135.7,134.3, 132.1, 131.8, 125.4, 123.4, 123.3, 122.9, 120.2, 109.7, 106.6, 61.1, 56.4, 42.5, 37.0. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 458.4860; Observed $[M+H]^+$ = 458.4943.

Synthesis of 1-(3-(1H-benzo[d]imidazol-2-yl)-5-methoxyphenyl)-1H-pyrrole-2,5-dione



To a solution of maleic anhydride (30 mg, 0.31 mmol) in DCM (3 mL) was added **70** (75 mg, 0.31 mmol). The solution was left to stir at rt for 1 h. Upon completion, the solution was reduced *in vacuo* and re-solubilized in acetic anhydride (3 mL) with sodium acetate (11 mg, 0.13 mmol). The reaction mixture was stirred at 100 °C for 1 h. The resultant mixture was poured onto a suspension of ice-cold sat. $NaHCO_3$ (15 mL). Once neutralized, the aqueous layer was extracted using DCM (2 x 20 mL). The organic layers were combined, washed with brine (20 mL) and dried over $MgSO_4$ before being reduced *in vacuo*. A golden oil was formed (100 mg) and purified via column chromatography using (petroleum spirits: EtOAc = 4:1) to afford title compound (54 mg, 55 %). R_f: 0.2. ^{1}H NMR (400 MHz; $CDCl_3$): δ 8.19-8.17 (m, 1H, Ar), 7.81-7.79 (m, 1H, Ar), 7.43-7.41 (m, 2H, Ar), 7.30 (dd, 1H, J = 2.4 Hz, 1.4 Hz, Ar), 7.23 (t, 1H, J = 1.6 Hz, Ar), 7.14 (dd, 1H, J = 2.4 Hz, 1.8 Hz, Ar), 6.87 (s, 2H, Ar), 3.91 (s, 3H, Me); ^{13}C NMR (101 MHz; MeOD): δ 145.9, 135.6, 134.1, 130.6, 126.8, 126.2, 121.0, 120.2, 115.6, 115.0, 113.0, 101.3, 56.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 320.0951; Observed $[M+H]^+$ = 320.1024.

5.4 Screening procedures

SPR assay

A Biacore T200 biosensor was employed to estimate binding affinities of R1-competing hits. 3D7 $PfAMA1_{[104-442]}$ was immobilized on the CM5 sensor chip using an amine-coupling methodology similar to that described by Harris *et al*.[3] Approximately, 10,000 RU of protein was coupled in a single flow cell (1,000 RU = 1 ng of protein per mm^2). A reference flow cell on the same chip was prepared by subjecting it to the identical amine coupling procedure with no AMA1 protein being injected. All SPR binding experiments were conducted at 25 °C with HBS-EP+ (50 mM HEPES, 150 mM NaCl, 3.4 mM ethylenediaminetetraacetic acid (EDTA), 0.05% Tween 20, 5% DMSO, pH 7.4) as the instrument running buffer at a constant flow rate of 60 μL min^{-1}. Immobilized AMA1

was equilibrated with approximately 30 blank buffer injections of 30 s contact time prior to the start of the binding experiments. Solvent correction curves[4] were obtained from a series of injections of running buffer containing 4.55 % to 5.95 % (v/v) DMSO. R1 peptide was injected over immobilized AMA1 at concentrations of 0, 15.6, 31.3, 62.5, 125, 250, 500, and 1000 nM with 2 min contact time. Fragment samples and elaborated molecules were screened at 50, 100 and 200 μM with 30 s contact times; better binders were screened across a wider concentration range of 12.5-200 μM. All binding data were processed and analyzed using Biacore T200 evaluation software version 1.0. The binding affinities of fragment hits were estimated based on the response (RU) from three different concentrations with a fixed fragment R_{max} (maximal binding capacity of AMA1 protein surface). The R_{max} value for the R1 peptide was determined experimentally by fitting the dose-response curves to a 1:1 steady state affinity model. R_{max} values for each fragment were adjusted according to the following normalization formula:

$$R_{max}[\text{fragment}] = R_{max}[\text{R1peptide}] \times MW[\text{fragment}] / MW[\text{R1 peptide}].$$

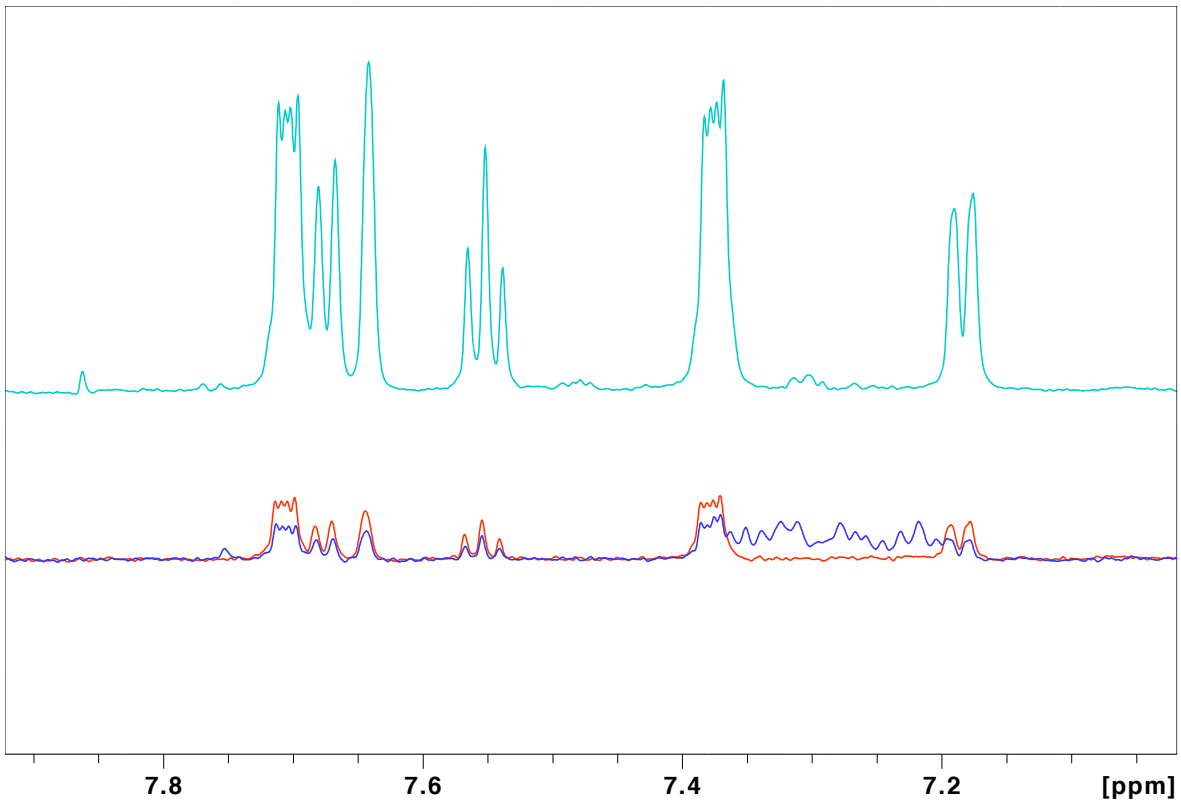
STD NMR Screen

The initial STD screen was conducted on an 1140-fragment library in 190 cocktails, each containing six fragments dissolved in DMSO-d_6. All macromolecular NMR were recorded by San Sui Lim. Samples were prepared by dissolving 3D7 *Pf*AMA1$_{[104-442]}$ and cocktails of six fragments in 20 mM phosphate buffer at pH 7.4 with 10 % D_2O and 1 % DMSO-d_6. The final concentrations of AMA1 and each fragment were 5 and 300 μM, respectively. All STD experiments were performed on a Bruker Avance III 600 MHz spectrometer at 10 °C with 128 scans.[2] Saturation was achieved with a 5 s train of 50 ms Gaussian pulses at 45 dB attenuation. The irradiation frequency of the saturation pulse train was changed after every scan (on- and off-resonance frequencies were -480 and -20000 Hz, respectively). The fragment screen was automated using IconNMR with temperature check (precision ± 0.1 K) and 3 min equilibration time between samples.

STD and CPMG R1 Competition Experiments

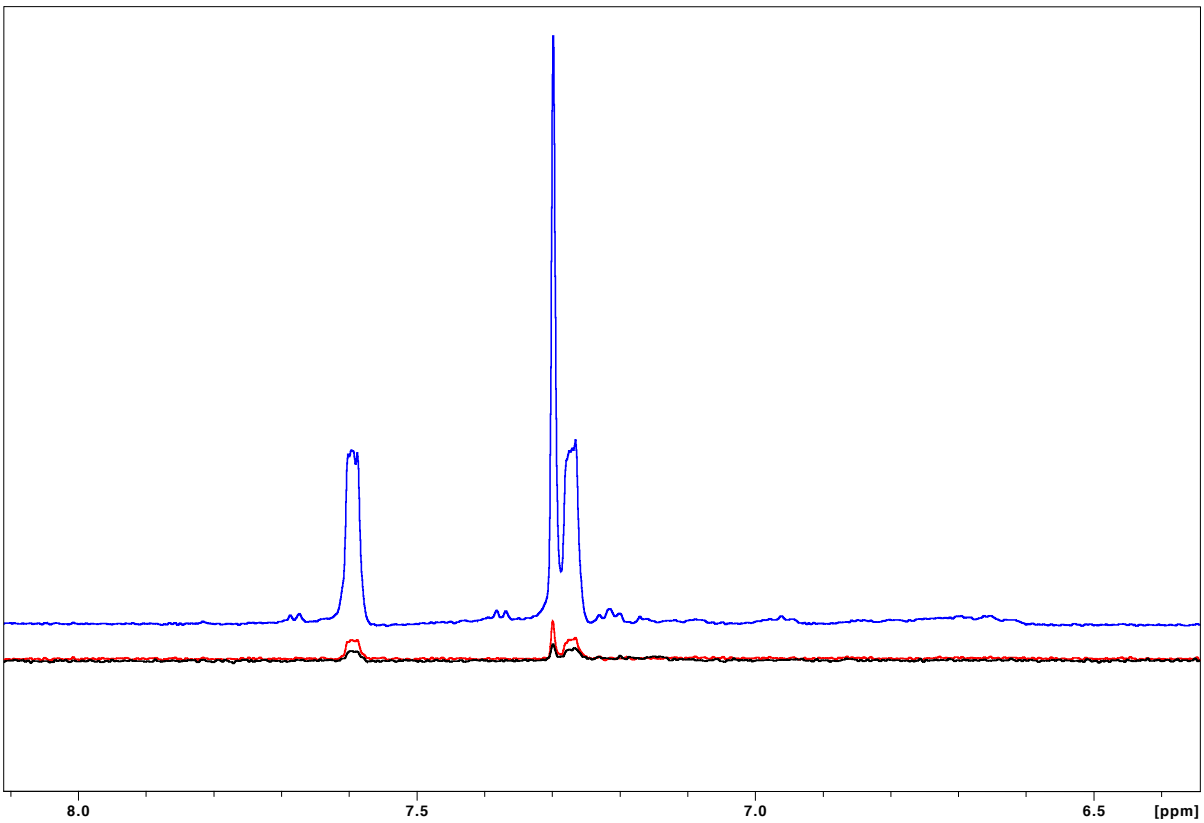
All macromolecular NMR were recorded in collaboration with Indu Chandrashekaran and San Sui Lim. Samples were prepared for STD and CPMG R1 competition experiments by dissolving 3D7 *Pf*AMA1$_{[104-442]}$ and individual fragments in 20 mM phosphate buffer at pH 7.4 containing 10 % D_2O and 1 % DMSO-d_6. The final concentrations of AMA1 and each fragment were 10 and 500 μM, respectively. All NMR experiments were performed on a Bruker Avance III 600 MHz spectrometer at 10 °C. In both STD and CPMG experiments, reference spectra for R1-free samples were first acquired before adding R1 from a 5 mM stock solution to give a final peptide

concentration of 100 μM. In the STD experiments, the parameters used to acquire the R1-containing and R1-free spectra were the same as those from the initial cocktail STD screen. For CPMG experiments, spin-locks of 0 and 0.6 s were applied to both samples with or without R1 peptide. A total of 32 scans were acquired for each sample. Small variations in signal intensities between different experiments were normalized manually using the spectra acquired at 0 s spin-lock. The recorded R1 competition STD NMR spectra of 3D7 *Pf*AMA1$_{[104-442]}$ and individual fragments **53**, **73** and **84** are shown in the following pages.



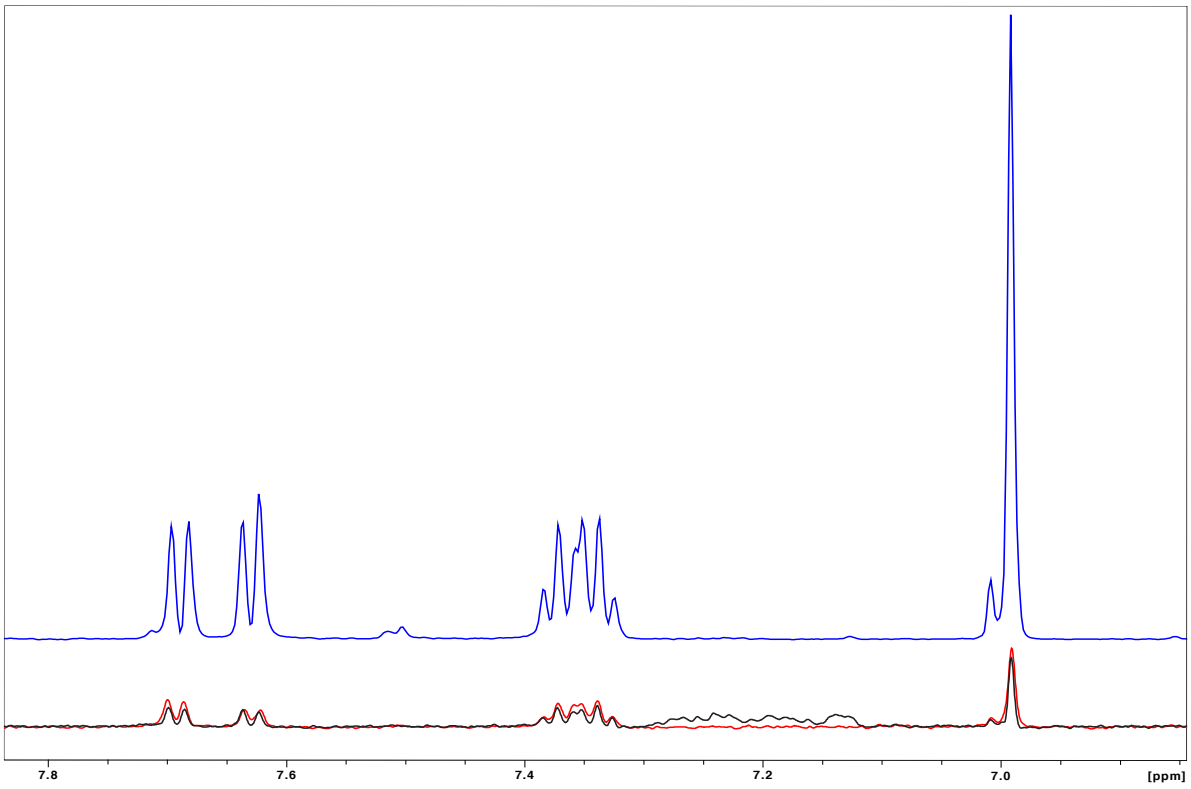
*Compound **53***

STD NMR data for compound **53** displaying the ^{1}H NMR reference spectrum (cyan) along with the R1 absent STD spectrum (red) and R1 present STD spectrum (blue). For clarity, both the R1 absent STD spectrum (red) and R1 present STD spectrum (blue) have been scaled. A reduction in the signal intensity results upon addition of R1 to the sample. The peptide is expected to compete with and displace any fragments bound to the hydrophobic cleft. This was observed as decreased saturation transfer from the protein to fragments and hence a reduction in the STD signal intensity in the R1 STD competition spectrum (blue) is visible.



*Compound **73***

STD NMR data for compound **73** displaying the compound's ^{1}H reference spectrum (blue) along with the R1 absent STD spectrum (red) and R1 present STD spectrum (black). For clarity, both the R1 absent STD spectrum (red) and R1 present STD spectrum (black) have been scaled. A similar reduction in signal intensity has resulted upon addition of R1 to NMR sample.



*Compound **84***

STD NMR data for compound **84** displaying the compound's ^{1}H reference spectrum (blue) along with the R1 absent STD spectrum (red) and R1 present STD spectrum (black). For clarity, both the R1 absent STD spectrum (red) and R1 present STD spectrum (black) have been scaled. A similar reduction in signal intensity has resulted upon addition of R1 to NMR sample.

5.5 Bibliography

(1) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *Journal of Organic Chemistry* **1997**, *62*, 7512–7515.
(2) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie, International Edition* **1999**, *38*, 1784-1788.
(3) Harris, K. S.; Casey, J. L.; Coley, A. M.; Masciantonio, R.; Sabo, J. K.; Keizer, D. W.; Lee, E. F.; McMahon, A.; Norton, R. S.; Anders, R. F.; Foley, M. Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infection and Immunity* **2005**, *73*, 6981–6989.
(4) Papalia, G. A.; Leavitt, S.; Bynum, M. A.; Katsamba, P. S.; Wilton, R.; Qiu, H.; Steukers, M.; Wang, S.; Bindu, L.; Phogat, S.; Giannetti, A. M.; Ryan, T. E.; Pudlak, V. A.; Matusiewicz, K.; Michelson, K. M.; Nowakowski, A.; Pham-Baginski, A.; Brooks, J.; Tieman, B. C.; Bruce, B. D.; Vaughn, M.; Baksh, M.; Cho, Y. H.; Wit, M. D.; Smets, A.; Vandersmissen, J.; Michiels, L.; Myszka, D. G. Comparative analysis of 10 small molecules binding to carbonic anhydrase II by different investigators using Biacore technology. *Analytical Biochemistry* **2006**, *359*, 94–105.

Appendix A - Publications

A.1 Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening

RESEARCH FRONT

CSIRO PUBLISHING

Aust. J. Chem. **2013**, *66*, 1530–1536
http://dx.doi.org/10.1071/CH13266

Full Paper

Development of Inhibitors of *Plasmodium falciparum* Apical Membrane Antigen 1 Based on Fragment Screening

San Sui Lim,[A] *Cael O. Debono,*[A] *Christopher A. MacRaild,*[A] *Indu R. Chandrashekaran,*[A] *Olan Dolezal,*[B] *Robin F. Anders,*[C] *Jamie S. Simpson,*[A] *Martin J. Scanlon,*[A,D] *Shane M. Devine,*[A] *Peter J. Scammells,*[A] *and Raymond S. Norton*[A,E]

[A]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Vic. 3052, Australia.
[B]CSIRO Materials Science and Engineering, Parkville, Vic. 3052, Australia.
[C]Department of Biochemistry, La Trobe University, Bundoora, Vic. 3086, Australia.
[D]Centre of Excellence for Coherent X-Ray Science, Monash University, Parkville, Vic. 3052, Australia.
[E]Corresponding author. Email: ray.norton@monash.edu

Apical membrane antigen 1 (AMA1) is an essential component of the moving junction complex used by *Plasmodium falciparum* to invade human red blood cells. AMA1 has a conserved hydrophobic cleft that is the site of key interactions with the rhoptry neck protein complex. Our goal is to develop small molecule inhibitors of AMA1 with broad strain specificity, which we are pursuing using a fragment-based approach. In our screening campaign, we identified fragments that bind to the hydrophobic cleft with a hit rate of 5 %. The high hit rate observed strongly suggests that a druggable pocket is present within the cleft.

Manuscript received: 22 May 2013.
Manuscript accepted: 29 July 2013.
Published online: 26 August 2013.

Introduction

Malaria is caused by parasites of the genus *Plasmodium* and is one of the most widespread infectious tropical diseases.[1] There are around 250 million clinical cases of malaria each year, resulting in almost one million deaths.[1,2] Amongst the malaria parasites, *P. falciparum* is the most deadly species, being responsible for 90 % of the total malaria-related deaths.[1]

Apical membrane antigen 1 (AMA1) is an essential component of the moving junction (MJ) used by *Plasmodium* merozoites to invade human red blood cells.[3] AMA1 has a conserved hydrophobic cleft that is the site of key interactions with the rhoptry neck (RON) protein complex that forms part of the moving junction (Fig. 1).[4–8] Peptides identified by phage display, such as R1, as well as monoclonal antibodies that target this site on AMA1, are able to inhibit red blood cell invasion, but usually in a strain-specific manner as numerous polymorphic residues are clustered at one end of the cleft (Fig. 2).[5,6,9,10]

AMA1 represents an attractive target for developing antimalarial drugs.[11] It is unique to *Plasmodium* species and other apicomplexan parasites such as *Toxoplasma gondii*. No homologues exist in the human host, facilitating the design of inhibitors that selectively target the malaria parasite. Importantly, AMA1 is expressed on the parasite surface in the human bloodstream, thus avoiding the challenges of delivering a drug across the numerous membranes that protect intra-cellular parasite targets. Moreover, this renders AMA1 inhibitors immune to parasite-mediated drug efflux mechanisms that are important sources of resistance to many existing malaria drugs.[12] Anti-AMA1 antibodies also inhibit sporozoite invasion, providing evidence that the protein is involved in the invasion of hepatocytes.[13] Therefore, a therapeutic agent targeting AMA1 has the potential not only to treat or suppress the development of symptomatic malaria, but also to prevent *Plasmodium* parasites from establishing initial liver-stage infections.

Our aim is to develop small molecule inhibitors of AMA1 using fragment-based ligand design (FBLD). Compared with existing peptides or antibodies targeting AMA1, small molecule inhibitors are preferred candidates for treating malaria because of their lower production costs and likely oral bioavailability.[14] In the fragment-based approach, the process of growing, linking, or merging hits to leads would allow us to install chemical functionalities that are complementary to a conserved binding pocket, important for the discovery of AMA1 inhibitors that have broad strain specificity. More importantly, FBLD has been used successfully in generating promising leads for difficult targets such as those involving protein–protein interactions.[15–19]

Results and Discussion

A diverse fragment library[20] was screened against 3D7 *Pf*AMA1$_{[104–442]}$ corresponding to domains I and II of the AMA1 ectodomain; these fragments were screened in cocktails of six compounds using saturation transfer difference (STD) NMR experiments (Fig. 3a–c).[21,22] Of the 1140 compounds screened, 208 fragments showed positive STD signals in the

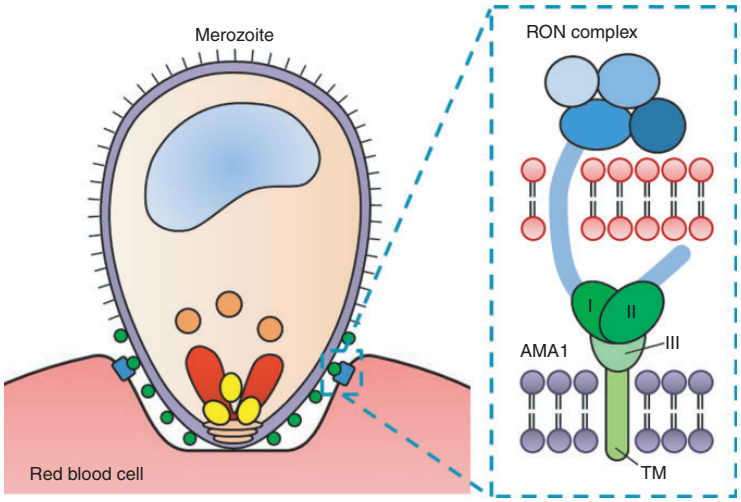


Fig. 1. Graphical representation showing apical membrane antigen 1 (AMA1) (green) interacting with the rhoptry neck (RON) protein complex (blue) to form the moving junction important for merozoite invasion of red blood cells. During invasion, the RON complex is inserted onto the red cell membrane (red) and AMA1 is located on the merozoite membrane surface (purple). AMA1 consists of three extracellular domains (I, II, and III), a transmembrane domain (TM), and a short cytoplasmic tail, with domains I and II directly involved in the binding of RON complex.

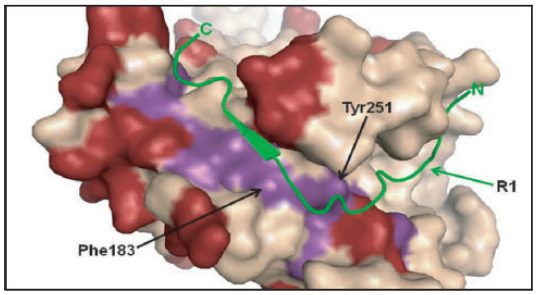


Fig. 2. Co-crystal complex of R1 and 3D7 *Pf*AMA1 DI + II. R1 (green) inhibits apical membrane antigen 1 (AMA1) (light brown) through its binding in the hydrophobic cleft (purple, PDB ID 3SRJ).[8] The R1 inhibitory effect is strain specific as the peptide interacts with several polymorphic residues (red) on AMA1. Residue numbers apply to AMA1. (Phe: phenylalanine, Tyr: tyrosine.)

presence of AMA1, corresponding to a primary hit rate of 18 %. Although STD is a rapid and sensitive method for identifying weak binders, it does not provide any information about the binding sites of these fragment hits.[22] To address this, we made use of R1, an inhibitory peptide that binds to the hydrophobic cleft of AMA1 with a binding affinity (K_D) of 100 nM.[9,23] The hits identified from the initial cocktail screen were followed up as individual compounds using R1 competition experiments to identify hits that bind to the AMA1 hydrophobic cleft.

These competition experiments were conducted using both STD experiments and transverse relaxation rate (T_2) measurements using Carr–Purcell–Meiboom–Gill (CPMG) experiments[24] in both the presence and absence of the competing R1 peptide. Higher concentrations of fragments and 3D7 *Pf*AMA1 (500 and 10 µM, respectively, compared with 300 and 5 µM, respectively, in the initial cocktail screens) were employed in an effort to increase the STD signal intensity and make any competitive effect easier to detect. Upon addition of R1 to these samples, the peptide is expected to compete with and displace any fragments bound to the hydrophobic cleft. This was observed as decreased saturation transfer from the protein to fragments and hence a reduction in STD signal intensities in the R1 STD competition experiments (Fig. 3d). In the case of CPMG, the smaller fraction of bound fragment results in reduced fragment transverse relaxation rates, and thus increased fragment signal after the CPMG period of 600 ms (Fig. 3e).

Among the 208 primary hits evaluated in the R1 competition assays, 65 showed evidence of competition by STD, whilst 98 appeared to compete with the peptide by CPMG. In total, 57 fragments were found to compete with R1 in both assays, corresponding to a 5 % hit rate overall (Fig. 4). We note that 90 compounds that were positive hits in the initial cocktail screen showed no binding activity on AMA1 when examined as individual fragments in the STD experiments. One possible explanation for this observation is that the fragments in the cocktail might have interacted with each other and AMA1 to form higher-order complexes.[25] Despite this intriguing possibility, these compounds were treated as false positives and were not tested further in the R1 competition experiments.

The R1-competing hits were not all structurally distinct, with some fragments having similar chemical structures that could be clustered into different series such as 2-aminothiazoles and 2-aryl furans, among others. Several physicochemical parameters of the hits have been examined to assess whether any properties are important for the compounds to bind to AMA1 (Table 1). From this analysis, there is an obvious difference in log *P* values between R1-competing hits and the library compounds, with the hits appearing to be more hydrophobic on average than the rest of the library compounds. This concurs with the general expectation that non-polar interactions will promote binding in a hydrophobic site. The hits also have slightly higher molecular weights, with the increase in size being attributable to increasing numbers of rings rather than additional rotatable bonds. This trend is consistent with

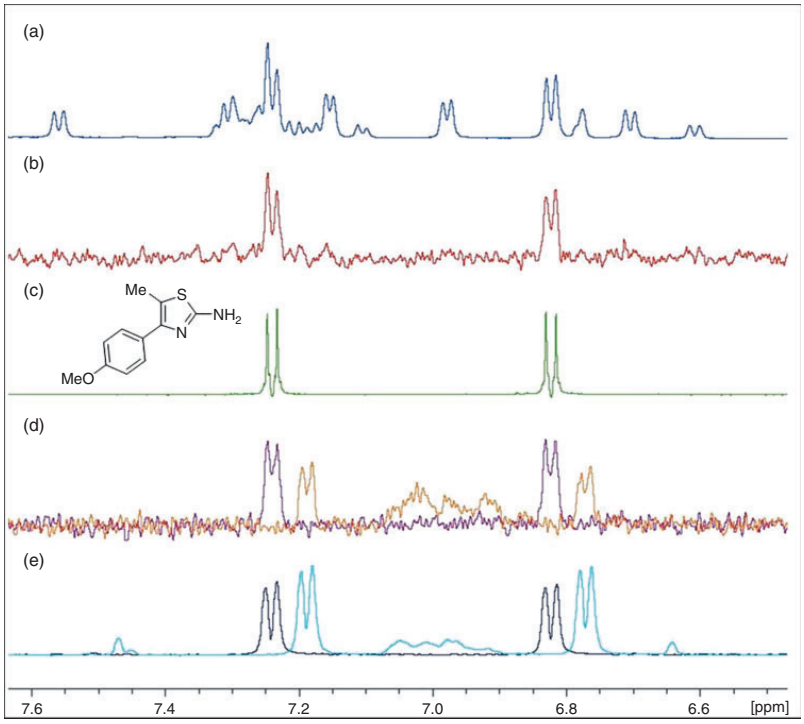


Fig. 3. Saturation transfer difference (STD) results from a cocktail of six fragments at 10°C, 20 mM phosphate buffer, pH 7.4. (a) ^{1}H NMR spectrum for cocktail of six fragments. (b) STD spectrum showing that one fragment out of the cocktail was identified as a hit, with STD signal intensity of 1.3 % (STD signal intensity is the percentage of proton signal reduction relative to the off-resonance spectrum). (c) Reference spectrum for the fragment hit. (d) STD competition NMR experiments. R1 peptide was used to identify fragments that bind to the hydrophobic cleft of 3D7 *Pf*AMA1 DI + II. STD spectra were acquired in the absence (purple) and presence (orange) of R1. (e) CPMG spectra for the fragment hit in the presence of AMA1 and in the absence (grey) and presence (cyan) of R1. In both (d) and (e), spectra in the presence of R1 are offset by −0.07 ppm for clarity. Proton signals between 6.9 and 7.1 ppm are resonances of the R1 peptide.

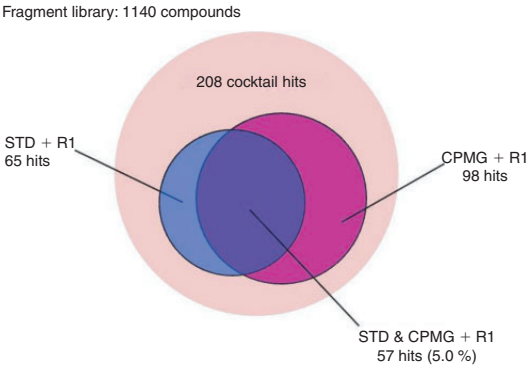


Fig. 4. Venn diagram showing the numbers of hits identified in the cocktail screen and R1 competition experiments. There were 57 hits that showed competition with the peptide in both saturation transfer difference (STD) and Carr–Purcell–Meiboom–Gill (CPMG) experiments.

other fragment screening results reported for protein–protein interactions, which generally have shallow pockets on their interacting surfaces and require more extended scaffolds to make sufficient contacts with the binding sites.[26,27] There is also a slight difference in the numbers of hydrogen bond acceptors and donors (Table 1).

Table 1. Comparison of physicochemical properties of R1-competing hits and library compounds (average values are shown for each property)

Log *P*: partition coefficient; TPSA: topological polar surface area

Properties	Hits	Fragment library
Molecular weight	221.7	210.6
Log *P*	2.16	1.39
TPSA	47.64	47.50
H bond acceptor	2.23	2.47
H bond donor	1.14	0.99
Rotatable bonds	1.98	1.97
Ring count	2.12	1.82

All R1-competing hits were characterised further using surface plasmon resonance (SPR). These compounds were screened against immobilized 3D7 *Pf*AMA1 at 50, 100, and 200 µM to estimate their binding affinities. Fragment concentrations of more than 200 µM were not tested since, at higher concentrations, many of the selected fragments displayed so-called 'SPR-undesirable-behaviours', similar to those

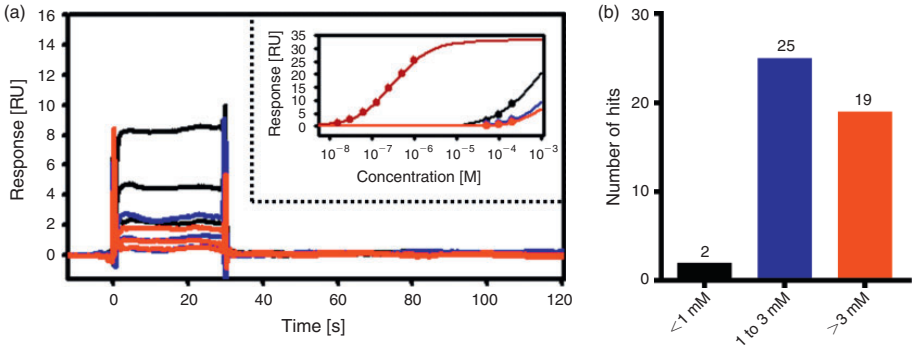


Fig. 5. (a) Surface plasmon resonance (SPR) sensorgram for R1-competing hits tested at three different concentrations (50, 100 and 200 μM). Shown are overlayed sensorgrams for fragments MIPS0000865 (black sensorgrams), MIPS0000713 (blue lines) and MIPS0000873 (red lines); the inset shows binding responses at equilibrium ($t = 20$–25 s, plotted against injected concentration) fitted to a Langmuir adsorption isotherm. R_{max} value (maximal binding response), derived by fitting the R1 peptide responses to the same isotherm (brown dots), was applied to non-saturating responses obtained with fragment hits. Using this approach, the estimated binding affinities (K_D values) for these three fragments were 0.6 mM (MIPS0000865), 2 mM (MIPS0000713) and 3 mM (MIPS0000873) (b) Distribution of the binding affinities (K_D) of R1-competing hits. A total of 46 compounds showed binding activity by SPR.

Table 2. Ligand efficiency (LE) of R1 competing hits

LE [kcal mol^{-1} heavy atom^{-1}]	Number of hits	Percentage [%]
LE <0.2	12	26.1
0.2 ≤ LE <0.3	31	67.4
LE ≥ 0.3	3	6.5

described by Giannetti et al.[28] Thus, for example, several fragments appeared to undergo concentration-dependent aggregation, which, at high (>200 μM) concentrations, resulted in non-stoichiometric binding. Of 57 NMR hits evaluated, 46 compounds showed binding in the SPR experiments. The 11 NMR hits not identified by SPR might have very weak interactions with the target protein that are beyond the detection limit of our SPR experiments. For all fragments tested, binding to AMA1 failed to reach saturation over the concentration range tested, consistent with the relatively weak affinity expected for primary fragment hits. Not unexpectedly, SPR sensorgrams (Fig. 5a) also revealed that the interaction between fragments and AMA1 displayed fast dissociation kinetics that could not be fitted to a kinetic binding model (the data collection rate in Biacore T200 is not sufficient to determine dissociation rate constants, k_d, that are greater than 0.5–1 s^{-1}). In order to rank the selected fragment hits, their binding responses were normalized using a scheme similar to that described by Giannetti et al.[28] Thus, the normalized maximal binding response (R_{max}), predicted from fitting the control compound (R1 peptide) sensorgrams, was applied to response curves obtained with fragments at non-saturating concentrations. To determine K_D values, the Langmuir adsorption isotherm was then fitted to the normalized binding data. Based on SPR experiments, all except two hits showed binding affinities weaker than 1 mM, with the strongest hits having K_D values of 600 μM (Fig. 5b). The weak binding activities observed are a consequence of the size of the fragments and are consistent with other fragment-based drug discovery projects reported in the literature.[15,29] Most of the fragments (67.4 %) bind AMA1 with ligand efficiencies (LE) of between 0.2 and 0.3 kcal mol^{-1} heavy atom^{-1} (Table 2). Three fragments that bind AMA1 have ligand efficiencies

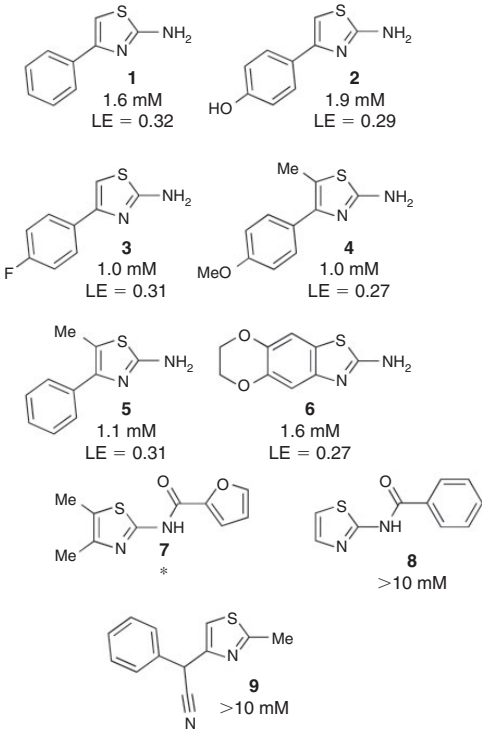


Fig. 6. Thiazole series identified from the R1 competition assay, showing binding affinities (K_D) determined by surface plasmon resonance (SPR) and ligand efficiencies (LE). *The binding affinity of **7** could not be determined as it gave an SPR response of zero at the three test concentrations.

of ≥0.3 kcal mol^{-1} heavy atom^{-1}, all of which belong to the 2-aminothiazole series. Indeed, from the total of 57 hits identified, nine have the common feature of a thiazole core, representing 15 % of the hits. These closely related R1-competing hits are shown in Fig. 6, along with their binding affinities as

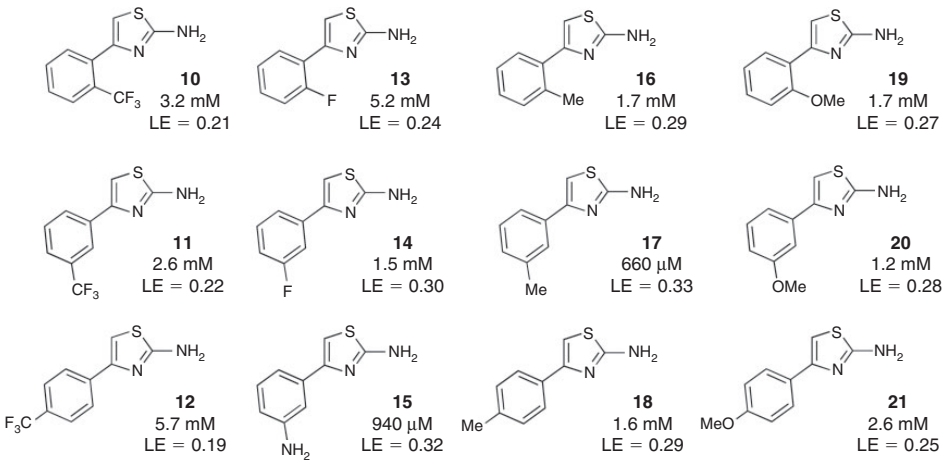


Fig. 7. Synthesised 2-aminothiazole series with binding affinities (K_D) determined by surface plasmon resonance (SPR) and ligand efficiencies (LE) indicated for each analogue.

determined by SPR. From this small set of molecules, some trends were evident. The thiazoles with the 2-amino moiety (**1**–**6**) have binding affinities ranging from 1.0 to 1.9 mM, whereas the 2-amidothiazoles **7** and **8** are virtually non-binders by SPR, with **7** showing no response. The 2-methylthiazole **9** has greatly reduced activity (>10 mM) compared with the 2-aminothiazoles. This 2-aminothiazole binding scaffold should thus serve as a good starting point for chemical elaboration.

Ligand efficiency has gained wide acceptance as a basis for initially ranking and tracking the progress of fragment elaboration.[30] The goal is to maintain the LE throughout the optimization process to give a compound with an LE > 0.3 kcal mol^{-1} heavy atom^{-1}. However, as the LE can decrease during fragment optimization, starting with a high LE hit will make it easier to grow the fragment into an inhibitor with drug-like properties.[30] We have begun to develop preliminary SAR around this scaffold by synthesizing a small library of 4-aryl substituted 2-aminothiazoles (see Supplementary Material), as shown in Fig. 7. Strongly electron-withdrawing substituents on the 4-aryl group, such as the trifluoromethyl group, are not well tolerated, with the *m*-substituent faring slightly better, with a binding affinity of 2.6 mM, over the *o*- and *p*-substituents. The *p*-fluoro (**3**) (1.0 mM) and *m*-fluoro (**14**) (1.5 mM) compounds were substantially better binders than the corresponding *o*-fluoro (**13**) compound (5.2 mM). In the case of the methoxy- and methyl-substituted analogues, *meta*-substituents were found to support the greatest activity, with binding affinities of 660 µM for **17** and 1.2 mM for **20**. The electron-donating affects of the *m*-amino group (**15**) also gave sub-millimolar affinity of 940 µM. Generally speaking, these molecules are weak binders of AMA1 and we are currently synthesizing analogues of this promising scaffold to enable evolution into higher affinity ligands.

Although various inhibitory peptides and antibodies that bind to the hydrophobic cleft of AMA1 have been described, the question of whether a druggable pocket for small molecule inhibitors exists has to date remained unclear.[5,6,8,9,31] Fragments are low-complexity molecules that allow the efficient sampling of large parts of chemical space. This allows the assessment of the protein druggability based on the fragment screening hit rate.[32] The high hit rate observed in our screening strongly suggests that at least one hot spot capable of binding small molecule ligands with high affinity is present within the AMA1 hydrophobic cleft, and our current efforts are directed towards exploiting this by enhancing the binding affinities of promising fragment hits. Efforts are also underway to identify the binding sites of candidate fragments within the hydrophobic cleft by means of high-resolution NMR spectroscopy and X-ray crystallography.

Conclusions

Our fragment screen has identified several *Pf*AMA1-binding scaffolds that bind to the AMA1 hydrophobic cleft. We observed a high hit rate of 5 %, which supports the existence of a druggable pocket within the hydrophobic cleft that is amenable to the design of small molecule inhibitors. The R1-competing hits discovered in the screening process are expected to serve as building blocks for the development of AMA1 inhibitors with broad strain specificity.

Experimental

High-Cell-Density 3D7 Pf*AMA1 DI+II Expression*

3D7 *Pf*AMA1$_{[104–442]}$ (corresponding to domains I + II) was expressed from a pPROEX HTb expression vector (Novagen) in *Escherichia coli* BL21 (DE3) using a high-cell-density methodology.[33] A single colony of the freshly transformed cells was inoculated into L-Broth containing 100 µg mL^{-1} ampicillin. The culture was grown overnight at 37°C with constant shaking at 225 rpm. After ~18 h, cells were harvested by centrifugation and resuspended in two volumes of optimized minimal medium.[33] The mixture was incubated with shaking at 37°C for 1 h before being induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and allowed to grow for 3 h before harvesting. The cell pellets were frozen at −20°C until further use.

Solubilization and Ni^{2+} *Affinity Chromatography*

Frozen cell pellets were thawed at room temperature for 30 min and resuspended with a small volume (2 mL) of cold phosphate

buffered saline (PBS). The mixture was then solubilized with buffer containing 7 M guanidine (Gdn)-HCl and 10 mM Tris (tris(hydroxymethyl)aminomethane) at pH 8.0. The total solubilization volume of mixture was 10-fold that of the weight of the cell paste, giving a final Gdn-HCl concentration of 6 M. The pH was adjusted to 8.0 using 1 M NaOH after 10 min, and incubated for another hour at room temperature with gentle shaking and intermittent 30 s sonication (10 times over 1 h). The mixture was then centrifuged at 20000*g* for 30 min. The supernatant was collected and sterile-filtered, and then 2 % (v/v) Ni^{2+}-charged chelating Sepharose (GE Healthcare) was added. After 2 h incubation, the supernatant was isolated from the resin. A further 2 % fresh Ni^{2+}-charged chelating Sepharose was added and left to stir overnight at room temperature. The mixture was transferred into a column housing and the flow through was collected. The resin was washed with 10 column volumes of guanidine buffers (6 M Gdn-HCl, 100 mM Na_2HPO_4, 10 mM Tris) of decreasing pH, starting at pH 8.0, then pH 6.3 and pH 5.9. Proteins bound to the Ni^{2+}-charged resin were eluted in 5×2 mL fractions of 6 M Gdn-HCl buffer, pH 4.5. Protein content was assessed by UV absorbance at 280 nm using a calculated extinction coefficient of 58000 M^{-1} cm^{-1}.

Refolding

The denatured protein solution was first diluted with 6 M Gdn-HCl buffer, pH 4.5, and then with 1/9 volume of 2 M Tris-HCl buffer, pH 8.0, to give a protein concentration of less than 2 mg mL^{-1}. Under constant stirring at 4°C, this solution was diluted 1 : 50 into the freshly made refold buffer (0.5 M urea, 100 mM NaCl, 20 mM Tris, pH 8.0) that had been filtered and purged with nitrogen gas for a minimum of 1 h. Reduced glutathione (GSH) was added to give a concentration of 2 mM and the mixture was stirred for 3–4 min. Oxidized glutathione (GSSG) was then added at a final concentration of 0.5 mM and the mixture stirred for another 1–2 min. The refold mixture was then incubated overnight at room temperature.

Ion-Exchange Chromatography

The refold mixture was passed through a sterile 0.2 μm filter, diluted 5-fold with 20 mM Tris, pH 8.0, and pumped through a 5 mL HiTRAP QFF (GE Healthcare) column. Refolded 3D7 *Pf*AMA1 was eluted using a linear gradient of 0 to 300 mM NaCl in 20 mM Tris, pH 8.0, over 40 column volumes. UV absorbance at 280 nm was measured to identify the desired product. Fractions containing 3D7 *Pf*AMA1 were combined and analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). Protein concentration was determined using UV absorbance at 280 nm. The product was then dialyzed against 4×100 volumes of 20 mM ammonium bicarbonate solution at 4°C over 2 days before it was lyophilized.

STD NMR Screen

The initial STD screen was conducted on an 1140-fragment library in 190 cocktails, each containing six fragments dissolved in d_6-DMSO. Samples were prepared by dissolving 3D7 *Pf*AMA1$_{[104–442]}$ and cocktails of six fragments in 20 mM phosphate buffer at pH 7.4 with 10 % D_2O and 1 % d_6-DMSO. The final concentrations of AMA1 and each fragment were 5 and 300 μM, respectively. All STD experiments were performed on a Bruker Avance III 600 MHz spectrometer at 10°C with 128 scans.[21] Saturation was achieved with a 5 s train of 50 ms Gaussian pulses at 45 dB attenuation. The irradiation frequency of the saturation pulse train was changed after every scan (on- and off-resonance frequencies were −480 and −20000 Hz, respectively). The fragment screen was automated using IconNMR with temperature check (precision ± 0.1 K) and 3 min equilibration time between samples.

STD and CPMG R1 Competition Experiments

Samples were prepared for STD and CPMG R1 competition experiments by dissolving 3D7 *Pf*AMA1$_{[104–442]}$ and individual fragments in 20 mM phosphate buffer at pH 7.4 containing 10 % D_2O and 1 % d_6-DMSO. The final concentrations of AMA1 and each fragment were 10 and 500 μM, respectively. All NMR experiments were performed on a Bruker Avance III 600 MHz spectrometer at 10°C. In both STD and CPMG experiments, reference spectra for R1-free samples were first acquired before adding R1 from a 5 mM stock solution to give a final peptide concentration of 100 μM. In the STD experiments, the parameters used to acquire the R1-containing and R1-free spectra were the same as those from the initial cocktail STD screen. For CPMG experiments, spin-locks of 0 and 0.6 s were applied to both samples with or without R1 peptide. A total of 32 scans were acquired for each sample. Small variations in signal intensities between different experiments were normalized manually using the spectra acquired at 0 s spin-lock.

SPR Screen

A Biacore T200 biosensor was employed to estimate binding affinities of R1-competing hits. 3D7 *Pf*AMA1$_{104–442]}$ was immobilized on the CM5 sensor chip using an amine-coupling methodology similar to that described by Harris et al.[31] Approximately 10000 RU of protein was coupled in a single flow cell (1000 RU = 1 ng of protein per mm^2). A reference flow cell on the same chip was prepared by subjecting it to the identical amine coupling procedure with no AMA1 protein being injected. All SPR binding experiments were conducted at 25°C with HBS-EP+ (50 mM 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid (HEPES), 150 mM NaCl, 3.4 mM ethylenediaminetetraacetic acid (EDTA), 0.05 % Tween 20, 5 % DMSO, pH 7.4) as the instrument running buffer at a constant flow rate of 60 μL min^{-1}. Immobilized AMA1 was equilibrated with 30 blank buffer injections of 30 s contact time before the start of the binding experiments. Solvent correction curves[34] were obtained from a series of injections of running buffer containing 4.55 to 5.95 % (v/v) DMSO. R1 peptide was injected over immobilized AMA1 at concentrations of 0, 15.6, 31.3, 62.5, 125, 250, 500, and 1000 nM with 2 min contact time. Fragment samples and elaborated molecules were screened at 50, 100, and 200 μM with 30 s contact times. All binding data were processed and analyzed using Biacore T200 evaluation software version 1.0. The binding affinities of fragment hits were estimated based on the response (RU) from three different concentrations with a fixed fragment R_{max} (maximal binding capacity of AMA1 protein surface). The R_{max} value for the R1 peptide was determined experimentally by fitting the dose–response curves to a 1 : 1 steady-state affinity model. R_{max} values for each fragment were adjusted according to the following normalization formula:

$$R_{\max[\text{fragment}]} = R_{\max[\text{R1 peptide}]} \times \text{MW}_{[\text{fragment}]}/\text{MW}_{[\text{R1 peptide}]}$$
$$(\text{MW} = \text{molecular weight}).$$

Supplementary Material

The experimental techniques required to synthesize compounds **10**–**21**, the ^{1}H NMR, ^{13}C NMR (and where appropriate, ^{19}F NMR) spectra, and HRMS spectra associated with these compounds are available on the Journal's website.

Acknowledgements

The authors thank Martin L. Williams and Mark D. Mulcair for their assistance in this project. This work was supported in part by an Australian National Health and Medical Research Council project grant (1025150). R.S.N. acknowledges fellowship support from the NHMRC.

References

[1] World Health Organization (WHO), *World Malaria Report* **2012**.
[2] N. J. White, A. M. Dondorp, A. Faiz, S. Mishra, T. T. Hien, *Lancet* **2012**, *380*, 559. doi:10.1016/S0140-6736(12)61321-X
[3] T. Triglia, J. Healer, S. R. Caruana, A. N. Hodder, R. F. Anders, B. S. Crabb, A. F. Cowman, *Mol. Microbiol.* **2000**, *38*, 706. doi:10.1046/J.1365-2958.2000.02175.X
[4] T. Bai, M. Becker, A. Gupta, P. Strike, V. J. Murphy, R. F. Anders, A. H. Batchelor, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12736. doi:10.1073/PNAS.0501808102
[5] A. M. Coley, A. Gupta, V. J. Murphy, T. Bai, H. Kim, R. F. Anders, M. Foley, A. H. Batchelor, *PLoS Pathog.* **2007**, *3*, e138. doi:10.1371/JOURNAL.PPAT.0030138
[6] K. A. Henderson, V. A. Streltsov, A. M. Coley, O. Dolezal, P. J. Hudson, A. H. Batchelor, A. Gupta, T. Bai, V. J. Murphy, R. F. Anders, M. Foley, S. D. Nuttall, *Structure* **2007**, *15*, 1452. doi:10.1016/J.STR.2007.09.011
[7] M. L. Tonkin, M. Roques, M. H. Lamarque, M. Pugnière, D. Douguet, J. Crawford, M. Lebrun, M. J. Boulanger, *Science* **2011**, *333*, 463. doi:10.1126/SCIENCE.1204988
[8] B. Vulliez-Le Normand, M. L. Tonkin, M. H. Lamarque, S. Langer, S. Hoos, M. Roques, F. A. Saul, B. W. Faber, G. A. Bentley, M. J. Boulanger, M. Lebrun, *PLoS Pathog.* **2012**, *8*, e1002755. doi:10.1371/JOURNAL.PPAT.1002755
[9] K. S. Harris, J. L. Casey, A. M. Coley, R. Masciantonio, J. K. Sabo, D. W. Keizer, E. F. Lee, A. McMahon, R. S. Norton, R. F. Anders, M. Foley, *Infect. Immun.* **2005**, *73*, 6981. doi:10.1128/IAI.73.10.6981-6989.2005
[10] E. F. Lee, S. Yao, J. K. Sabo, W. D. Fairlie, R. A. Stevenson, K. S. Harris, R. F. Anders, M. Foley, R. S. Norton, *Biopolymers* **2011**, *95*, 354. doi:10.1002/BIP.21582
[11] C. A. MacRaild, R. F. Anders, M. Foley, R. S. Norton, *Curr. Top. Med. Chem.* **2011**, *11*, 2039. doi:10.2174/156802611796575885
[12] C. P. Sanchez, A. Dave, W. D. Stein, M. Lanzer, *Int. J. Parasitol.* **2010**, *40*, 1109. doi:10.1016/J.IJPARA.2010.04.001
[13] O. Silvie, J. F. Franetich, S. Charrin, M. S. Mueller, A. Siau, M. Bodescot, E. Rubinstein, L. Hannoun, Y. Charoenvit, C. H. Kocken, A. W. Thomas, G. J. van Gemert, R. W. Sauerwein, M. J. Blackman, R. F. Anders, G. Pluschke, D. Mazier, *J. Biol. Chem.* **2004**, *279*, 9490. doi:10.1074/JBC.M311331200
[14] S. Kar, S. Kar, *Nat. Rev. Drug Discov.* **2010**, *9*, 511. doi:10.1038/NRD3207
[15] M. Congreve, G. Chessari, D. Tisi, A. J. Woodhead, *J. Med. Chem.* **2008**, *51*, 3661. doi:10.1021/JM8000373
[16] D. A. Erlanson, R. S. McDowell, T. O'Brien, *J. Med. Chem.* **2004**, *47*, 3463. doi:10.1021/JM040031V
[17] D. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Nat. Rev. Drug Discov.* **2004**, *3*, 660. doi:10.1038/NRD1467
[18] D. E. Scott, A. G. Coyne, S. A. Hudson, C. Abell, *Biochemistry* **2012**, *51*, 4990. doi:10.1021/BI3005126
[19] J. A. Wells, C. L. McClendon, *Nature* **2007**, *450*, 1001. doi:10.1038/NATURE06526
[20] B. C. Doak, C. J. Morton, J. S. Simpson, M. J. Scanlon, *Aust. J. Chem.* **2013**, *66*, 1465. doi:10.1071/CH13280
[21] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784. doi:10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q
[22] B. Meyer, T. Peters, *Angew. Chem. Int. Ed.* **2003**, *42*, 864. doi:10.1002/ANIE.200390233
[23] D. Richard, C. A. MacRaild, D. T. Riglar, J. Chan, M. Foley, J. Baum, S. A. Ralph, R. S. Norton, A. F. Cowman, *J. Biol. Chem.* **2010**, *285*, 14815. doi:10.1074/JBC.M109.080770
[24] P. J. Hajduk, E. T. Olejniczak, S. W. Fesik, *J. Am. Chem. Soc.* **1997**, *119*, 12257. doi:10.1021/JA9715962
[25] B. J. Davis, D. A. Erlanson, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2844. doi:10.1016/J.BMCL.2013.03.028
[26] I. J. Chen, R. Hubbard, *J. Comput. Aided Mol. Des.* **2009**, *23*, 603. doi:10.1007/S10822-009-9280-5
[27] S. Surade, T. L. Blundell, *Chem. Biol.* **2012**, *19*, 42. doi:10.1016/J.CHEMBIOL.2011.12.013
[28] A. M. Giannetti, B. D. Koch, M. F. Browner, *J. Med. Chem.* **2008**, *51*, 574. doi:10.1021/JM700952V
[29] P. Hajduk, J. Greer, *Nat. Rev. Drug Discov.* **2007**, *6*, 211. doi:10.1038/NRD2220
[30] S. Schultes, C. de Graaf, E. E. J. Haaksma, I. J. P. de Esch, R. Leurs, O. Krämer, *Drug Discov. Today. Technol.* **2010**, *7*, e157. doi:10.1016/J.DDTEC.2010.11.003
[31] K. S. Harris, J. L. Casey, A. M. Coley, J. A. Karas, J. K. Sabo, Y. Y. Tan, O. Dolezal, R. S. Norton, A. B. Hughes, D. Scanlon, M. Foley, *J. Biol. Chem.* **2009**, *284*, 9361. doi:10.1074/JBC.M808762200
[32] P. J. Hajduk, J. R. Huth, S. W. Fesik, *J. Med. Chem.* **2005**, *48*, 2518. doi:10.1021/JM049131R
[33] A. Sivashanmugam, V. Murray, C. Cui, Y. Zhang, J. Wang, Q. Li, *Protein Sci.* **2009**, *18*, 936. doi:10.1002/PRO.102
[34] G. A. Papalia, S. Leavitt, M. A. Bynum, P. S. Katsamba, R. Wilton, H. Qiu, M. Steukers, S. Wang, L. Bindu, S. Phogat, A. M. Giannetti, T. E. Ryan, V. A. Pudlak, K. Matusiewicz, K. M. Michelson, A. Nowakowski, A. Pham-Baginski, J. Brooks, B. C. Tieman, B. D. Bruce, M. Vaughn, M. Baksh, Y. H. Cho, M. D. Wit, A. Smets, J. Vandersmissen, L. Michiels, D. G. Myszka, *Anal. Biochem.* **2006**, *359*, 94. doi:10.1016/J.AB.2006.08.021

A.2 Ligand-Induced Conformational Change of *Plasmodium falciparum* AMA1 Detected using ^{19}F NMR

Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

Ligand-Induced Conformational Change of *Plasmodium falciparum* AMA1 Detected Using ^{19}F NMR

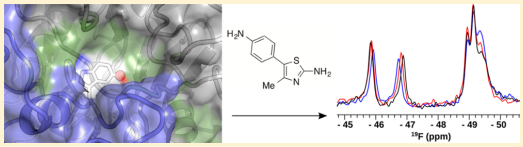
Xiaopeng Ge,[†,§] Christopher A. MacRaild,[‡,§] Shane M. Devine,[‡] Cael O. Debono,[‡] Geqing Wang,[‡] Peter J. Scammells,[‡] Martin J. Scanlon,[‡] Robin F. Anders,[†] Michael Foley,*[,†] and Raymond S. Norton*[,‡]

[†]Department of Biochemistry, La Trobe University, Melbourne 3086, Victoria, Australia

[‡]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville 3052, Victoria, Australia

Supporting Information

ABSTRACT: We established an efficient means of probing ligand-induced conformational change in the malaria drug target AMA1 using ^{19}F NMR. AMA1 was labeled with 5-fluorotryptophan (5F-Trp), and the resulting 5F-Trp resonances were assigned by mutagenesis of the native Trp residues. By introducing additional Trp residues at strategic sites within a ligand-responsive loop, we detected distinct conformational consequences when various peptide and small-molecule ligands bound AMA1. Our results demonstrate an increase in flexibility in this loop caused by the native ligand, as inferred from, but not directly observed in, crystal structures. In addition, we found evidence for long-range allosteric changes in AMA1 that are not observed crystallographically. This method will be valuable in ongoing efforts to identify and characterize therapeutically relevant inhibitors of protein−protein interactions involving AMA1 and is generalizable to the study of ligand-induced conformational change in a wide range of other drug targets.



INTRODUCTION

There are approximately 200 million cases of malaria each year, resulting in over 600000 deaths, mostly caused by *Plasmodium falciparum* infections.[1] The RTS,S malaria vaccine is showing some efficacy in an ongoing phase III trial,[2−4] but antimalarial drugs will continue to play a critical role in controlling and treating the disease for the foreseeable future. However, the repeated emergence of drug-resistant parasites has made it essential that the development of new antimalarial drugs continues.[5,6] A range of new potential drug targets has been identified as a result of sequencing the parasite's genome and from a more detailed understanding of the parasites' life cycle at the molecular level.[7−9] To cause disease, parasite merozoites must invade host erythrocytes, a process that involves initial attachment, reorientation, tight-junction formation, and engagement of the parasite actin−myosin motor to drive the parasite into the erythrocyte.[10,11] The mechanisms underlying these processes are unique to *Plasmodium* and related parasites, making them potential targets for the development of new antimalarial drugs.[12,13]

Apical membrane antigen 1 (AMA1) is a well-characterized protein involved in parasite invasion that is being studied as a potential target of new antimalarial drugs.[13] After release from the merozoite micronemes, *Pf*AMA1 forms a complex with another parasite antigen, *Pf*RON2, which is located at the moving junction.[14,15] Recent conditional knockdown studies have generated some controversy regarding the precise role of this complex,[16,17] but vaccine studies have nonetheless shown that targeting AMA1 can markedly inhibit development of asexual blood-stages of *P. falciparum*,[18−20] and antibodies and peptides targeting AMA1 have been shown to inhibit erythrocyte invasion by blocking the AMA1−RON2 interaction.[21−23] Structures of the AMA1 ectodomain revealed the presence of a hydrophobic cleft that is conserved across apicomplexan parasites[24] and has been suggested to be the principal ligand-binding site on AMA1.[21,25] The recent co-crystal structure of AMA1 with RON2L, a peptide corresponding to the extracellular region of RON2, confirmed that the hydrophobic cleft is the site of RON2 binding.[26,27] This structure also revealed a conformational change in which RON2L displaces the AMA1 domain II (DII) loop to expose a positively charged binding site that interacts with a complementary charged region at the N-terminus of peptide RON2L.[26,27] R1, an AMA1-binding peptide selected from a peptide library displayed on phage,[25,28] competes with RON2L for binding to this site and induces a similar conformational change of the AMA1 DII loop.[27]

A diverse range of ligands, including the peptides RON2L and R1, monoclonal antibodies and shark antibodies (IgNARs), bind to the hydrophobic cleft and block the interaction between AMA1 and RON2.[25,27,29−31] This has stimulated research to identify inhibitors of this interaction that may be valuable drug leads. Two approaches, high-throughput screening and fragment-based ligand discovery, have been used to identify small molecules that bind to AMA1 and compete with peptide RON2L or R1.[32,33] To understand the mechanisms by which these small molecules bind to AMA1, and to better

Received: March 13, 2014
Published: July 28, 2014

dx.doi.org/10.1021/jm500390g | J. Med. Chem. 2014, 57, 6419−6427

evaluate their potential as inhibitors of the AMA1−RON2 interaction, we have explored methods that would allow detection of conformational changes of the DII loop induced by ligand binding.

^{19}F NMR spectroscopy has proved sensitive to ligand-induced conformational changes in a range of other proteins of biological and therapeutic interest.[34−37] In this study, we have applied ^{19}F NMR spectroscopy to the detection of ligand-induced movements of the DII loop of *Pf*AMA1. Tryptophan residues were inserted into specific sites within the DII loop by site-directed mutagenesis, then all Trp in the protein were substituted with 5-fluorotryptophan (5F-Trp) by expression in the presence of glyphosate. ^{19}F NMR was then used to identify small AMA1-binding molecules among our current fragment hits[32] that induced a conformational change in AMA1 similar to that of known inhibitors. This approach will be used to characterize other AMA1-binding molecules that are being assessed as drug leads. More generally, it highlights the power of ^{19}F NMR as an efficient probe of protein−ligand binding and associated conformational change.

RESULTS AND DISCUSSION

Construction of AMA1 Mutants. Because none of the four Trp residues in AMA1 is in the flexible DII loop, mutant forms of 3D7 *Pf*AMA1$_{108-434}$ (corresponding to domains I and II of the AMA1 ectodomain) were created, either by replacing individual aromatic residues in the loop with Trp or by inserting an additional Trp in the middle of the loop. The insertion site was chosen at a highly exposed site, between two small residues (A372 and S373), where insertion was deemed least likely to perturb the structure of the DII loop (Figure 1A). Each of the mutant proteins was expressed on the surface of phage[38] so that they could readily be assessed for their ability to bind the conformationally dependent monoclonal antibodies (mAb) 1F9 and 4G2[31] and peptide RON2L (Table 1). All three reagents bound the replacement mutant F367W and the insertion mutant IS373W equivalently to wild-type AMA1 (Figure 1). In contrast, the binding of the mAbs and peptide to the other replacement mutants was reduced (F360W, F379W, and F385W) or abolished (Y353W and Y390W). Consequently, mutants F367W and IS373W were selected for the ^{19}F NMR studies described below. These mutants, together with wild-type AMA1 DI+II, were expressed in *Escherichia coli* as 5F-Trp-labeled recombinant proteins by growing the bacteria in minimal media supplemented with glyphosate to inhibit aromatic amino acid metabolism,[39] and with 5F-Trp, Phe, and Tyr.

Expression, Purification, and Characterization of F-Trp-AMA1 and Its Mutants. The 5F-Trp-labeled recombinant proteins were isolated and refolded following the previously published procedure,[40] modified by the addition of 0.6 M L-arginine to the refolding buffer. This resulted in a high yield of each form of AMA1 DI+II, all of which were judged to be >95% pure by nonreducing SDS-PAGE. Mass spectrometry confirmed the identity of the expressed constructs and was consistent with 5F-Trp incorporation to the level of approximately 85% (data not shown). The conformationally dependent mAb 1F9 reacted well on Western blots with all three forms of AMA1, indicating the presence of correctly refolded protein (Figure 2A). The binding affinities of 5F-Trp-AMA1 for the peptides RON2L and R1 were found to be 5 and 23 nM, respectively, using SPR measurements. The replacement mutant F367W bound both peptides with affinities nearly identical to those of wild-type AMA1 DI+II, whereas the binding affinities of the insertion mutant IS373W were slightly reduced (Figure 2B). Moreover, 5F-Trp-AMA1 and the two mutants gave similar circular dichroism (Supporting Information, Figure S1) and 1-D ^{1}H NMR spectra (Supporting

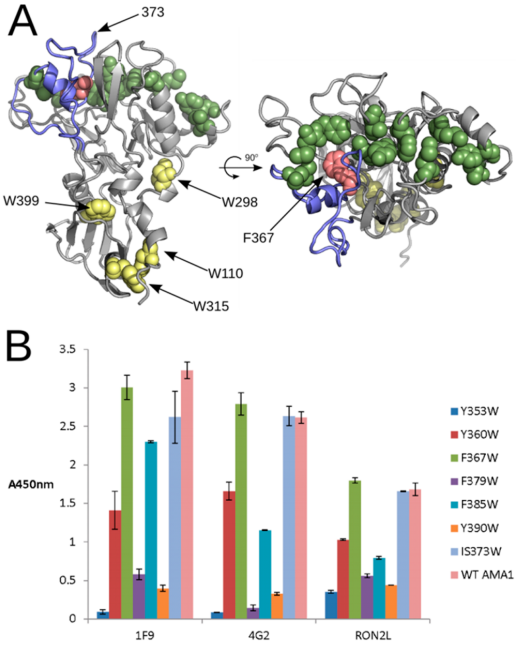


Figure 1. (A) Location of tryptophan residues and the domain II loop in AMA1. The structure of AMA1 DI+II (PDB ID 2Z8V) showing the locations of the four tryptophan residues, W110, W298, W315, and W399 (yellow). None of the four Trp residues is located in the domain II loop (blue) or hydrophobic cleft (green). Also shown are the sites of Trp mutation, F367 (red), and insertion, S373. (B) mAb reactivity of AMA1 DII loop mutants displayed on the surface of phage. Phage displaying wild-type and mutant forms of AMA1 were allowed to bind to immobilized mAb 1F9, mAb 4G2, and RON2L peptide. The average of duplicated wells is shown, and error bars indicate the range.

Table 1. Comparison of the Reactivity of Phage-Displayed AMA1 DII Loop Mutants and Wild-Type AMA1[a]

	reactivity with AMA1-binding ligands		
phage-displayed AMA1	mAb 1F9	mAb 4G2	peptide RON2L
wild-type AMA1	+++	+++	+++
replacement mutant Y353W	−	−	−
replacement mutant Y360W	+	++	++
replacement mutant F367W*	+++	+++	+++
replacement mutant F379W	+	−	+
replacement mutant F385W	++	+	+
replacement mutant Y390W	−	−	−
insertion mutant IS373W*	+++	+++	+++

[a]* indicates the two mutant forms of AMA1 DI+II selected for expression in *E. coli* as 5F-Trp recombinant proteins. +++ indicates an OD >2.5 for mAbs and >1.5 for peptide, ++ indicates an OD >1.5 for mAbs and >1 for peptide, + indicates an OD >0.5 for mAbs and >0.5 for peptide, − indicates an OD <0.5 for mAbs and <0.5 for peptide.

dx.doi.org/10.1021/jm500390g | J. Med. Chem. 2014, 57, 6419−6427

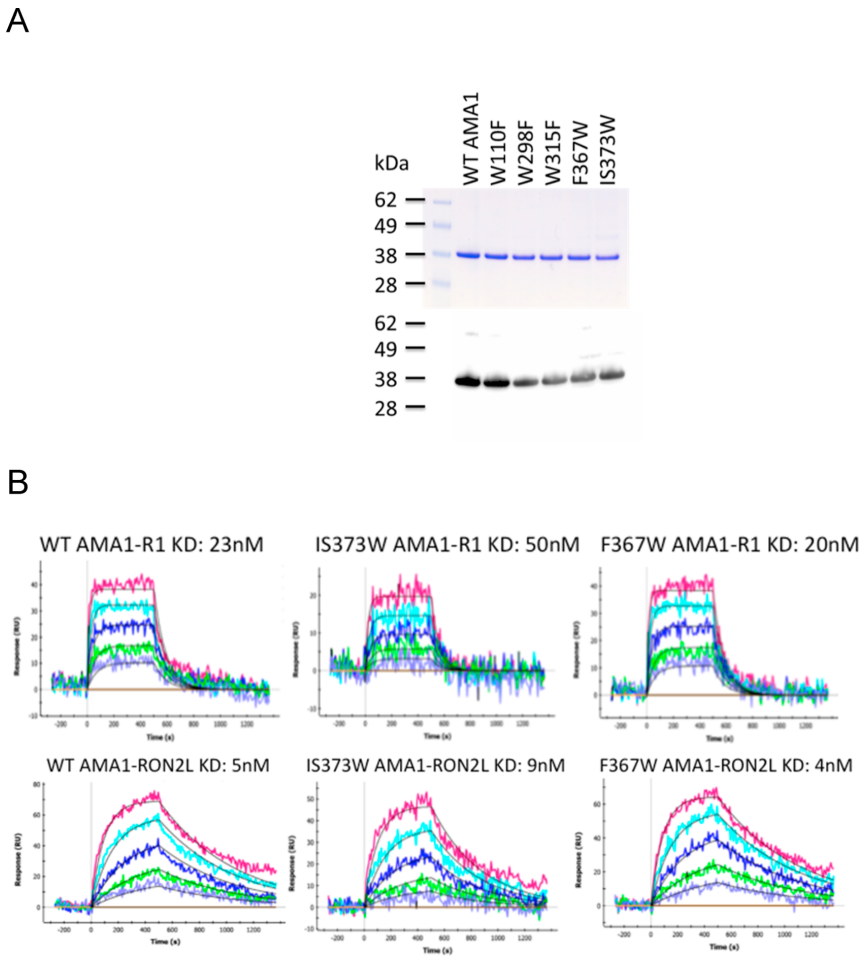


Figure 2. 5F-Trp-labeled AMA1 and its mutant forms are highly pure and functional. (A) 2 μg of 5F-Trp-labeled proteins were subjected to SDS-PAGE followed by Coomassie blue staining (top panel). After transferring to nitrocellulose membrane, the protein reacted with mAb 1F9 (bottom panel). (B) SPR sensorgrams for the binding of 5F-Trp-labeled AMA1 proteins to peptides RON2L and R1 (50−3.1 nM).

Information, Figure S2). This indicated that all three 5F-Trp-labeled proteins were folded correctly and potentially suitable for use in ^{19}F NMR studies to monitor movement of the DII loop induced by ligand binding.

Assignment of 5F-Trp ^{19}F Resonances in Native AMA1. The NMR spectrum of 5F-Trp-AMA1 DI+II consists of four peaks between −45 and −50 ppm, consistent with the presence of four Trp residues in this construct (Figure 3A). To assign these peaks to specific 5F-Trp residues, additional AMA1 mutants were constructed in which the native Trp residues were individually mutated to Phe or Tyr. These mutants were expressed as 5F-Trp-labeled recombinant proteins and judged to be >95% pure on a nonreducing SDS-PAGE (data not shown). Three mutants, W110F, W298F, and W315F, were correctly folded, as judged by reactivity with mAb 1F9 on Western blots (Figure 2A) and ^{1}H NMR spectra (Supporting Information, Figure S3), but attempts to mutate W399 failed to yield correctly folded AMA1.

Each peak in the ^{19}F NMR spectrum of 5F-Trp-AMA1 DI+II was assigned to one of the native Trp residues by identifying the missing peak in the successful Trp to Phe point mutants. W110, W315, and W298 in 5F-Trp-AMA1 were unequivocally assigned to resonances at chemical shifts of −46.8, −48.9, and −49.1 ppm, respectively (Figure 3A), and the assignment of W399 to the resonance at −45.8 ppm was made by elimination. The W110 peak shifted upfield in the ^{19}F NMR spectrum of 5F-Trp-AMA1[W315F], whereas the W315 peak shifted downfield in 5F-Trp-AMA1[W110F] (Figure 3A); these two Trp residues are within 4 Å of one another in the AMA1 structure (Supporting Information, Figure S1), so this mutual perturbation was not unexpected.

^{19}F NMR Spectra of AMA1 DII Loop Mutants. When the insertion and replacement mutant proteins (IS373W and

dx.doi.org/10.1021/jm500390g | J. Med. Chem. 2014, 57, 6419−6427

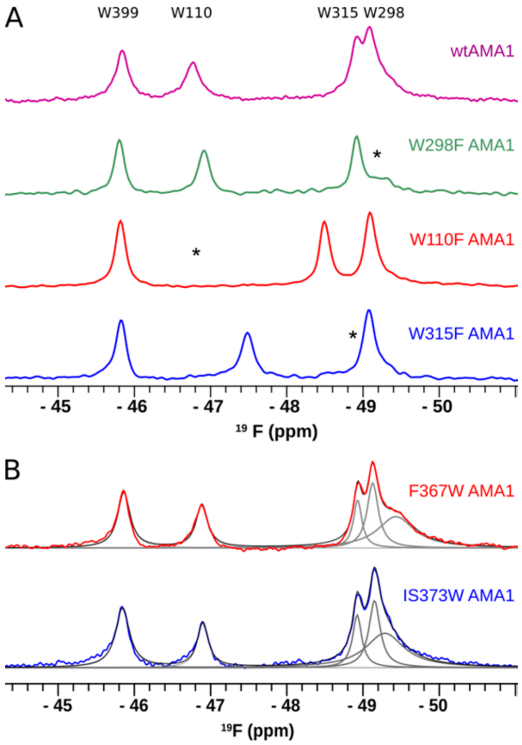


Figure 3. ^{19}F NMR spectra of 5F-Trp-AMA1 and its mutants. (A) Assignments of the native Trp resonances using Trp-to-Phe mutants. The expected location of the peak missing in each mutant spectrum is indicated by an asterisk, with the resulting resonance assignments as labeled. (B) Spectra of 5F-Trp F367W AMA1 (red) and 5F-Trp IS373W AMA1 (blue) are fit to a sum of five Lorentzian peaks (black; individual components in gray), showing that both spectra are well described by the four wild-type 5F-Trp signals plus a single broad peak from the introduced Trp.

F367W, respectively) were examined by ^{19}F NMR, no additional sharp peak was seen in either spectrum (Figure 3B). In the spectrum of 5F-Trp-AMA1[F367W], there was a broad peak partially overlapping the W298 resonance, at around −49.4 ppm, which is attributable to the 5F-Trp introduced at position 367. Its line shape is presumably the result of conformational exchange in the environment of W367 on microsecond to millisecond time scales. The significant overlap of the W367 and W298 resonances precludes a detailed characterization of this exchange. In contrast, 5F-Trp-AMA1-[IS373W] has a ^{19}F NMR spectrum almost identical to that of 5F-Trp-AMA1 (Figure 3B), the only difference being a small increase in signal intensity in the region around the W298 resonance. This indicates an almost complete overlap of the introduced W373 resonance with that of W298. Despite the fact that the 5F-Trp signals introduced in these constructs are, to a greater or lesser extent, obscured by overlap, ligand-induced changes to the conformation and/or dynamics of the DII loop are expected to cause detectable changes to the chemical shift and or line shape of these signals. For this reason, both 5F-Trp-AMA[F367W] and 5F-Trp-AMA1[IS373W] were used in further NMR studies to assess whether either mutant would serve as a probe of conformational change in the DII loop.

^{19}F NMR Spectra Reveal Conformational Change in the DII Loop. As a first step in establishing the sensitivity of ^{19}F NMR spectra of 5F-Trp-AMA1[F367W] and 5F-Trp-AMA1[IS373W] to conformational change in the DII loop, we tested the response of these spectra to binding of the peptide ligands RON2L and R1. A significant change in the ^{19}F NMR of 5F-Trp-AMA1[F367W] was observed in the presence of RON2L, with the broad peak at −49.4 ppm, attributed to the introduced W367, becoming sharper (Figure 4A). The sharpening may reflect a quenching of the exchange process that causes the W367 resonance in the absence of peptide to be

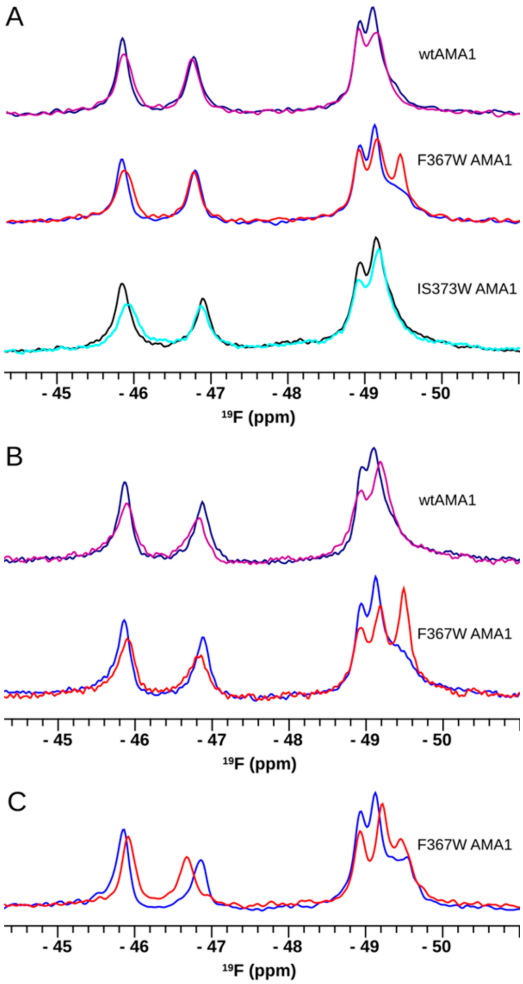


Figure 4. ^{19}F NMR spectra of F-Trp-AMA1 and its DII loop mutants in complex with RON2L (A), R1 (B), and trRON2 (C). Spectra of wild-type AMA1 DI+II in the presence and absence of peptide are in purple and navy, F367W AMA1 in red and blue, and IS373W AMA1 in black and cyan. R1 and RON2 are present at 120 μM (20% molar excess with respect to AMA1) and trRON2 at 500 μM (5-fold excess).

dx.doi.org/10.1021/jm500390g | J. Med. Chem. 2014, 57, 6419−6427

unusually broad or a significant increase in the flexibility of W367 relative to the rest of AMA1. In the light of X-ray crystallographic studies on cocrystals of AMA1 and RON2L, which indicate that the DII loop is displaced from the hydrophobic cleft by the binding of the peptide and is not resolved in the electron density,[27] we consider the latter explanation more likely. The chemical shift of the W367 resonance in the presence of RON2L is essentially identical to that of free 5F-Trp under these conditions (data not shown), consistent with the observed signal arising from an unstructured and solvent-exposed loop.

Similar sharpening of the peak attributed to the inserted W373 was seen in the ^{19}F NMR spectrum of the insertion mutant 5F-Trp-AMA1[IS373W] upon binding of the RON2L peptide. However, this effect is obscured by overlap with the W298 peak, which itself appears to be slightly broadened by the addition of RON2L (Figure 4A). In light of this result, the insertion mutant was not used for further NMR studies of the ligand-induced conformational change of the AMA1 DII loop. When we examined the effect of RON2L binding on the ^{19}F NMR spectrum of 5F-Trp-AMA1, minor changes were also seen in the resonances of the wild-type tryptophans. Specifically, we observed slight broadening and chemical shift changes to the peaks of W298 and W399, which presumably reflect changes to the conformation or dynamics of AMA1. Such changes are subtle, as they are not obviously resolved in the crystal structure of the AMA1−RON2L complex, thus emphasizing the extreme sensitivity of ^{19}F NMR signals to changes of this type.

When R1 was bound to 5F-Trp-AMA1[F367W], we saw similar spectral changes to those observed in the presence of RON2L. The line width of the 5F-Trp 367 resonance decreased, such that it became the sharpest peak in the spectrum (Figure 4B, Table S1, Supporting Information). This supports our interpretation that the decrease in line width is attributable to an increase in conformational flexibility of W367 on a nanosecond time scale. Under the alternative explanation, that the line width is decreased by the quenching of intermediate time scale exchange processes, we would expect the W367 resonance to reach approximately the same line shape as the other (structured) tryptophans in AMA1 DI+II. Furthermore, the chemical shift of the 5F-Trp 367 resonance is identical in the presence of R1 and RON2L, as well as other ligands (below), as would be expected if the DII loop is displaced, rather than making direct interaction with any of the ligands studied. Thus, the sharpening of the W367 resonance is indicative of the enhanced mobility of this region of the DII loop in AMA1 in the bound state. This had been inferred from crystallographic studies[27] but not hitherto observed directly.

To further explore the sensitivity of the introduced 5F-W367 resonance to conformational changes in the DII loop, we used a truncated RON2 peptide (trRON2), which contained the C-terminal disulfide-bonded loop of RON2L but not the N-terminal negatively charged residues. An examination of the structure of the AMA1−RON2L complex suggested that this trRON2 would not directly contact or displace the DII loop and, therefore, it would be expected to induce changes in the DII loop distinct from those induced by the longer RON2L peptide. As a result of the loss of the interactions made by the N-terminal residues of RON2L, the affinity of trRON2 is significantly reduced relative to the longer RON2 peptide, with a K_d measured for the trRON2−3D7 AMA1 interaction by SPR of approximately 40 μM (data not shown). To account for this reduced affinity and to ensure that the interaction was near saturation, trRON2 was added to 5F-Trp-AMA1[F367W] in significant excess. Despite this, trRON2 induced much less sharpening of the W367 resonances than when RON2L or R1 was the ligand (Figure 4C). It is possible, therefore, that trRON2 increases the flexibility of the DII loop, albeit to a much lesser extent than do RON2L and R1. Alternatively, the changes induced by trRON2 may reflect a quenching of the exchange process that broadens the W367 resonance in the absence of ligand. In either case, this indicates that trRON2 does indeed induce a distinct conformational change in the region around W367. The ^{19}F-NMR spectrum of the 5F-Trp-AMA1[F367W]−trRON2L complex provided additional evidence of small, ligand-induced structural changes elsewhere in AMA1 DI+II; again, there was an upfield shift of the resonances of W399 and W298 as well as a more marked downfield shift of the W110 resonance.

Characterization of Nonpeptidic Inhibitors Using ^{19}F NMR. We then explored the utility of this ^{19}F NMR-based probe of the DII loop conformation in current efforts to develop inhibitors of the AMA1−RON2 interaction as potential antimalarial drugs. To this end, we tested a series of molecules from our current fragment-based drug discovery efforts directed against AMA1.[32] Each of these molecules binds AMA1 DI+II and is competed by R1, indicating that they bind in the vicinity of the hydrophobic cleft. However, the structures of R1 and RON2L in complex with AMA1 identify an extensive area of interaction that includes a number of defined subsites, all of which represent possible binding sites for these fragments.[13,41] It remains unclear to which of these sites each fragment binds, and their impact on the conformation of AMA1 has not yet been determined.

To monitor the conformational changes of AMA1 upon the binding of these compounds, the spectra of 5F-Trp-AMA1-[F367W] were recorded with increasing amounts of the compounds of interest (Figure 5). Each of these compounds was tested at concentrations well above their respective K_d, as determined by SPR.[32] For a subset of tested compounds, we detected no significant change in the line shape of the 5F-Trp 367 resonance, suggesting that these molecules do not alter the conformation of the DII loop. In contrast, other compounds, exemplified here by **3**, caused a concentration-dependent decrease in the W367 line width, albeit to a much lesser extent than the peptide ligands R1 and RON2L (Figure 5). This may reflect the displacement of the DII loop by **3**, as occurred for the peptides, with the reduced magnitude of the effect being a consequence of the lower affinity of the small molecule. On the other hand, the change in line shape caused by **3** may be due to the quenching of exchange processes that contribute to the intrinsic line width of the W367 signal. In either case, these data indicate that **3** binds at the AMA1 hydrophobic cleft and modulates the conformational behavior of the DII loop. In contrast, although **1** and **2** both bind at the hydrophobic cleft, they do not affect the DII loop, suggesting that they bind to a distinct region of the cleft from the binding site of **3**.

Both **2** and **3** caused concentration-dependent downfield shifts of the W110 peak. Similar shifts were observed in the presence of trRON2 and, to a lesser extent, in the presence of both R1 and RON2L. **1**, in contrast, caused an upfield shift of this peak. These changes indicate some change to the local environment of W110 upon binding of each of these molecules. Each of these ligands has been shown directly (by crystallography) or indirectly (by competition with R1) to

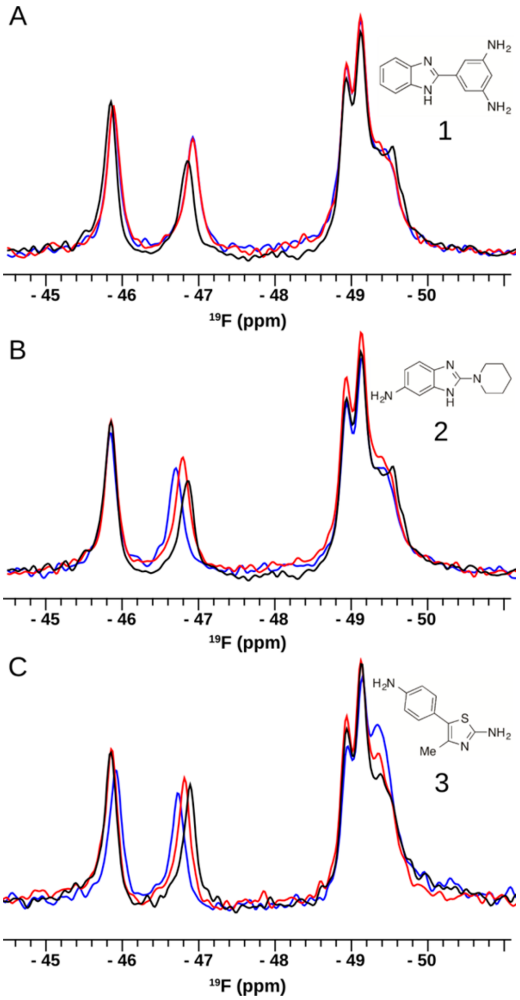


Figure 5. ^{19}F NMR spectra of 5-5F-Trp-F367W AMA1 bound to small molecule ligands **1** (A), **2** (B), and **3** (C). Spectra are shown for free F-Trp-F367W AMA1 (black) and in the presence of 1 mM (red) and 3 mM (blue) ligand.

bind to the hydrophobic cleft, more than 30 Å from W110. Although we cannot rule out the possibility of a secondary binding site for each of these ligands close to W110 on DII, the generality of this effect across peptide and small molecule ligands suggests a long-range conformational coupling between the two sites, such that the binding of these ligands at the hydrophobic cleft induces conformational change at this distant site. The detection of allosteric structural change of this nature by ^{19}F NMR is well established.[34,42] As noted above, any such conformational coupling is not evident from the available crystal structures, suggesting that it may be dynamic in nature. It is, however, consistent with the extensive changes seen in the ^{1}H−^{15}N HSQC spectrum of AMA1 DI+II on binding these ligands (our unpublished data).

CONCLUSIONS

Ligand-induced conformational changes in proteins are widespread. They are frequently functionally important and often present challenges for the development of inhibitors. As such, there is demand for simple and efficient means to characterize these changes. Overwhelmingly, such methods depend on the introduction of specific probes at strategic sites within the target protein, typically by the introduction and labeling of thiols in the form of free Cys residues. The presence of disulfides in the target protein complicates such strategies, as correct protein refolding and labeling specificity are difficult to achieve. We have addressed these issues here in the development of a simple and generally applicable ^{19}F NMR method involving the site-specific introduction of Trp residues and labeling with 5F-Trp. The environmental sensitivity of ^{19}F chemical shifts permits specificity with respect to native Trp residues. Applied to the emerging malaria drug target, AMA1, this method provides direct evidence for a ligand-induced increase in the flexibility of the DII loop that flanks and partially obscures the binding site of the native ligand, RON2. In addition, it identifies distinct conformational consequences of binding ligands from our current fragment-based drug discovery efforts, suggesting that these fragments may be binding at distinct subsites within the hydrophobic cleft of AMA1. This method will be an important tool in ongoing efforts to develop and exploit AMA1 ligands as inhibitors of host cell invasion by the malaria parasite.

EXPERIMENTAL SECTION

Materials. RON2L (DITQQAKDIGAGPVASCFTTRMSPPQQICLNSVVNTALS), trRON2 (CFTTRMSPPQQIC), and R1 (VFAEFLPLFSKFGSRMHILK) were sourced commercially at >90% purity (GL Biochem, China). Both peptides with two cysteine residues were cyclized with correctly formed intramolecular disulfide bonds. AMA1 binding fragments 5-(1*H*-benzo[*d*]imidazo-2-yl)benzene-1,3-diamine (**1**),[43] 2-(piperidin-1-yl)-1*H*-benzo[*d*]imidazol-6-amine (**2**),[44] and 5-(4-aminophenyl)-4-methylthiazol-2-amine (**3**)[45] were synthesized by established procedures.

Phage-Displayed AMA1 Mutants. All AMA1 mutant genes in the pHENH6 vector were transformed into *Escherichia coli* TG1 strain, and the cells were grown at 37 °C in 2YT medium with shaking until the culture reached log phase. The culture was subsequently left stationary at room temperature for 15 min to allow the regeneration of the F-pilus. Phages were then rescued from the *E. coli* cells by adding 10^8 helper phage, followed by incubation at 37 °C stationary for 1 h. Finally, 70 μg/mL of kanamycin was added to the culture, which was then incubated at 37 °C overnight with shaking. The overnight culture was centrifuged at 7700*g* for 10 min, and phage in the supernatant was precipitated using 5× PEG solution (30% (w/v) PEG8000, 2.6 M NaCl). Following incubation at 4 °C for 2 h, the precipitated phage were centrifuged at 15000*g* for 30 min, and the phage pellet was resuspended in 1 mL of PBS and stored at −80 °C for subsequent analysis.

Phage ELISA. Microtiter plates were coated with 0.2 μg/well of mAbs 1F9 or 4G2 overnight at 4 °C. For peptide RON2L, a neutravidin plate was coated overnight with 0.2 μg/well of biotinylated RON2L peptide in PBS/0.1% Tween/0.05% BSA. Phage expressing the desired proteins were diluted 1 in 10 in PBS−Tween and allowed to interact with the immobilized antigen for 1 h with shaking. After washing, the phage were subsequently detected using a horseradish peroxidase (HRP)-conjugated anti-M13 monoclonal antibody. Binding was visualized using 100 μL/well of 3,3′,5,5′-tetramethylbenzidine solution, and the reaction was stopped by adding 100 μL/well of 1 M HCl. The absorbance was then read at 450 nm.

Recombinant 5F-Trp-AMA1 and Its Mutants Proteins. The genes for wild-type 3D7 *Pf*AMA1$_{108-434}$ and its mutants W110F,

W298F, W315F, F367W, and IS373W were cloned into the expression vector pET26b(+) and transformed into *E. coli* BL21 (DE3) strain. Bacteria were grown overnight at 37 °C in LB medium, containing 100 μg/mL of ampicillin, and subsequently diluted 1 in 10 in M9 medium supplemented with 2% glucose (w/v), 1× BME vitamins, 1× trace elements, 2 mM $MgSO_4$, and 0.1 mM $CaCl_2$. Then 1 g/L of glyphosate, 50 mg/L of Phe, 50 mg/L of Tyr, and 36 mg/L of 5F-Trp were added to the medium when the cell density reached ∼0.4. After shaking for 1 h, the expression of 5F-Trp-labeled proteins was induced by addition of 1 mM (final) IPTG for 3 h, then cells were harvested by centrifugation at 5000*g* for 15 min.

Protein Refolding, Purification, and Characterization. 5F-Trp-labeled AMA1 and its mutants were expressed, solubilized, and isolated as described.[20] The recombinant proteins were subsequently refolded overnight in the conventional refolding buffer with the addition of 0.6 M L-Arg.[46] To correctly form the five intramolecular disulfide bonds in AMA1 DI+II,[47] reduced glutathione and oxidized glutathione were then added to the refolding buffer at final concentrations of 0.1 and 0.025 M, respectively. After refolding, the proteins were purified through a HiTRAP QFF (GE Healthcare) ion-exchange column and the sample buffer was changed to 20 mM phosphate buffer, pH 7.4.

The binding of conformationally sensitive mAbs to F-Trp-AMA1 and its mutants was assessed using Western blot. All protein samples were run on SDS-PAGE and transferred to a nitrocellulose membrane. The membrane was subsequently probed with 1 μg/mL of monoclonal antibody 1F9, and proteins were detected using HRP-conjugated antimouse antibody (1 in 2000 dilution, Millipore).

Surface Plasmon Resonance. All SPR analyses were performed using ProteOn XPR 36 platform (BioRad) and carried out in PBS−Tween (0.1%) at 25 °C using a constant flow rate of 30 μL/min. AMA1 was immobilized on a GLC sensor chip following the manufacturer's recommended protocol. Peptides RON2L and R1 were flowed across the surface of the chip at a constant flow rate for 8 min, then dissociated from the AMA1 surface for 800 s by flowing PBS−Tween across the surface of the chip. Peptides were tested at a range of concentrations (50−3.125 nM) and a lane without AMA1 was included as a negative control. Signals generated by buffer-only or DMSO-only control experiments were subtracted from the resulting sensorgrams to indicate the binding of the peptides to AMA1. All peptide samples were diluted in PBS−Tween, and the surface of the chip was regenerated by injecting 30 μL of 0.85% phosphoric acid (flow rate = 100 μL/min) in all channels in the vertical orientation of the ProteOn XPR36 fluidics.

NMR Spectroscopy. ^{19}F NMR spectra were acquired on a Bruker Avance III spectrometer operating at 564 MHz (600 MHz ^1H frequency) and at 25 °C using a TCI triple-resonance cryoprobe, with the ^1H channel tuned for ^{19}F. Samples contained ∼100 μM 5F-Trp AMA1 in 20 mM sodium phosphate, pH 7.4, with 0.005% trifluoroethanol as internal reference for ^{19}F chemical shift (−1.375 ppm relative to trace TFA in water) and signal intensity and 7% ^2H_2O for the spectrometer lock signal. R1 and RON2L were added from 1 mM stocks in 20 mM sodium phosphate, pH 7.4, to a final peptide:AMA1 ratio of 1.2:1, while trRON2 was used at a ratio of 5:1 to account for its lower affinity. Small molecule ligands of AMA1 were added from 100 mM stocks in H_2O (**2**), ethanol (**1**), or DMSO (**3**). Spectra were typically acquired as 2K scans of 8K complex points over a 150 ppm sweep width, for a total experimental time of less than 40 min. Spectra were processed using 20 Hz exponential line-broadening prior to Fourier transformation. Line widths were estimated by least-squares fits to a sum of Lorentzians.

■ ASSOCIATED CONTENT

Supporting Information

Table showing ^{19}F line widths of 5F-Trp resonances of F367W AMA1 in the presence of RON2L and R1; figure showing the location of the native Trp residues in AMA1, and figures showing ^1H NMR spectra of AMA1 mutants. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*For R.S.N.: phone, +61 3 9903 9167; fax, +61 3 9903 9582; E-mail, ray.norton@monash.edu.

*For M.F.: phone, +61 3 9479 2158; E-mail, m.foley@latrobe.edu.au.

Author Contributions

§X.G. and C.A.M. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Krishnarjuna Bankala for assistance with mass spectrometry. This work was supported in part by an Australian National Health and Medical Research Council (NHMRC) project grant (1025150). R.S.N. acknowledges fellowship support from the NHMRC.

■ ABBREVIATIONS USED

5F-Trp, 5-fluorotryptophan; AMA1, apical membrane antigen 1; DII, domain II; DI+II, domains I and II; mAb, monoclonal antibody

■ REFERENCES

(1) *World Malaria Report 2013*; World Health Organization: Geneva, 2013.

(2) Stoute, J. A.; Slaoui, M.; Heppner, D. G.; Momin, P.; Kester, K. E.; Desmons, P.; Wellde, B. T.; Garcon, N.; Krzych, U.; Marchand, M.; RTS,S Malaria Vaccine Evaluation Group. A preliminary evaluation of a recombinant circumsporozoite protein vaccine against *Plasmodium falciparum* malaria. *N. Engl. J. Med.* **1997**, *336*, 86−91.

(3) Casares, S.; Brumeanu, T. D.; Richie, T. L. The RTS,S malaria vaccine. *Vaccine* **2010**, *28*, 4880−4894.

(4) Agnandji, S. T.; Lell, B.; Fernandes, J. F.; Abossolo, B. P.; Methogo, B. G.; Kabwende, A. L.; Adegnika, A. A.; Mordmuller, B.; Issifou, S.; Kremsner, P. G.; Sacarlal, J.; Aide, P.; Lanaspa, M.; Aponte, J. J.; Machevo, S.; Acacio, S.; Bulo, H.; Sigauque, B.; Macete, E.; Alonso, P.; Abdulla, S.; Salim, N.; Minja, R.; Mpina, M.; Ahmed, S.; Ali, A. M.; Mtoro, A. T.; Hamad, A. S.; Mutani, P.; Tanner, M.; Tinto, H.; D'Alessandro, U.; Sorgho, H.; Valea, I.; Bihoun, B.; Guiraud, I.; Kabore, B.; Sombie, O.; Guiguemde, R. T.; Ouedraogo, J. B.; Hamel, M. J.; Kariuki, S.; Oneko, M.; Odero, C.; Otieno, K.; Awino, N.; McMorrow, M.; Muturi-Kioi, V.; Laserson, K. F.; Slutsker, L.; Otieno, W.; Otieno, L.; Otsyula, N.; Gondi, S.; Otieno, A.; Owira, V.; Oguk, E.; Odongo, G.; Woods, J. B.; Ogutu, B.; Njuguna, P.; Chilengi, R.; Akoo, P.; Kerubo, C.; Maingi, C.; Lang, T.; Olotu, A.; Bejon, P.; Marsh, K.; Mwambingu, G.; Owusu-Agyei, S.; Asante, K. P.; Osei-Kwakye, K.; Boahen, O.; Dosoo, D.; Asante, I.; Adjei, G.; Kwara, E.; Chandramohan, D.; Greenwood, B.; Lusingu, J.; Gesase, S.; Malabeja, A.; Abdul, O.; Mahende, C.; Liheluka, E.; Malle, L.; Lemnge, M.; Theander, T. G.; Drakeley, C.; Ansong, D.; Agbenyega, T.; Adjei, S.; Boateng, H. O.; Rettig, T.; Bawa, J.; Sylverken, J.; Sambian, D.; Sarfo, A.; Agyekum, A.; Martinson, F.; Hoffman, I.; Mvalo, T.; Kamthunzi, P.; Nkomo, R.; Tembo, T.; Tegha, G.; Tsidya, M.; Kilembe, J.; Chawinga, C.; Ballou, W. R.; Cohen, J.; Guerra, Y.; Jongert, E.; Lapierre, D.; Leach, A.; Lievens, M.; Ofori-Anyinam, O.; Olivier, A.; Vekemans, J.; Carter, T.; Kaslow, D.; Leboulleux, D.; Loucq, C.; Radford, A.; Savarese, B.; Schellenberg, D.; Sillman, M.; Vansadia, P. A phase 3 trial of RTS,S/AS01 malaria vaccine in African infants. *N. Engl. J. Med.* **2012**, *367*, 2284−2295.

(5) Biamonte, M. A.; Wanner, J.; Le Roch, K. G. Recent advances in malaria drug discovery. *Bioorg. Med. Chem. Lett.* **2013**, *23*, 2829−2843.

(6) Anthony, M. P.; Burrows, J. N.; Duparc, S.; Moehrle, J. J.; Wells, T. N. The global pipeline of new medicines for the control and elimination of malaria. *Malar. J.* **2012**, *11*, 316.

(7) Yanow, S. K.; Purcell, L. A.; Lee, M.; Spithill, T. W. Genomics-based drug design targets the AT-rich malaria parasite: implications for antiparasite chemotherapy. *Pharmacogenomics* **2007**, *8*, 1267−1272.
(8) Lucet, I. S.; Tobin, A.; Drewry, D.; Wilks, A. F.; Doerig, C. *Plasmodium* kinases as targets for new-generation antimalarials. *Future Med. Chem.* **2012**, *4*, 2295−2310.
(9) Cowman, A. F.; Berry, D.; Baum, J. The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J. Cell Biol.* **2012**, *198*, 961−971.
(10) Aikawa, M.; Miller, L. H.; Johnson, J.; Rabbege, J. Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J. Cell Biol.* **1978**, *77*, 72−82.
(11) Pinder, J. C.; Fowler, R. E.; Dluzewski, A. R.; Bannister, L. H.; Lavin, F. M.; Mitchell, G. H.; Wilson, R. J.; Gratzer, W. B. Actomyosin motor in the merozoite of the malaria parasite, *Plasmodium falciparum*: implications for red cell invasion. *J. Cell Sci.* **1998**, *111*, 1831−1839.
(12) Santos, G.; Torres, N. V. New targets for drug discovery against malaria. *PLoS One* **2013**, *8*, e59968.
(13) MacRaild, C. A.; Anders, R. F.; Foley, M.; Norton, R. S. Apical membrane antigen 1 as an anti-malarial drug target. *Curr. Top Med. Chem.* **2011**, 2039−2047.
(14) Tyler, J. S.; Boothroyd, J. C. The C-terminus of *Toxoplasma* RON2 provides the crucial link between AMA1 and the host-associated invasion complex. *PLoS Pathog.* **2011**, *7*, e1001282.
(15) Lamarque, M.; Besteiro, S.; Papoin, J.; Roques, M.; Vulliez-Le Normand, B.; Morlon-Guyot, J.; Dubremetz, J.-F.; Fauquenoy, S.; Tomavo, S.; Faber, B. W.; Kocken, C. H.; Thomas, A. W.; Boulanger, M. J.; Bentley, G. A.; Lebrun, M. The RON2−AMA1 interaction is a critical step in moving junction-dependent invasion by Apicomplexan parasites. *PLoS Pathog.* **2011**, *7*, e1001276.
(16) Bargieri, D. Y.; Andenmatten, N.; Lagal, V.; Thiberge, S.; Whitelaw, J. A.; Tardieux, I.; Meissner, M.; Menard, R. Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion. *Nature Commun.* **2013**, *4*, 2552.
(17) Yap, A.; Azevedo, M. F.; Gilson, P. R.; Weiss, G. E.; O'Neill, M. T.; Wilson, D. W.; Crabb, B. S.; Cowman, A. F. Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cell Microbiol.* **2014**, *16*, 642−656.
(18) Spring, M. D.; Cummings, J. F.; Ockenhouse, C. F.; Dutta, S.; Reidler, R.; Angov, E.; Bergmann-Leitner, E.; Stewart, V. A.; Bittner, S.; Juompan, L.; Kortepeter, M. G.; Nielsen, R.; Krzych, U.; Tierney, E.; Ware, L. A.; Dowler, M.; Hermsen, C. C.; Sauerwein, R. W.; de Vlas, S. J.; Ofori-Anyinam, O.; Lanar, D. E.; Williams, J. L.; Kester, K. E.; Tucker, K.; Shi, M.; Malkin, E.; Long, C.; Diggs, C. L.; Soisson, L.; Dubois, M. C.; Ballou, W. R.; Cohen, J.; Heppner, D. G., Jr. Phase 1/2a study of the malaria vaccine candidate apical membrane antigen-1 (AMA-1) administered in adjuvant system AS01B or AS02A. *PLoS One* **2009**, *4*, e5254.
(19) Thera, M. A.; Doumbo, O. K.; Coulibaly, D.; Laurens, M. B.; Ouattara, A.; Kone, A. K.; Guindo, A. B.; Traore, K.; Traore, I.; Kouriba, B.; Diallo, D. A.; Diarra, I.; Daou, M.; Dolo, A.; Tolo, Y.; Sissoko, M. S.; Niangaly, A.; Sissoko, M.; Takala-Harrison, S.; Lyke, K. E.; Wu, Y.; Blackwelder, W. C.; Godeaux, O.; Vekemans, J.; Dubois, M.-C.; Ballou, W. R.; Cohen, J.; Thompson, D.; Dube, T.; Soisson, L.; Diggs, C. L.; House, B.; Lanar, D. E.; Dutta, S.; Heppner, D. G.; Plowe, C. V. A field trial to assess a blood-stage malaria vaccine. *N. Engl. J. Med.* **2011**, *365*, 1004−1013.
(20) Hodder, A. N.; Crewther, P. E.; Anders, R. F. Specificity of the protective antibody response to apical membrane antigen 1. *Infect. Immun.* **2001**, *69*, 3286−3294.
(21) Richard, D.; MacRaild, C. A.; Riglar, D. T.; Chan, J.-A.; Foley, M.; Baum, J.; Ralph, S. A.; Norton, R. S.; Cowman, A. F. Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. *J. Biol. Chem.* **2010**, *285*, 14815−14822.
(22) Dutta, S.; Dlugosz, L. S.; Drew, D. R.; Ge, X.; Ababacar, D.; Rovira, Y. I.; Moch, J. K.; Shi, M.; Long, C. A.; Foley, M.; Beeson, J. G.; Anders, R. F.; Miura, K.; Haynes, J. D.; Batchelor, A. H. Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1. *PLoS Pathog.* **2013**, *9*, e1003840.
(23) Srinivasan, P.; Beatty, W. L.; Diouf, A.; Herrera, R.; Ambroggio, X.; Moch, J. K.; Tyler, J. S.; Narum, D. L.; Pierce, S. K.; Boothroyd, J. C.; Haynes, J. D.; Miller, L. H. Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13275−13280.
(24) Bai, T.; Becker, M.; Gupta, A.; Strike, P.; Murphy, V. J.; Anders, R. F.; Batchelor, A. H. Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 12736−12741.
(25) Harris, K. S.; Casey, J. L.; Coley, A. M.; Masciantonio, R.; Sabo, J. K.; Keizer, D. W.; Lee, E. F.; McMahon, A.; Norton, R. S.; Anders, R. F.; Foley, M. Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* **2005**, *73*, 6981−6989.
(26) Tonkin, M. L.; Roques, M.; Lamarque, M. H.; Pugniere, M.; Douguet, D.; Crawford, J.; Lebrun, M.; Boulanger, M. J. Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide. *Science* **2011**, *333*, 463−467.
(27) Vulliez-Le Normand, B.; Tonkin, M. L.; Lamarque, M. H.; Langer, S.; Hoos, S.; Roques, M.; Saul, F. A.; Faber, B. W.; Bentley, G. A.; Boulanger, M. J.; Lebrun, M. Structural and functional insights into the malaria parasite moving junction complex. *PLoS Pathog.* **2012**, *8*, e1002755.
(28) Harris, K. S.; Casey, J. L.; Coley, A. M.; Karas, J. A.; Sabo, J. K.; Tan, Y.; Dolezal, O.; Norton, R. S.; Hughes, A. B.; Scanlon, D.; Foley, M. Rapid optimisation of a peptide inhibitor of malaria parasite invasion by comprehensive *N*-methyl scanning. *J. Biol. Chem.* **2009**, *284*, 9361−9371.
(29) Coley, A. M.; Gupta, A.; Murphy, V. J.; Bai, T.; Kim, H.; Foley, M.; Anders, R. F.; Batchelor, A. H. Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody. *PLoS Pathog.* **2007**, *3*, 1308−1319.
(30) Henderson, K. A.; Streltsov, V. A.; Coley, A. M.; Dolezal, O.; Hudson, P. J.; Batchelor, A. H.; Gupta, A.; Bai, T.; Murphy, V. J.; Anders, R. F.; Foley, M.; Nuttall, S. D. Structure of an IgNAR−AMA1 complex: targeting a conserved hydrophobic cleft broadens malarial strain recognition. *Structure* **2007**, *15*, 1452−1466.
(31) Collins, C. R.; Withers-Martinez, C.; Bentley, G. a.; Batchelor, A. H.; Thomas, A. W.; Blackman, M. J. Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J. Biol. Chem.* **2007**, *282*, 7431−7441.
(32) Lim, S. S.; Debono, C. O.; MacRaild, C. A.; Chandrashekaran, I. R.; Dolezal, O.; Anders, R. F.; Simpson, J. S.; Scanlon, M. J.; Devine, S. M.; Scammells, P. J.; Norton, R. S. Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening. *Aust. J. Chem.* **2013**, *66*, 1530−1536.
(33) Srinivasan, P.; Yasgar, A.; Luci, D. K.; Beatty, W. L.; Hu, X.; Andersen, J.; Narum, D. L.; Moch, J. K.; Sun, H.; Haynes, J. D.; Maloney, D. J.; Jadhav, A.; Simeonov, A.; Miller, L. H. Disrupting malaria parasite AMA1−RON2 interaction with a small molecule prevents erythrocyte invasion. *Nature Commun.* **2013**, *4*, 2261.
(34) Liu, J. J.; Horst, R.; Katritch, V.; Stevens, R. C.; Wüthrich, K. Biased signaling pathways in β2-adrenergic receptor characterized by ^{19}F-NMR. *Science* **2012**, *335*, 1106−1110.
(35) Pomerantz, W. C.; Wang, N.; Lipinski, A. K.; Wang, R.; Cierpicki, T.; Mapp, A. K. Profiling the dynamic interfaces of fluorinated transcription complexes for ligand discovery and characterization. *ACS Chem. Biol.* **2012**, *7*, 1345−1350.
(36) Luck, L. A.; Barse, J. L.; Luck, A. M.; Peck, C. H. Conformational changes in the human estrogen receptor observed by ^{19}F NMR. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 988−991.
(37) Evanics, F.; Kitevski, J. L.; Bezsonova, I.; Forman-Kay, J.; Prosser, R. S. ^{19}F NMR studies of solvent exposure and peptide

binding to an SH3 domain. *Biochim. Biophys. Acta* **2007**, *1770*, 221−230.
(38) Coley, A. M.; Campanale, N. V.; Casey, J. L.; Hodder, A. N.; Crewther, P. E.; Anders, R. F.; Tilley, L. M.; Foley, M. Rapid and precise epitope mapping of monoclonal antibodies against *Plasmodium falciparum* AMA1 by combined phage display of fragments and random peptides. *Protein Eng.* **2001**, *14*, 691−698.
(39) Kim, H. W.; Perez, J. A.; Ferguson, S. J.; Campbell, I. D. The specific incorporation of labelled aromatic amino acids into proteins through growth of bacteria in the presence of glyphosate. Application to fluorotryptophan labelling to the H(+)-ATPase of *Escherichia coli* and NMR studies. *FEBS Lett.* **1990**, *272*, 34−36.
(40) Gupta, A.; Bai, T.; Murphy, V.; Strike, P.; Anders, R. F.; Batchelor, A. H. Refolding, purification, and crystallization of apical membrane antigen 1 from *Plasmodium falciparum*. *Protein Expression Purif.* **2005**, *41*, 186−198.
(41) Tonkin, M. L.; Crawford, J.; Lebrun, M. L.; Boulanger, M. J. *Babesia divergens* and *Neospora caninum* apical membrane antigen 1 (AMA1) structures reveal selectivity and plasticity in apicomplexan parasite host cell invasion. *Protein Sci.* **2012**, *22*, 114−127.
(42) Song, L.; Teng, Q.; Phillips, R. S.; Brewer, J. M.; Summers, A. O. ^{19}F-NMR reveals metal and operator-induced allostery in MerR. *J. Mol. Biol.* **2007**, *371*, 79−92.
(43) Mikroyannidis, J. A. Aromatic polyamides and polyimides with benzimidazole or benzoxazinone pendent groups prepared from 5-(2-benzimidazole)- or 5-(2-benzoxazinone)-1,3-phenylenediamine. *Polymer* **1996**, *37*, 2715−2721.
(44) Carpenter, A. J.; Al-Barazanji, K. A.; Barvian, K. K.; Bishop, M. J.; Britt, C. S.; Cooper, J. P.; Goetz, A. S.; Grizzle, M. K.; Hertzog, D. L.; Ignar, D. M.; Morgan, R. O.; Peckham, G. E.; Speake, J. D.; Swain, W. R. Novel benzimidazole-based MCH R1 antagonists. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4994−5000.
(45) Lee, B. W.; Lee, S. D. [5,5] Sigmatropic shift of *N*-phenyl-*N*′-(2-thiazolyl)hydrazines and *N*,*N*′-bis(2-thiazolyl)hydrazines into 2-amino-5-(*p*-aminophenyl)thiazoles and 5,5′-bis(2-aminothiazole) derivatives. *Tetrahedron Lett.* **2000**, *41*, 3883−3886.
(46) Anders, R. F.; Crewther, P. E.; Edwards, S.; Margetts, M.; Matthew, M. L.; Pollock, B.; Pye, D. Immunisation with recombinant AMA-1 protects mice against infection with *Plasmodium chabaudi*. *Vaccine* **1998**, *16*, 240−247.
(47) Hodder, A. N.; Crewther, P. E.; Matthew, M. L.; Reid, G. E.; Moritz, R. L.; Simpson, R. J.; Anders, R. F. The disulfide bond structure of *Plasmodium* apical membrane antigen-1. *J. Biol. Chem.* **1996**, *271*, 29446−29452.

A.3 Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials

Research article

Journal of Molecular Recognition

Received: 14 September 2015, Accepted: 28 November 2015, Published online in Wiley Online Library: 24 January 2016

(wileyonlinelibrary.com) DOI: 10.1002/jmr.2529

Solution NMR characterization of apical membrane antigen 1 and small molecule interactions as a basis for designing new antimalarials

Bankala Krishnarjuna[a†], San Sui Lim[a†], Shane M. Devine[a], Cael O. Debono[a], Raymond Lam[a], Indu R. Chandrashekaran[a], Garima Jaipuria[b], Hiromasa Yagi[a], Hanudatta S. Atreya[b], Martin J. Scanlon[a], Christopher A. MacRaild[a], Peter J. Scammells[a] and Raymond S. Norton[a]*

***Plasmodium falciparum* apical membrane antigen 1 (*Pf*AMA1) plays an important role in the invasion by merozoites of human red blood cells during a malaria infection. A key region of *Pf*AMA1 is a conserved hydrophobic cleft formed by 12 hydrophobic residues. As anti-apical membrane antigen 1 antibodies and other inhibitory molecules that target this hydrophobic cleft are able to block the invasion process, *Pf*AMA1 is an attractive target for the development of strain-transcending antimalarial agents. As solution nuclear magnetic resonance spectroscopy is a valuable technique for the rapid characterization of protein–ligand interactions, we have determined the sequence-specific backbone assignments for *Pf*AMA1 from two *P. falciparum* strains, FVO and 3D7. Both selective labelling and unlabelling strategies were used to complement triple-resonance experiments in order to facilitate the assignment process. We have then used these assignments for mapping the binding sites for small molecules, including benzimidazoles, pyrazoles and 2-aminothiazoles, which were selected on the basis of their affinities measured from surface plasmon resonance binding experiments. Among the compounds tested, benzimidazoles showed binding to a similar region on both FVO and 3D7 *Pf*AMA1, suggesting that these compounds are promising scaffolds for the development of novel *Pf*AMA1 inhibitors. Copyright © 2016 John Wiley & Sons, Ltd.**

Additional supporting information may be found in the online version of this article at the publisher's web site.

Keywords: AMA1; NMR; resonance assignments; isotopic labelling; SPR; fragments

INTRODUCTION

Malaria is a significant global health burden, causing nearly 600 000 deaths per annum (WHO, 2014). The causative agents are *Plasmodium* parasites, with the majority of the disease burden arising from infection by *Plasmodium falciparum*. Importantly, this parasite has started to develop resistance to available antimalarial drugs (White, 2004), including first-line therapy artemisinin combination therapies (Breman, 2012). Hence, there is a pressing need to develop new antimalarial agents targeting resistant strains of *P. falciparum* parasites. A key step in the invasion by *Plasmodium* parasites of host red blood cells during the asexual blood stage of malaria is the formation of a highly conserved functional complex called the moving junction (Aikawa *et al.*, 1978). Two parasite proteins, *P. falciparum* apical membrane antigen 1 (*Pf*AMA1) and *P. falciparum* rhoptry neck protein 2 (*Pf*RON2), play a crucial role in the formation of the moving junction (Lamarque *et al.*, 2011; Srinivasan *et al.*, 2011; Tonkin *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012). During host cell invasion, *Pf*RON2 is secreted onto the red blood cell membrane and forms strong interactions with *Pf*AMA1 expressed on the parasite surface membrane (Vulliez-Le Normand *et al.*, 2012). Both *Pf*AMA1 and *Pf*RON2 are essential for the invasion process, and hence, small molecule inhibitors targeting this complex have considerable potential to be used as new antimalarials to combat resistant *P. falciparum* parasites (MacRaild *et al.*, 2011).

Plasmodium falciparum apical membrane antigen 1 is a type I integral membrane protein with three extracellular domains (DI + II + III). *Pf*AMA1 has a hydrophobic cleft, which is the site of interactions with *Pf*RON2 (Lamarque *et al.*, 2011; Srinivasan *et al.*, 2011; Tonkin *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012). The opposing ends of the cleft possess different properties that play

* *Correspondence to: R. S. Norton, Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia.*
E-mail: ray.norton@monash.edu

† *These authors contributed equally.*

a *B. Krishnarjuna, S. S. Lim, S. M. Devine, C. O. Debono, R. Lam, I. R. Chandrashekaran, H. Yagi, M. J. Scanlon, C. A. MacRaild, P. J. Scammells, R. S. Norton*
Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria, 3052, Australia

b *G. Jaipuria, H. S. Atreya*
NMR Research Centre, Indian Institute of Science, Bangalore 560012, India

distinct functional roles in the complex formation (Bai *et al.*, 2005). At one end, there is a conserved domain II loop (DII loop) that undergoes conformational changes upon *Pf*RON2 binding. The monoclonal antibody 4G2, which binds to the base of this DII loop, exhibits strain-transcending inhibitory effects on red cell invasion by *P. falciparum* parasites (Collins *et al.*, 2007). At the opposite end, there are numerous polymorphic residues that surround the *Pf*AMA1 hydrophobic cleft (Bai *et al.*, 2005; Lim *et al.*, 2014). These polymorphic residues are important in mediating immune escape from natural antibodies in the human host, and they play important roles in limiting the cross-strain inhibitory activities of ligands such as R1 (Harris *et al.*, 2005; Harris *et al.*, 2009) and F1 (Li *et al.*, 2002; Keizer *et al.*, 2003) peptides, as well as the monoclonal antibody 1 F9 (Coley *et al.*, 2006; Coley *et al.*, 2007).

A fragment screening campaign has identified a range of scaffolds that bind to the *Pf*AMA1 hydrophobic cleft as defined by competition with the peptide R1, which is known to bind to this cleft (Harris *et al.*, 2005; Harris *et al.*, 2009; Richard *et al.*, 2010; Lee *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012; Lim *et al.*, 2013; Ge *et al.*, 2014; Wang *et al.*, 2014). Whether these compounds interact with some of the polymorphic residues surrounding the cleft is an important question as compounds that interact with such residues would be considered poor starting points for chemical elaboration given that our goal is to develop strain-transcending inhibitors of *Pf*AMA1. Therefore, it is critical to determine the specific binding sites of the R1-competing fragments early in the lead discovery pipeline. The aim of this work is to use nuclear magnetic resonance (NMR) chemical shift perturbations (CSPs) to identify the residues that are perturbed upon fragment binding and to map them onto relevant *Pf*AMA1 crystal structures to define the binding sites.

MATERIALS AND METHODS

Protein expression and purification

Uniformly ^{15}N-labelled, ^{13}C-labelled, ^{2}H-labelled FVO *Pf*AMA1$_{[104–438]}$ and 3D7 *Pf*AMA1$_{[104–438]}$ proteins were expressed and purified as described previously (Lim *et al.*, 2013). Briefly, the protein was expressed in *Escherichia coli* BL21 (DE3) strain cultured in $^{2}H_2O$-based M9 medium containing ^{13}C-glucose and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources, respectively, using the high-cell-density method (Sivashanmugam *et al.*, 2009). The cells were lysed in 20 mM Tris pH 8.0 containing lysozyme, and the protein was purified as inclusion bodies, followed by solubilization in 6 M Gdn-HCl. The protein was purified under denaturing conditions on a Ni^{2+}–nitrilotriacetic acid (NTA) affinity column and subjected to refolding for 24 h in a buffer containing 20 mM Tris (pH 8.0), 0.5 M urea, 50 mM NaCl, 100 mM reduced glutathione and 25 mM oxidized glutathione. The folded material was purified by anion-exchange chromatography, subjected to tobacco etch virus (TEV) cleavage to remove the 6His tag, and finally purified by anion-exchange chromatography. The fractions containing pure protein according to sodium dodecyl sulfate polyacrylamide gel electrophoresis analysis were lyophilized in 20 mM ammonium bicarbonate buffer and stored at −80 °C. Uniformly ^{2}H-labelled and selectively ^{15}N-Lys-labelled FVO *Pf*AMA1 was prepared by adding 200 mg/l of ^{15}N-labelled Lys (Sigma-Aldrich, St. Louis, MO, USA) into the high-cell-density $^{2}H_2O$-based M9 medium 1 h prior to isopropyl β-D-1-thiogalactopyranoside (IPTG) induction. Similarly Lys, Arg, Asn and Ile, Leu, Val selectively unlabelled 3D7 *Pf*AMA1 samples were prepared by expressing the protein in $^{2}H_2O$-based M9 medium containing 1 g/l of the respective unlabelled amino acids and 1 g/l of $^{15}NH_4Cl$ (Krishnarjuna *et al.*, 2011).

Surface plasmon resonance binding experiments

The binding affinities of small molecules including benzimidazoles (MIPS-0001404, MIPS-0000865 and MIPS-0008405), pyrazoles (MIPS-0001160 and MIPS-0001176) and 2-aminothiazoles (MIPS-0000620 and MIPS-0008939) (Figure 1) were evaluated using a BIAcore T200 biosensor. 3D7 *Pf*AMA1 was immobilized on a CM5 sensor chip using amine coupling to achieve an immobilization level of 10 000 RU. A reference flow cell was prepared in a similar fashion, but in the absence of protein. R1 peptide (GL Biochem, Shanghai, China) was injected over the immobilized surface at a flow rate of 60 µl/min in an HBS-EP buffer (50 mM HEPES, 150 mM NaCl, 3.4 mM EDTA, 0.05% Tween 20, pH 7.4) containing 5% dimethyl sulfoxide at concentrations ranging from 31.3 to 1000 nM with a contact time of 2 min. Compounds were screened in the same buffer at concentrations ranging from 7 to 250 µM. All sensorgrams were double referenced and solvent corrected prior to analysis. Binding affinities were estimated based on steady-state analysis of the dose–response using a fixed small molecule R_{max}. The small molecule R_{max} was calculated with Eqn 1 using the R_{max} of R1 peptide derived from fitting the binding responses of R1 at steady state to a 1:1 binding model

$$R_{max[\text{small molecule}]} = R_{max[\text{R1 peptide}]} \times \left(MW_{[\text{small molecule}]} / MW_{[\text{R1 peptide}]}\right)$$
$$MW = \text{Molecular weight} \quad (1)$$

Experimental procedures for the synthesis of small molecules used in this work are described in the Supporting Information.

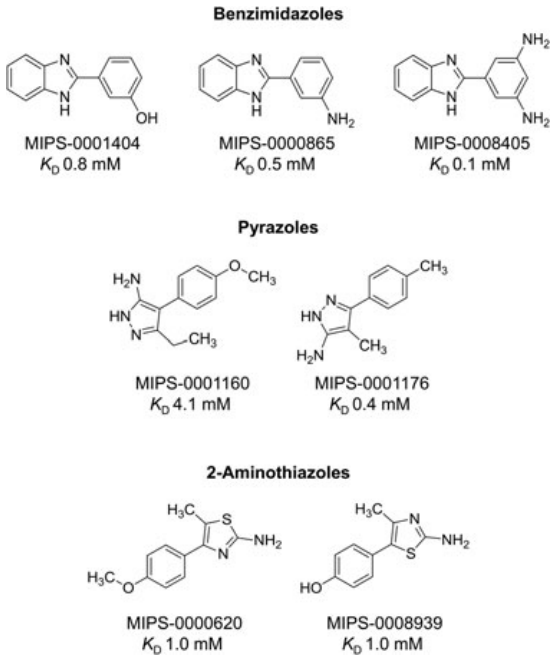


Figure 1. Structures of compounds, with their affinities (K_D) for apical membrane antigen 1 as measured by surface plasmon resonance.

Journal of Molecular Recognition

Nuclear magnetic resonance spectroscopy and resonance assignments

Nuclear magnetic resonance spectra for FVO *Pf*AMA1 was acquired at 35 °C on a Bruker Avance III 600-MHz spectrometer (Billerica, MA, USA) and for 3D7 *Pf*AMA1 at 30 °C on a Bruker Avance III 800-MHz spectrometer, in both cases with transverse relaxation optimized spectroscopy (TROSY). CSP experiments for FVO and 3D7 *Pf*AMA1 were carried out on a Bruker Avance III 600-MHz spectrometer at 35 and 30 °C, respectively. The buffer used for acquisition of NMR spectra was 20 mM phosphate buffer, pH 6.8 (FVO) and pH 7.0 (3D7), containing 50 mM L-arginine, 50 mM L-glutamic acid, 0.2% (w/v) protease inhibitors cocktail (Roche, Mannheim, Germany), 0.01% (w/v) sodium azide and 10% (v/v) ^2H_2O. Buffer at pH 7.0 was used for 3D7 *Pf*AMA1 instead of pH 6.8 as this protein precipitates at pH 6.8. Two-dimensional [^{1}H-^{15}N]-TROSY spectra were acquired with 32 scans at 2048 and 256 datapoints for the ^{1}H and ^{15}N dimensions, respectively. For FVO *Pf*AMA1, triple-resonance experiments including HNCA and HN(CO)CA spectra were acquired with a sample concentration of 150 μM and 32 scans, and HN(CA)CB and HN(COCA)CB spectra were acquired with a sample concentration of 250 μM and 64 scans. All 3D NMR experiments were recorded with 2048, 46 and 80 datapoints for ^{1}H, ^{15}N and ^{13}C dimensions, respectively. For 3D7 *Pf*AMA1, HNCA and HN(CO)CA spectra were acquired with 16 and 32 scans, respectively, and HN(CA)CB and HN(COCA)CB spectra were recorded with 24 scans. All 3D NMR experiments for 3D7 *Pf*AMA1 were recorded with a sample concentration of 250 μM and using non-uniform sampling conditions with 2048, 64, 120 datapoints for the ^{1}H, ^{15}N and ^{13}C dimensions, respectively. Data for FVO *Pf*AMA1 were processed using Bruker Topspin 3.2, and non-uniform sampling data for 3D7 *Pf*AMA1 were reconstructed with compressed sensing using qMDD and analysed using Sparky (v3.114; University of California, San Francisco).

The stability of FVO *Pf*AMA1 at 35 °C was monitored using 2D [^{1}H-^{15}N]-TROSY experiments at days 0, 3 and 6. Similarly, the stability of 3D7 *Pf*AMA1 at 30 °C was monitored at days 0, 1 and 2. Two-dimensional [^{1}H-^{15}N]-TROSY spectra were acquired with eight scans and 1024 and 64 datapoints for the ^{1}H and ^{15}N dimensions, respectively, and the spectra were processed as described earlier.

Phase-modulated CLEAN chemical exchange (CLEANEX) TROSY (Hwang *et al.*, 1998) experiments were recorded for 3D7 and FVO *Pf*AMA1 at 35 °C on a Bruker Avance 600-MHz spectrometer to identify amide protons in rapid chemical exchange with water. The mixing time for the CLEANEX spinlock was set to 45 ms. Spectra were acquired with 128 scans and 2048 and 256 datapoints for the ^{1}H and ^{15}N dimensions, respectively. The spectra were recorded on ^{2}H-labelled and ^{15}N-labelled 3D7 *Pf*AMA1 and FVO *Pf*AMA1 samples in the absence and presence of *Pf*RON2 peptide in 20 mM sodium phosphate buffer, pH 7.4 at sample concentrations of 100 and 70 μM, respectively.

Chemical shift perturbations

Two-dimensional [^{1}H-^{15}N]-TROSY spectra of ^{15}N-labelled and ^{2}H-labelled *Pf*AMA1 in the presence and absence of small molecules were recorded with 128 scans and 256 increments. The concentrations of *Pf*AMA1 and small molecule were 25 μM and 2 mM, respectively. The buffer used in these experiments contained 20 mM phosphate buffer, pH 6.8 (FVO) and pH 7.0 (3D7), with 50 mM L-arginine, 50 mM L-glutamic acid and 10% (v/v) ^2H_2O. Weighting of chemical shifts from ^{1}H and ^{15}N resonances was performed using the following equation (Williamson, 2013):

$$\text{Weighted CSP } (\Delta\delta) = \sqrt{\frac{1}{2}\left[\delta_H{}^2 + (0.14 \cdot \delta_N)^2\right]} \qquad (2)$$

where δ_H and δ_N are the CSP (ppm) at ^{1}H and ^{15}N dimensions, respectively. Small molecules were considered to be binding to a specific site on *Pf*AMA1 based on two criteria: (1) weighted CSP >0.01 ppm and (2) more than one peak perturbed in the spectrum.

RESULTS

Backbone resonance assignments

Both the FVO and 3D7 *Pf*AMA1 (DI + II) constructs consisted of 335 residues, including 19 proline residues, with molecular mass of around 38 kDa. Biomolecules of this size have fast T_2 relaxation rate with resulting signal broadening. To overcome this issue, a combination of protein deuteration (Gardner & Kay, 1998; Sattler & Fesik, 1996) and the TROSY scheme (Riek *et al.*, 2000; Salzmann *et al.*, 1998) in pulse sequences was used to improve the sensitivity of all heteronuclear NMR experiments (Figure S1). In our previous studies (Lim *et al.*, 2014), we observed 261 (~83%) and 250 (~79%) of the 316 amide backbone peaks expected for FVO (Figure 2) and 3D7 (Figure 3) *Pf*AMA1, respectively. Overall, the 2D [^{1}H-^{15}N]-TROSY spectra of both proteins showed well-dispersed amide backbone peaks that are consistent with a single folded conformer.

To determine if the protein samples (FVO and 3D7 *Pf*AMA1) were suitable for use in 3D NMR experiments with long acquisition times (HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB), their stability was assessed using 2D [^{1}H-^{15}N]-TROSY experiments. No significant changes of the resonances were observed in the FVO *Pf*AMA1 2D [^{1}H-^{15}N]-TROSY spectrum over a period of 6 days at 35 °C, indicating that the FVO *Pf*AMA1 sample possesses good long-term stability (Figure S2). In contrast, it was found that a few new resonances appeared over time in the 3D7 *Pf*AMA1 spectrum (Figure S2), which could be a consequence of partial unfolding of the protein with time or limited proteolysis (despite the presence of protease inhibitors).

In the FVO *Pf*AMA1 3D NMR spectra, totals of 246 $C\alpha_i$ and 244 $C\alpha_{i-1}$ resonances were observed in the HNCA and HN(CO)CA, respectively. Owing to the lower sensitivity, fewer signals were detected in the $C\beta$-detected experiments, with 205 $C\beta_i$ and 202 $C\beta_{i-1}$ signals observed in the HN(CA)CB and HN(COCA)CB, respectively. An example of an assignment strip plot from the 3D HNCA, HN(CO)CA, HN(CA)CB and HN(COCA)CB spectra recorded on FVO *Pf*AMA1 is shown in Figure S3. Two-dimensional [^{1}H-^{15}N]-TROSY spectra recorded on the ^{15}N-Lys selectively labelled *Pf*AMA1 were also used to minimize ambiguity in the amide backbone resonance assignments (Figure 4). Lysine was selected because it is one of the most abundant residues in *Pf*AMA1 that can be selectively labelled without the problem of scrambling to other amino acids during expression in *E. coli* (Muchmore *et al.*, 1989; Goto & Kay, 2000; Ohki & Kainosho, 2008; Krishnarjuna *et al.*, 2011; Jaipuria *et al.*, 2012).

In addition to manual peak assignments, the assignments were also confirmed by the Resonance Assignment by chemical Shift Prediction algorithm (MacRaild & Norton, 2014). A total of

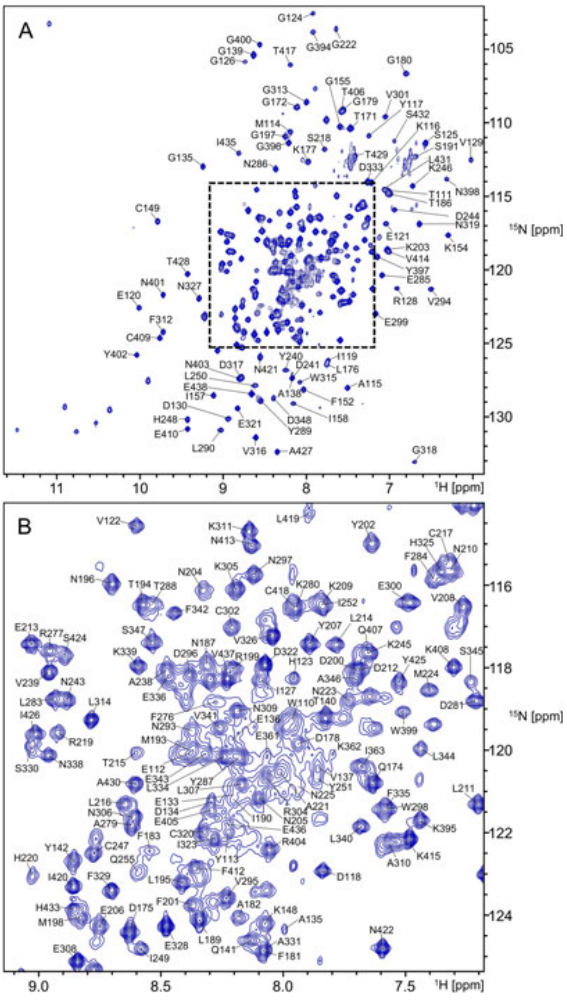


Figure 2. Two-dimensional [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectrum of FVO *Plasmodium falciparum* apical membrane antigen 1 (DI + DII) with resonance assignments shown. (A) Full spectrum. (B) Enlarged view of the middle region (dotted box in A). The experiments were performed at 35 °C, pH 6.8, on a Bruker Avance III 600-MHz spectrometer.

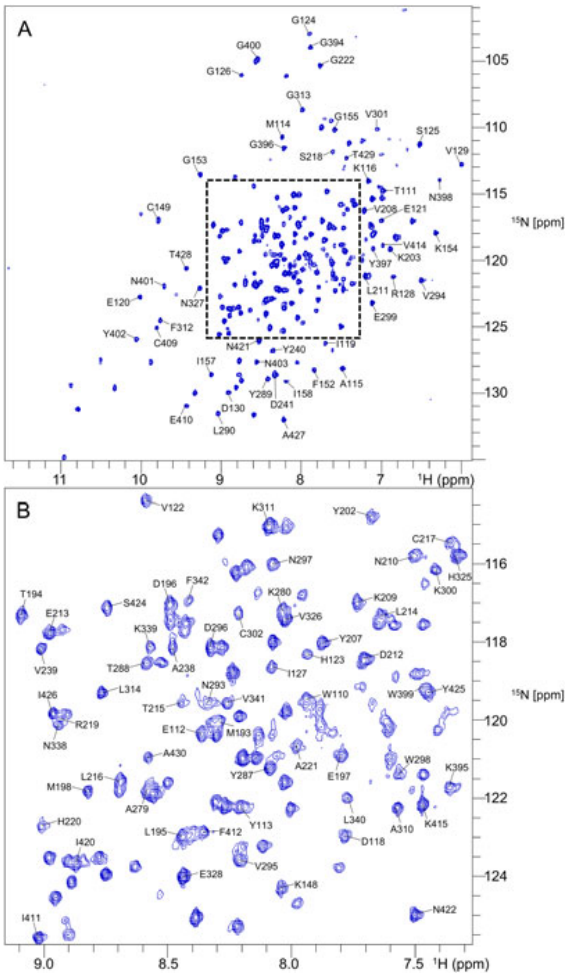


Figure 3. Two-dimensional [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectrum of 3D7 *Plasmodium falciparum* apical membrane antigen 1 (DI + DII) with resonance assignments shown. (A) Full spectrum. (B) Enlarged view of the middle region (dotted box in A). The experiments were performed at 30 °C, pH 7.0, on a Bruker Avance III 800-MHz spectrometer.

220 peaks in 2D [^{1}H-^{15}N]-TROSY could be assigned for FVO *Pf*AMA1 (BMRB ID 26648), which corresponds to 84% of the observed peaks (Figures 2, 5A and 5B). Notably, eight of the 12 residues in the hydrophobic cleft and 31 of the 40 polymorphic residues, including one of the most important residues that mediate immune escape against antibodies, Gly197, were assigned (Takala *et al.*, 2009; Ouattara *et al.*, 2013). Most of the unassigned resonances (except Arg143 and Val151 in structured regions) correspond to residues in the loop regions (104–108, 131–132, 144, 146–147, 159–167, 169, 227–236, 253–254, 257–259, 261–275, 291–292, 332, 349, 351–360, 365–380 and 382–393) of FVO *Pf*AMA1 (Lim *et al.*, 2014). These unassigned regions include the entire If loop (residues 264–273), most of the Ie loop (residues 224–238) and the DII loop (residues 351–387) that surround the *Pf*AMA1 hydrophobic cleft (Figure 5B). Assignments could not be obtained in these regions because of weak or missing resonances in the NMR spectra that were probably a result of peak broadening caused by intermediate conformational exchange or the relatively high amide proton exchange in solution at pH 6.8. Previous studies (Lim *et al.*, 2014) found that the 3D7 *Pf*AMA1 spectrum had a greater number of missing resonances than did the FVO *Pf*AMA1 spectrum owing to broadening by intermediate timescale conformational exchange and amide proton exchange at pH 7 (Figures 2 and 3).

To further probe the number of amide protons in rapid exchange with solvent, CLEANEX experiments (Hwang *et al.*, 1998) were performed on both 3D7 and FVO *Pf*AMA1. More peaks were observed in the CLEANEX spectrum of 3D7 *Pf*AMA1 than in that of FVO *Pf*AMA1 (Figure S4A). In the presence of a peptide corresponding to the apical membrane antigen 1 binding site on *Pf*RON2 (Tonkin *et al.*, 2011), we observed an increase in the number of resonances in the CLEANEX spectra of both 3D7 and FVO *Pf*AMA1 (Figure S4B and S4C). Most of these newly observed resonances are highly overlapped in the centre of the spectrum, suggesting that they may be from amides in the DII loop,

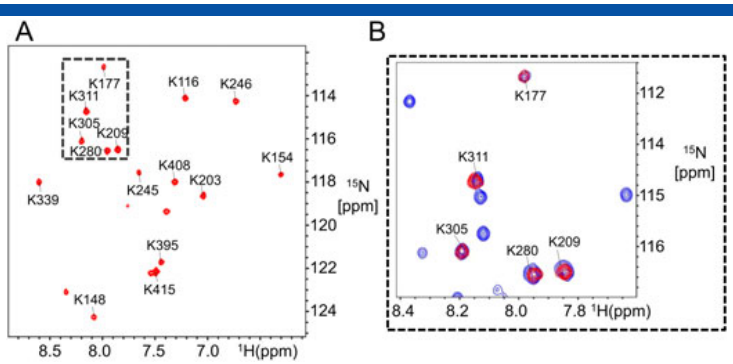


Figure 4. Lysine selective labelling of FVO *Plasmodium falciparum* apical membrane antigen 1 (*Pf*AMA1). (A) 2D [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectrum of ^{15}N-Lys selectively labelled FVO PfAMA1. (B) Overlay of the 2D [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectra of ^{15}N-Lys selectively labelled (red) and uniformly ^{15}N-labelled (blue) FVO *Pf*AMA1 samples.

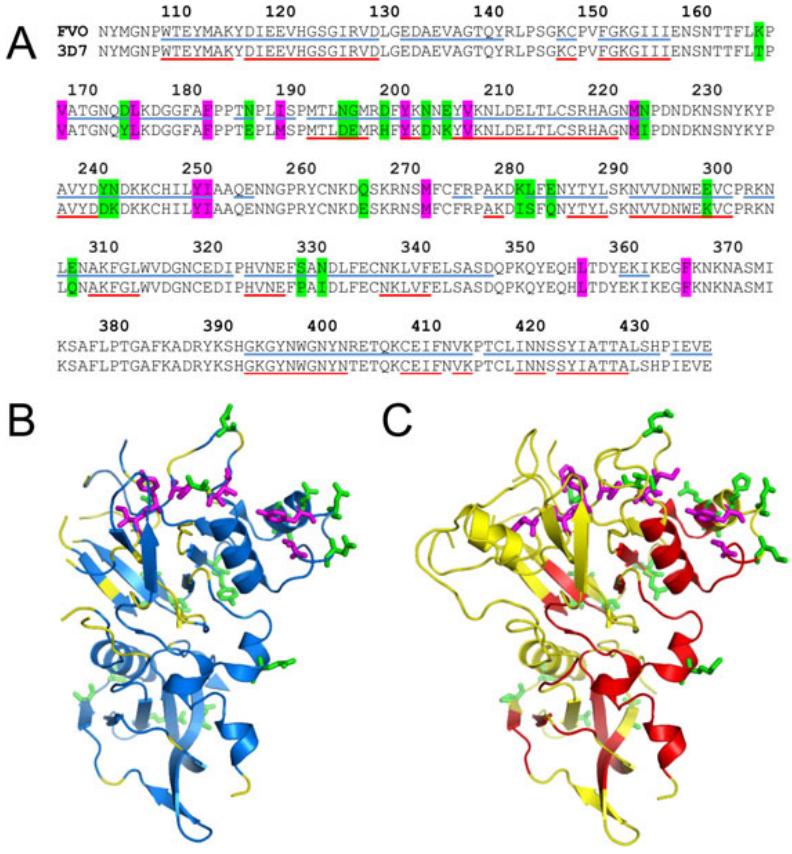


Figure 5. Comparison of nuclear magnetic resonance assignments of FVO *Plasmodium falciparum* apical membrane antigen 1 (*Pf*AMA1) (DI + DII) and 3D7 *Pf*AMA1 (DI + DII). (A) Amino acid sequences of FVO and 3D7 *Pf*AMA1. Residues that line the hydrophobic cleft are coloured purple. Residues that are different between the two forms of *Pf*AMA1 are highlighted in green. Assigned residues on FVO and 3D7 *Pf*AMA1 are underlined in blue and red, respectively. The crystal structures of (B) FVO *Pf*AMA1 (Lim *et al.*, 2014) (PDB ID 4R1A) and (C) 3D7 *Pf*AMA1 (Bai *et al.*, 2005) (PDB ID 1Z40). The residues that line the cleft are shown as purple sticks. Residues that are different between the two forms of *Pf*AMA1 are shown as green sticks. Regions that are not assigned are coloured yellow in both structures. Assigned regions are coloured blue and red for FVO and 3D7 *Pf*AMA1, respectively.

which is displaced by *Pf*RON2 binding to the hydrophobic cleft to become more flexible and solvent exposed (Ge *et al.*, 2014). Some peaks observed in CLEANEX are difficult to observe in 2D [^{1}H-^{15}N]-TROSY and other H^{N}-detected triple-resonance NMR experiments

because of their rapid exchange with solvent. These undesirable properties in 3D7 *Pf*AMA1, together with signal overlap and poor stability in solution, resulted in poor-quality triple-resonance NMR spectra, which presented a major hurdle in assigning the amide backbone resonances. To partially overcome this problem, multiple selectively unlabelled 3D7 *Pf*AMA1 (Lys, Arg, Asn and Ile, Leu, Val) samples were prepared, and 2D [^{1}H-^{15}N]-TROSY spectra were recorded for each sample (Figure S5) (Krishnarjuna *et al.*, 2011). The unlabelling approach was employed in this case because of the cost of preparing multiple specifically labelled samples.

Residue-specific assignments were obtained by overlaying the spectra of selectively unlabelled 3D7 *Pf*AMA1 samples on that of uniformly ^{15}N-labelled 3D7 *Pf*AMA1. The missing resonances in the spectra of selectively unlabelled samples were assigned to the corresponding amino acid (Figure S5) that was added to M9 medium in an unlabelled (^{14}N) form. These residue-specific assignments were used as starting points to obtain sequential assignments and to confirm resonance assignments from triple-resonance experiments (Figure S6 and Table S1). For 3D7 *Pf*AMA1, totals of 212 $C\alpha_i$ and 207 $C\alpha_{i-1}$ signals were observed in HNCA and HN(CO)CA spectra, respectively, and 190 $C\beta_i$ and 150 $C\beta_{i-1}$ signals, respectively, in HN(CA)CB and HN(COCA)CB spectra. Although many resonances were not observed in the 3D7 *Pf*AMA1 spectra, 114 (~45% of the observed peaks in 2D [^{1}H-^{15}N]-TROSY), including two from residues in the cleft, were assigned (Figures 3, 5A and 5C). The assigned chemical shifts (HN, N, C_α and C_β) for 3D7 *Pf*AMA1 are summarized in Table S1.

Chemical shift perturbations

Selected small molecule classes that displayed some structure–activity relationship in the initial fragment screening campaign (Lim *et al.*, 2013) were chosen for further evaluation in the CSP study. These scaffolds include benzimidazoles, pyrazoles and 2-aminothiazoles. These series bind weakly to 3D7 *Pf*AMA1 with K_D values ranging from 0.1 to 4.1 mM, as determined by surface plasmon resonance (SPR) (Figure 1 and Figure S7). The SPR sensorgrams indicate that all of the compounds bind to *Pf*AMA1 with very fast association and dissociation rates (Figure S7). Amongst the different small molecule classes tested, the benzimidazole series showed higher binding affinities than the pyrazoles and 2-aminothiazoles (Figure 1). Of the compounds tested in 2D [^{1}H-^{15}N]-TROSY experiments, all but MIPS-0000620 induced CSPs of more than one backbone amide resonance of FVO *Pf*AMA1 (Table S2). The number of peaks perturbed by MIPS-0000865, MIPS-0008405, MIPS-0001404, MIPS-0001160, MIPS-0001176 and MIPS-0008939 ranges from 5 to 19, as summarized in Table S2.

Resonances perturbed by the benzimidazoles MIPS-0000865, MIPS-0001404 and MIPS-0008405 were mapped onto both ends (DII loop and polymorphic end) of the *Pf*AMA1 hydrophobic cleft (Table S2). Larger numbers and magnitudes of CSPs were found at the DII loop end of the hydrophobic cleft for these compounds, suggesting that this region is the primary binding site (Figures 5B and 6). All resonances (except D175 for MIPS-0008405 and MIPS-0000865) perturbed by these compounds

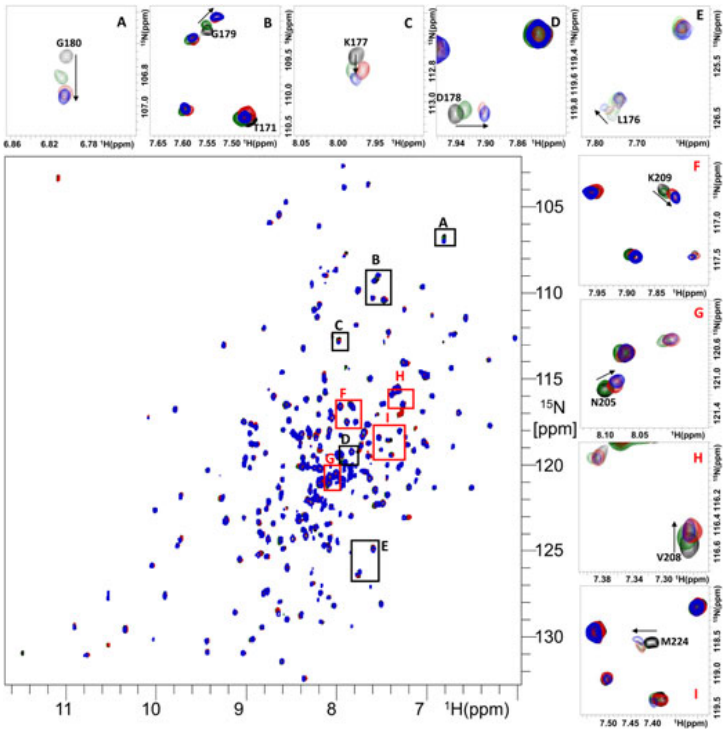


Figure 6. Two-dimensional [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy perturbations of FVO *Plasmodium falciparum* apical membrane antigen 1 (black), by MIPS-0001404 (green), MIPS-0000865 (red) and MIPS-0008405 (blue). Perturbed residues in A–E and F–I were mapped onto the conserved and polymorphic ends of the hydrophobic cleft, respectively. The black arrows indicate changes in the chemical shifts.

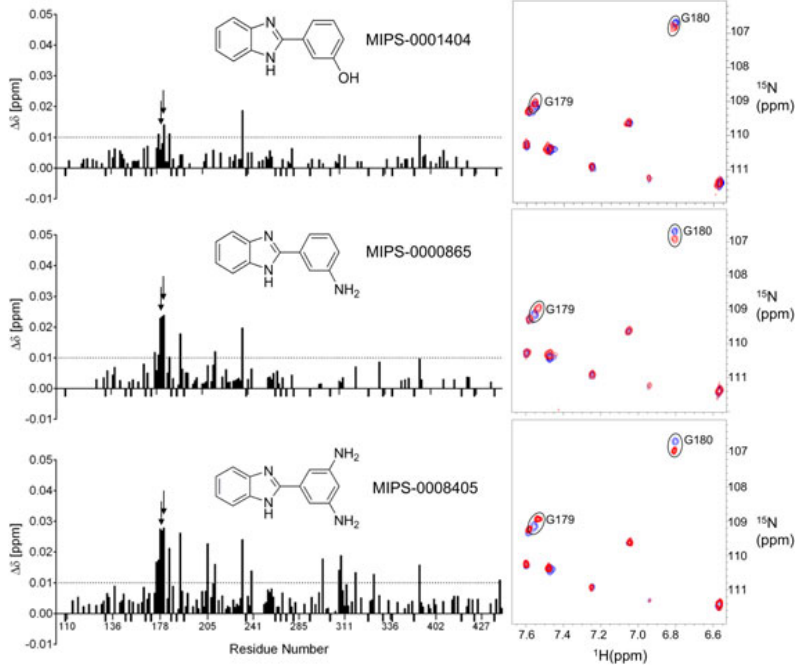


Figure 7. Chemical shift perturbations ($\Delta\delta$) of FVO *Plasmodium falciparum* apical membrane antigen 1 upon binding to benzimidazoles. Left: chemical shift perturbations per residue induced by the compounds shown. Gly179 and Gly180 are highlighted with arrows. Right: expansion of 2D [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectra highlighting perturbed resonances of Gly179 and Gly180. Blue and red spectra were recorded in the absence and presence of the benzimidazole shown in the left panel. The thick negative bars indicate the regions of unassigned residues (BMRB ID 26648).

correspond to conserved residues near the end of DII loop. Overall, the magnitude of the CSPs induced by benzimidazole compounds, particularly MIPS-0008405, was larger (Figures 6 and 7) compared with that of the CSPs induced by pyrazoles and 2-aminothiazoles (Table S2). MIPS-0008405 perturbed more resonances than other benzimidazoles (Figure 7). The perturbed residues are mapped onto the crystal structure of FVO *Pf*AMA1 (PDB ID 4R1A) in (Figure 8A and 8B). These results were also compared with the R1 and *Pf*RON2 peptide binding sites on 3D7 *Pf*AMA1 (Figure 8C and 8D). Small-molecule binding sites overlap partially with the binding sites of either R1 or *Pf*RON2. Only MIPS-0008405, which showed higher affinity for 3D7 *Pf*AMA1 in SPR binding experiments (Figure 1), was used in 2D [^{1}H-^{15}N]-TROSY experiment with 3D7 *Pf*AMA1; CSPs are compared with those for FVO *Pf*AMA1 in Figure 9. The similar perturbations observed in the spectra of FVO *Pf*AMA1 and 3D7 *Pf*AMA1 imply that FVO and 3D7 *Pf*AMA1 have very similar binding sites for this compound.

The pyrazoles MIPS-0001160 and MIPS-0001176 also showed binding to the DII loop end of the cleft. Apart from their interactions with polymorphic residue 172, these small molecules contact mostly conserved residues in this region on *Pf*AMA1 (Table S2). MIPS-0001160 also interacts with residues 419 and 420, which are located outside the hydrophobic cleft.

The 2-aminothiazole MIPS-0000620 induced CSPs at only one resonance, corresponding to Thr171 at the DII end of the cleft, which is also part of the binding site of R1 peptide. We have found that MIPS-0000620 competes with R1 (Lim *et al.*, 2013), and this observation, together with the 2D [^{1}H-^{15}N]-TROSY perturbation results, suggests a significant binding activity of MIPS-0000620 at the DII end of the hydrophobic cleft. Unlike MIPS-0000620, the other 2-aminothiazole, MIPS-0008939, showed extensive CSPs from multiple sites on *Pf*AMA1 that include both the DII loop and polymorphic ends of the cleft (Table S2 of the Supporting Information). This observation is in agreement with our previous study, which showed that 2-aminothiazoles can be problematic promiscuous binders (Devine *et al.*, 2015).

DISCUSSION

Structural information on ligand binding may be obtained using either X-ray crystallography or NMR spectroscopy. The former is usually the method of choice as it provides detailed structural information about ligand–protein interactions that is essential for designing potent inhibitors (Arkin & Wells, 2004; Congreve *et al.*, 2008). However, the study of small molecule-protein crystals in a static environment may not always be fully informative given that the *in vivo* interactions occur in a dynamic solution environment (Ishima & Torchia, 2000; Davis *et al.*, 2003; Davis *et al.*, 2008). Moreover, in the case of apical membrane antigen 1, there is evidence of conformational flexibility (Ge *et al.*, 2014; Lim *et al.*, 2014) and ligand-induced conformational changes (Tonkin *et al.*, 2011; Vulliez-Le Normand *et al.*, 2012; Ge *et al.*, 2014). Mapping CSPs in 2D [^{1}H-^{15}N]-TROSY spectra of the target protein represents a rapid and informative means of defining ligand binding sites in solution (Shuker *et al.*, 1996; Williamson, 2013) and can serve as a valuable complement to available crystallographic data. In situations where an accurate small molecule

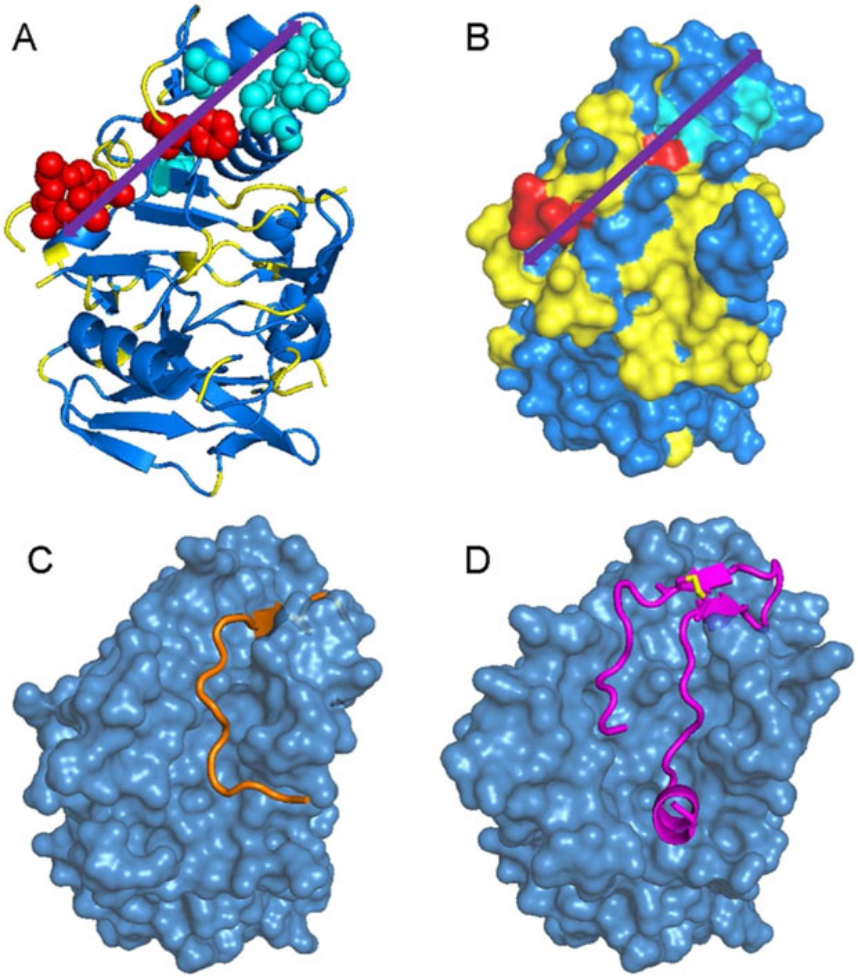


Figure 8. Mapping the chemical shift perturbations of benzimidazoles (MIPS-0001404, MIPS-0000865 and MIPS-0008405) onto the FVO *Pf*AMA1 crystal structure (Lim *et al.*, 2014) (PDB ID 4R1A). (A) Cartoon view and (B) surface view. The hydrophobic cleft is indicated by purple lines. The blue and yellow regions correspond to assigned and unassigned resonances, respectively. Resonances from residues at the conserved and polymorphic ends of the cleft that were perturbed by these compounds in 2D [^{1}H-^{15}N]-transverse relaxation optimized spectroscopy spectra are coloured red and cyan, respectively. Three compounds perturb resonances from residues at two sites in the cleft, with primary and secondary binding sites towards the conserved and polymorphic ends of the cleft, respectively. (C) Crystal structure of 3D7 *Pf*AMA1 (sky blue) in the presence of R1 peptide (orange) (PDB ID 3SRJ) (Vulliez-Le Normand *et al.*, 2012) and (D) crystal structure of 3D7 *Pf*AMA1 (sky blue) in the presence of *P. falciparum* rhoptry neck protein 2 (magenta) with a disulfide bond highlighted in yellow (PDB ID 3ZWZ). All structures were aligned using the PyMOL (version 1.2r3pre Schrodinger, LLC. Coordinates) alignment function.

binding pose is difficult to obtain, such as low-resolution crystal structures or missing fragment electron densities, mapping the binding sites of fragment hits using CSPs will provide valuable guidance to subsequent medicinal chemistry elaboration.

Two-dimensional [^{1}H-^{15}N]-TROSY perturbation experiments can be conducted with relatively small amounts of *Pf*AMA1 (~0.3 mg). The FVO and 3D7 *Pf*AMA1 backbone amide resonances in 2D [^{1}H-^{15}N]-TROSY spectra were assigned using triple-resonance NMR experiments, complemented by both selective labelling and selective unlabelling methods (Jaipuria *et al.*, 2012; Krishnarjuna *et al.*, 2011). Using all these approaches, we could assign ~84% and ~45%, respectively, of the peaks in 2D [^{1}H-^{15}N]-TROSY spectra of FVO and 3D7 *Pf*AMA1. The assignments for FVO *Pf*AMA1 include most of the conserved hydrophobic cleft residues, as well as the polymorphic residues surrounding the cleft.

In both 3D7 and FVO *Pf*AMA1, the number of resonances observable in CLEANEX spectra increased upon addition of *Pf*RON2 peptide (Vulliez-Le Normand *et al.*, 2012). We believe that these resonances arise from amides in the flexible DII loop, which is displaced by *Pf*RON2 binding to the hydrophobic cleft (Ge *et al.*, 2014; Lim *et al.*, 2014). Several factors restricted the number of backbone assignments for 3D7 *Pf*AMA1, including the instability of the protein (Figure S2), line broadening due to intermediate exchange (Lim *et al.*, 2014) and the greater extent of solvent exchange identified

Journal of Molecular Recognition

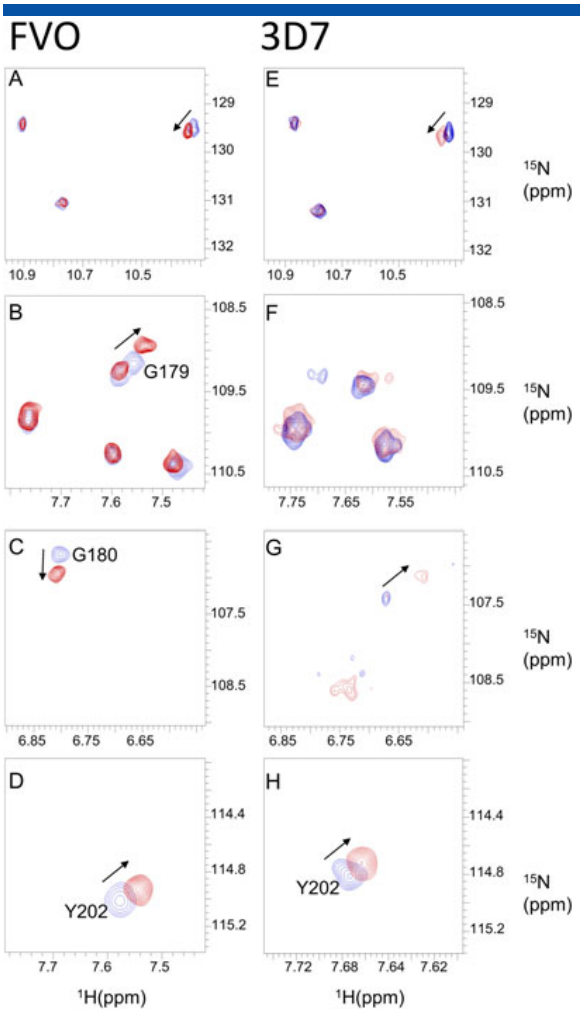


Figure 9. Comparison of the chemical shift perturbations in the spectra of FVO *Plasmodium falciparum* apical membrane antigen 1 (*Pf*AMA1) and 3D7 *Pf*AMA1 in the presence of MIPS-0008405. Left panel (A–D) represents the different regions of the FVO *Pf*AMA1 spectra, and right panel (E–H) represents the same regions for 3D7 *Pf*AMA1 spectra. Blue and red spectra are in the absence and presence, respectively, of MIPS-0008405. The black arrows indicate changes in the chemical shifts.

in CLEANEX spectra (Figure S4). For this reason, the binding sites were mapped on FVO *Pf*AMA1 rather than 3D7 *Pf*AMA1 for most of the small molecules tested here.

The DII end of the hydrophobic cleft has been proposed as a promising site for small-molecule intervention owing to its largely conserved nature (Bai *et al.*, 2005; MacRaild *et al.*, 2011; Lim *et al.*, 2014). The benzimidazole and pyrazole series showed binding to this region and therefore represent promising starting scaffolds for the design of *Pf*AMA1 lead compounds. Although the 2-aminothiazole series also showed binding to the DII end of the cleft, their promiscuous binding, as shown in previous (Devine *et al.*, 2015) and current studies, makes them poor starting points for subsequent fragment elaboration. MIPS-0008405 and MIPS-0000865 small molecules showed interaction with residue 175, which is one of the very few polymorphic residues found at this end of the cleft. This site can vary between Asp and Tyr in *Pf*AMA1 and is known to limit the cross reactivity of R1 (3D7 *Pf*AMA1-specific peptide inhibitor) against different strains of *Pf*AMA1, with the Y175D substitution reducing R1 peptide binding (Chesne-Seck *et al.*, 2005; Vulliez-Le Normand *et al.*, 2012). Asp is the predominant residue in the *Pf*AMA1 alleles, and as a result of this trait, potent inhibitory activity of R1 is limited to a small subset of *Plasmodium* strains. Two small molecules, MIPS-0008405 and MIPS-0000865, were identified initially from screening against 3D7 *Pf*AMA1, which has Tyr residue at position 175 (Lim *et al.*, 2013). We have shown that these compounds also bound to FVO *Pf*AMA1, indicating that they are able to bind to *Pf*AMA1 with either Asp or Tyr residues at position 175, perhaps through their interaction with the backbone. The benzimidazole series also perturbed resonances near the polymorphic end of the cleft. As opposed to the conserved end of the cleft, this binding site represents a less attractive site for fragment development (MacRaild *et al.*, 2011) as the polymorphic residues would limit the possible chemical space that can be exploited for the design of potent inhibitors with broad strain specificity. It is also possible that the observed CSPs in this region are a result of allosteric effects or non-specific interactions upon small molecule binding to the conserved region at the end of the cleft. Future analogue synthesis would aim to introduce functional groups that favour binding to the DII end of the cleft and disfavour binding to the polymorphic end of the cleft.

Both of the pyrazole compounds contact a dimorphic residue at position 172 in the DII region that can vary between Gly and Glu, with these two residues occurring at approximately the same frequency (Gly = 46% and Glu = 54%). Both 3D7 and FVO *Pf*AMA1 have Gly at this position. It is unclear at this stage if MIPS-0001160 and MIPS-0001176 could still bind to *Pf*AMA1 when Gly is substituted by Glu; this would require further investigation with either *Pf*AMA1 mutation experiments or binding studies of *Pf*AMA1 alleles carrying Glu172.

CONCLUSION

In conclusion, we have successfully assigned 84% of the expected backbone resonances for FVO *Pf*AMA1 and used these for mapping the binding sites for several small molecules. This information will aid in elaborating these small molecules into high-affinity ligands that bind to the conserved regions of *Pf*AMA1. Among the compounds tested, MIPS-0008405 showed stronger affinity by SPR and the largest perturbations in the *Pf*AMA1 spectrum. This suggests that benzimidazoles would be suitable as a basis for designing new *Pf*AMA1 inhibitors targeting the conserved regions of the hydrophobic cleft.

Acknowledgements

We thank Robin Anders for valuable discussions. This work was supported in part by an Australian National Health and Medical Research Council (NHMRC) project grant (1025150). R. S. N. acknowledges fellowship support from the NHMRC. The facilities provided by NMR Research Centre at IISc, supported by the Department of Science and Technology (DST), India, are gratefully acknowledged.

REFERENCES

Aikawa M, Miller L, Johnson J, Rabbege J. 1978. Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J. Cell Biol.* **77**: 72–82.

Arkin MR, Wells JA. 2004. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3**: 301–317.

Bai T, Becker M, Gupta A, Strike P, Murphy VJ, Anders RF, Batchelor AH. 2005. Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U. S. A.* **102**: 12736–12741.

Breman JG. 2012. Resistance to artemisinin-based combination therapy. *Lancet Infect. Dis.* **12**: 820–822.

Chesne-Seck M-L, Pizarro JC, Normand BV-L, Collins CR, Blackman MJ, Faber BW, Remarque EJ, Kocken CHM, Thomas AW, Bentley GA. 2005. Structural comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites. *Mol. Biochem. Parasitol.* **144**: 55–67.

Coley AM, Gupta A, Murphy VJ, Bai T, Kim H, Anders RF, Foley M, Batchelor AH. 2007. Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody. *PLoS Pathog.* **3**: e138.

Coley AM, Parisi K, Masciantonio R, Hoeck J, Casey JL, Murphy VJ, Harris KS, Batchelor AH, Anders RF, Foley M. 2006. The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infect. Immun.* **74**: 2628–2636.

Collins CR, Withers-Martinez C, Bentley GA, Batchelor AH, Thomas AW, Blackman MJ. 2007. Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J. Biol. Chem.* **282**: 7431–7441.

Congreve M, Chessari G, Tisi D, Woodhead AJ. 2008. Recent developments in fragment-based drug discovery. *J. Med. Chem.* **51**: 3661–3680.

Davis AM, St-Gallay SA, Kleywegt GJ. 2008. Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discov. Today* **13**: 831–841.

Davis AM, Teague SJ, Kleywegt GJ. 2003. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed.* **42**: 2718–2736.

Devine SM, Mulcair MD, Debono CO, Leung EWW, Nissink JWM, Lim SS, Chandrashekaran IR, Vazirani M, Mohanty B, Simpson JS, Baell JB, Scammells PJ, Norton RS, Scanlon MJ. 2015. Promiscuous 2-aminothiazoles (PrATs): a frequent hitting scaffold. *J. Med. Chem.* **58**: 1205–1214.

Gardner KH, Kay LE. 1998. The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**: 357–406.

Ge X, MacRaild CA, Devine SM, Debono CO, Wang G, Scammells PJ, Scanlon MJ, Anders RF, Foley M, Norton RS. 2014. Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ^{19}F NMR. *J. Med. Chem.* **57**: 6419–6427.

Goto NK, Kay LE. 2000. New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.* **10**: 585–592.

Harris KS, Casey JL, Coley AM, Karas JA, Sabo JK, Tan YY, Dolezal O, Norton RS, Hughes AB, Scanlon D, Foley M. 2009. Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning. *J. Biol. Chem.* **284**: 9361–9371.

Harris KS, Casey JL, Coley AM, Masciantonio R, Sabo JK, Keizer DW, Lee EF, McMahon A, Norton RS, Anders RF, Foley M. 2005. Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun.* **73**: 6981–6989.

Hwang T-L, van Zijl PM, Mori S. 1998. Accurate quantitation of water-amide proton exchange rates using the phase-modulated CLEAN chemical exchange (CLEANEX-PM) approach with a fast-HSQC (FHSQC) detection scheme. *J. Biomol. NMR* **11**: 221–226.

Ishima R, Torchia DA. 2000. Protein dynamics from NMR. *Nat. Struct. Biol.* **7**: 740–743.

Jaipuria G, Krishnarjuna B, Mondal S, Dubey A, Atreya HS. 2012. Amino acid selective labeling and unlabeling for protein resonance assignments, Isotope labeling in biomolecular NMR, Atreya HS (ed). Springer Netherlands. pp 95–118.

Keizer DW, Miles LA, Li F, Nair M, Anders RF, Coley AM, Foley M, Norton RS. 2003. Structures of phage-display peptides that bind to the malarial surface protein, apical membrane antigen 1, and block erythrocyte invasion. *Biochemistry* **42**: 9915–9923.

Krishnarjuna B, Jaipuria G, Thakur A, D'Silva P, Atreya HS. 2011. Amino acid selective unlabeling for sequence specific resonance assignments in proteins. *J. Biomol. NMR* **49**: 39–51.

Lamarque M, Besteiro S, Papoin J, Roques M, Normand BV-L, Morlon-Guyot J, Dubremetz J-F, Fauquenoy S, Tomavo S, Faber BW, Kocken CH, Thomas AW, Boulanger MJ, Bentley GA, Lebrun M. 2011. The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog.* **7**: e1001276.

Lee EF, Yao S, Sabo JK, Fairlie WD, Stevenson RA, Harris KS, Anders RF, Foley M, Norton RS. 2011. Peptide inhibitors of the malaria surface protein, apical membrane antigen 1: identification of key binding residues. *Biopolymers* **95**: 354–364.

Li F, Dluzewski A, Coley AM, Thomas A, Tilley L, Anders RF, Foley M. 2002. Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *J. Biol. Chem.* **277**: 50303–50310.

Lim SS, Debono CO, MacRaild CA, Chandrashekaran IR, Dolezal O, Anders RF, Simpson JS, Scanlon MJ, Devine SM, Scammells PJ, Norton RS. 2013. Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening. *Aust. J. Chem.* **66**: 1530–1536.

Lim SS, Yang W, Krishnarjuna B, Kannan Sivaraman K, Chandrashekaran IR, Kass I, MacRaild CA, Devine SM, Debono CO, Anders RF, Scanlon MJ, Scammells PJ, Norton RS, McGowan S. 2014. Structure and dynamics of apical membrane antigen 1 from *Plasmodium falciparum* FVO. *Biochemistry* **53**: 7310–7320.

MacRaild CA, Norton RS. 2014. RASP: rapid and robust backbone chemical shift assignments from protein structure. *J. Biomol. NMR* **58**: 155–163.

MacRaild CA, Anders RF, Foley M, Norton RS. 2011. Apical membrane antigen 1 as an anti-malarial drug target. *Curr. Top. Med. Chem.* **11**: 2039–2047.

Muchmore DC, McIntosh LP, Russell CB, Anderson DE, Dahlquist FW. 1989. Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance, Methods enzymol, Norman JO, Thomas LJ (eds). Academic Press: London. pp 44–73.

Ohki S-y, Kainosho M. 2008. Stable isotope labeling methods for protein NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* **53**: 208–226.

Ouattara A, Takala-Harrison S, Thera MA, Coulibaly D, Niangaly A, Saye R, Tolo Y, Dutta S, Heppner DG, Soisson L, Diggs CL, Vekemans J, Cohen J, Blackwelder WC, Dube T, Laurens MB, Doumbo OK, Plowe CV. 2013. Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: vaccine development implications. *J. Infect. Dis.* **207**: 511–519.

Richard D, MacRaild CA, Riglar DT, Chan J-A, Foley M, Baum J, Ralph SA, Norton RS, Cowman AF. 2010. Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites. *J. Biol. Chem.* **285**: 14815–14822.

Riek R, Pervushin K, Wüthrich K. 2000. TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends Biochem. Sci.* **25**: 462–468.

Salzmann M, Pervushin K, Wider G, Senn H, Wüthrich K. 1998. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 13585–13590.

Sattler M, Fesik SW. 1996. Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure* **4**: 1245–1249.

Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**: 1531–1534.

Sivashanmugam A, Murray V, Cui CX, Zhang YH, Wang JJ, Li QQ. 2009. Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci.* **18**: 936–948.

Srinivasan P, Beatty WL, Diouf A, Herrera R, Ambroggio X, Moch JK, Tyler JS, Narum DL, Pierce SK, Boothroyd JC, Haynes JD, Miller LH. 2011. Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 13275–13280.

Takala SL, Coulibaly D, Thera MA, Batchelor AH, Cummings MP, Escalante AA, Ouattara A, Traoré K, Niangaly A, Djimdé AA, Doumbo OK, Plowe

CV. 2009. Extreme polymorphism in a vaccine antigen and risk of clinical malaria: implications for vaccine development. *Sci. Transl. Med.* **1**: 2ra5–2ra5.
Tonkin ML, Roques M, Lamarque MH, Pugnière M, Douguet D, Crawford J, Lebrun M, Boulanger MJ. 2011. Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide. *Science* **333**: 463–467.
Vulliez-Le Normand B, Tonkin ML, Lamarque MH, Langer S, Hoos S, Roques M, Saul FA, Faber BW, Bentley GA, Boulanger MJ, Lebrun M. 2012. Structural and functional insights into the malaria parasite moving junction complex. *PLoS Pathog.* **8**: e1002755.
Wang G, MacRaild CA, Mohanty B, Mobli M, Cowieson NP, Anders RF, Simpson JS, McGowan S, Norton RS, Scanlon MJ. 2014. Molecular insights into the interaction between *Plasmodium falciparum* apical membrane antigen 1 and an invasion-inhibitory peptide. *PLoS One* **9**: e109674.
White NJ. 2004. Antimalarial drug resistance. *J. Clin. Invest.* **113**: 1084–1092.
WHO. 2014. World malaria report. *WHO Library Cataloguing*. 1–242.
Williamson MP. 2013. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**: 1–16.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web site.

A.4 A critical evaluation of pyrrolo[2,3-*d*]pyrimidine-4-amines as *Plasmodium falciparum* apical membrane antigen 1 (AMA1) inhibitors

MedChemComm

CONCISE ARTICLE

View Article Online
View Journal

Cite this: DOI: 10.1039/c4md00090k

A critical evaluation of pyrrolo[2,3-*d*]pyrimidine-4-amines as *Plasmodium falciparum* apical membrane antigen 1 (AMA1) inhibitors†

Shane M. Devine,‡*[a] San Sui Lim,‡[a] Indu R. Chandrashekaran,[a] Christopher A. MacRaild,[a] Damien R. Drew,[b] Cael O. Debono,[a] Raymond Lam,[a] Robin F. Anders,[c] James G. Beeson,[b] Martin J. Scanlon,[ad] Peter J. Scammells[a] and Raymond S. Norton[a]

Received 28th February 2014
Accepted 23rd July 2014

DOI: 10.1039/c4md00090k

www.rsc.org/medchemcomm

We have determined that a previously reported class of pyrrolo[2,3-*d*]pyrimidine-4-amines exhibit low binding to apical membrane antigen 1 (AMA1) and suffer from unattractive qualities, such as aggregation. We attempted to remove these traits by generating molecules with improved solubility, but this did not translate into enhanced binding affinity or inhibition of parasite growth in erythrocytes. These results indicate that anti-malarial activity is not primarily due to inhibition of AMA1 function, but mediated by an alternate or additional mechanism of action.

Introduction

Malaria is a major health concern for many of the world's most vulnerable and impoverished societies, affecting peoples throughout sub-Saharan Africa and Southeast Asia, in particular. This infectious disease continues to threaten a large number of people, with more than 40% of the global population at risk of infection and the cause of 2% of human mortalities worldwide.[1] Malaria is caused by intracellular parasites of the genus *Plasmodium*, which are transmitted to humans when a female *Anopheles* mosquito takes a blood meal. All of the symptoms of malaria arise as a result of the parasite's asexual reproductive cycle, that occurs within erythrocytes of the human host.[2] For this reason existing treatments target this blood-stage infection.

Treatment strategies for malaria have changed markedly over recent decades, as the parasite has developed resistance to previously effective drugs. The current frontline approach is that of the artemisinin-combination therapies (ACTs).[1] However, recent evidence shows that resistance to artemisinin is emerging in Southeast Asia, particularly along the Thailand–Myanmar border.[3] Although there are other potential therapeutics in the pipeline, these are not sufficiently advanced to be therapeutically relevant at this stage. Moreover, the current portfolio of candidates in development, although improving, is still lacking in diversity.[4] Therefore, there is a clear and present need for new targets for drugs to combat malaria.[5]

Mature apical membrane antigen 1 (AMA1), a 66 kDa type 1 integral membrane protein with a short well-conserved cytoplasmic region, forms a complex with parasite rhoptry neck (RON) proteins as part of a moving junction that forms between the invading parasite and the host cell.[6,7] Crystal structures of the ectodomain of AMA1 from *Plasmodium falciparum* and related species reveal three domains, including two closely-packed PAN domains.[8,9] From these PAN domains, seven flexible loops extend to surround a large hydrophobic cleft consisting of 12 well-defined and conserved residues.[8] This cleft is the site of interaction between AMA1 and RON2,[6,10] and numerous inhibitory peptides and antibodies have also been shown to target this site.[11–13] Although recent studies have used genetic knockouts to challenge previous evidence that the AMA1–RON2 interaction is essential to host cell invasion by *P. falciparum* and related parasites in cell culture,[14,15] it remains clear that diverse inhibitors of this interaction do inhibit invasion.[6,11,12,16–19] Moreover, the strong conservation of the AMA1–RON2 interaction in the Apicomplexa phylum,[7,20] even in the face of strong selective pressure from host immune systems,[21,22] implies an important functional role *in vivo*.

On this basis, we and others have proposed AMA1 as a potential drug target against malaria.[13,23,24] Several additional factors contribute to its attractiveness in this regard. Firstly, there are no human homologues of AMA1 or the RON proteins. Secondly, inhibitors of the AMA1–RON2 interaction will have their site of action in the bloodstream, thereby avoiding the

[a]*Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia. E-mail: shane.devine@monash.edu*

[b]*Centre for Biomedical Research, Burnet Institute, Melbourne, Victoria 3004, Australia*

[c]*Department of Biochemistry, La Trobe University, Melbourne, Victoria 3086, Australia*

[d]*Centre of Excellence for Coherent X-Ray Science, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c4md00090k

‡ These authors contributed equally to this paper.

difficulties associated with targeting the intracellular stages of the malaria parasite. This point is particularly relevant in light of the role of drug transporters in mediating resistance to known anti-malarials.[25]

Recently, Srinivasan *et al.* reported the first example of small-molecule inhibitors of the AMA1–RON2 complex, identified *via* an AlphaScreen assay of a ~21 000 member library, utilising a truncated RON2 peptide.[26] This screen identified seven molecules, three of which blocked merozoite invasion *in vitro* with IC_{50} values in the range 21–29 μM (Fig. 1).

Re-synthesis of NCGC00015280 and chemical elaboration of this scaffold identified two related molecules NCGC00262650 and NCGC00262654 that showed enhanced inhibition with reported IC_{50} values of 9.8 μM and 6 μM, respectively, compared to the re-synthesised NCGC00015280's activity of 30 μM (Fig. 2).

Srinivasan *et al.* present an array of cell-based assays to support their proposed mode of action for these compounds. They also attempted to demonstrate a direct interaction with AMA1 by surface plasmon resonance (SPR).[26] However, these experiments were compromised by the poor solubility of the compounds, and the shape and concentration dependence of the observed sensorgrams are consistent with low-affinity super-stoichiometric interaction with the SPR biosensor surface, rather than the high-affinity stoichiometric interaction expected of a specific inhibitor. Two of the identified compounds have calculated partition co-efficients (clog *P*) that fall outside the traditional Lipinski 'rule of 5' upper limit of clog $P = 5$.[27] Another useful metric of inhibitor quality is the lipophilic ligand efficiency (LLE_{AT}), as described by Astex, which incorporates the number of heavy atoms as an indicator of lipophilicity.[28] By this metric as well, these molecules are far from the attractive value of ≥0.3 kcal mol^{-1} per heavy atom (Fig. 2). To assess the suitability of this class of compounds for further development as AMA1 inhibitors, we synthesised a panel of analogues and studied their solution behaviour and interactions with AMA1.

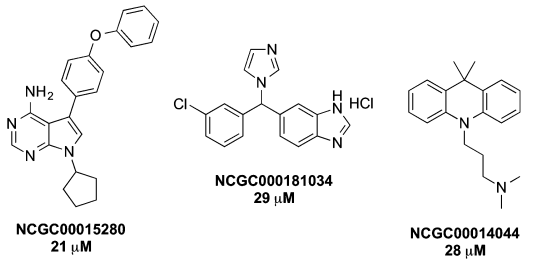


Fig. 1 Putative AMA1 inhibitors identified by Srinivasan *et al.*[26]

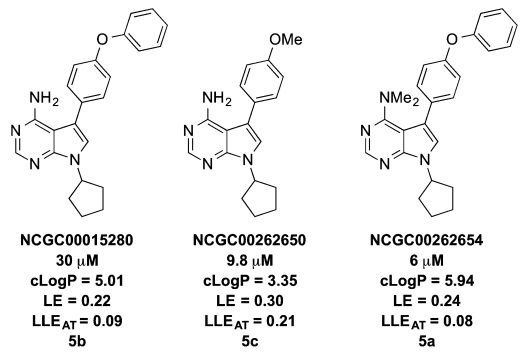


Fig. 2 Elaboration of NCGC00015280 by Srinivasan *et al.*[26]

Results

We endeavoured to evaluate the pyrrolo[2,3-*d*]pyrimidine scaffold by using our methods for monitoring AMA1 ligand binding[29] and to ameliorate the unfavourable traits, such as the high clog *P* and low LLE_{AT} values, whilst maintaining activity. Although compound lipophilicity is frequently an important factor for inhibitor potency, it can also contribute to off-target effects, leading to unwanted toxicity.[30] Small molecules that exhibit poor solubilities in aqueous buffers are often identified as promiscuous hits in a wide range of assays, complicating the interpretation of binding and/or activity data.[30–33] Understanding the physicochemical properties of a compound in an aqueous environment is therefore essential in evaluating chemical entities as lead candidates. To assess these issues for the class of compounds reported by Srinivasan *et al.*,[26] **5a–5c** were synthesised (see ESI† for details). They were then assessed at a range of concentrations (5, 10, 20, 40 and 80 μM) using 1D ^{1}H NMR spectroscopy, following the methods outlined by LaPlante *et al.*[31,32] This allowed direct observations of the aggregation propensities of these molecules in aqueous buffers.

The ^{1}H NMR spectra of **5a** and **5b** (Fig. 3A and B) show clear evidence of extensive aggregation. Both compounds gave very weak NMR signals that did not increase with concentration. As there were no visible precipitates in these samples, these observations suggest that the compounds were self-associating to form colloidal aggregates over the entire concentration range tested.[31] These large aggregates tumble more slowly than non-aggregating compounds and therefore exhibit faster NMR relaxation and signals that are broadened beyond detection.

To confirm this interpretation of these data, surfactant (Tween 20) was added into these samples. The NMR signals were consistent with those expected for these molecules, albeit with significant residual line broadening (top panel, Fig. 3A and B). These observations may be explained by the fact that surfactant is capable of dissociating large assemblies into smaller entities with better relaxation properties for NMR detection.[32] In contrast to **5a** and **5b**, **5c** gave sharp proton signals with peak intensities that increased with concentration over the range tested (Fig. 3C).

However, the concentration dependence of the peak intensities is not linear (Fig. 3D), while chemical shifts for a number of resonances showed a weak concentration dependence (Fig. 3E), indicating that compound **5c** also aggregates in aqueous solution, albeit to a much lesser extent than do **5a** and **5b**. The signals of compound **5c** were shifted and broadened in

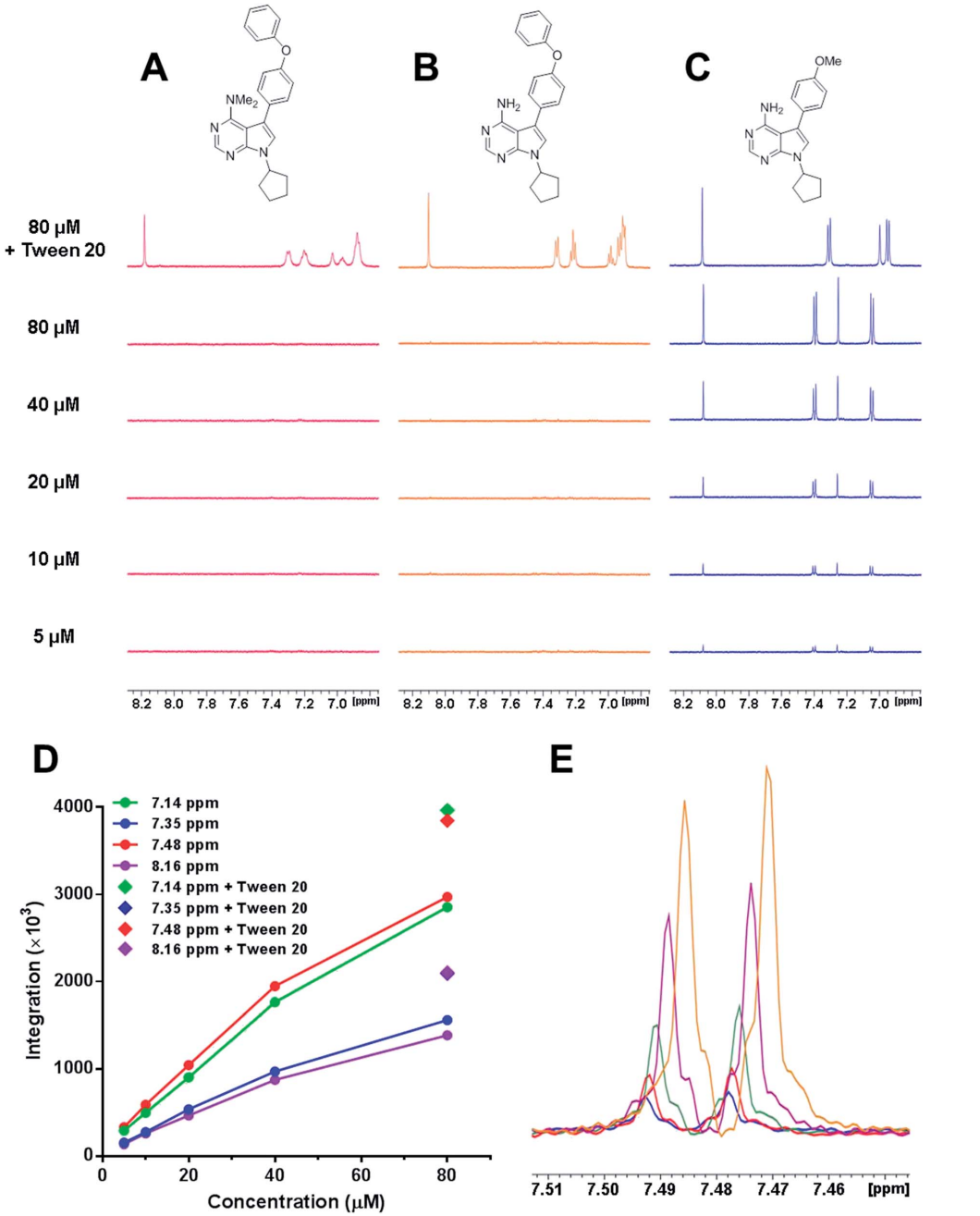


Fig. 3 1D ^{1}H NMR spectra of (A) **5a**, (B) **5b** and (C) **5c** at different concentrations and in the presence of 0.05% Tween 20. (D) Peak intensities of **5c** proton signals (7.14, 7.35, 7.48 and 8.16 ppm) at increasing compound concentrations and with the addition of surfactant. (E) Concentration dependence of the chemical shifts of **5c** at 5, 10, 20, 40 and 80 μM, respectively.

the presence of Tween 20, presumably as a result of the interactions between the compound and surfactant. Furthermore, peak intensities increased in the presence of the surfactant (Fig. 3D), again indicating the presence of large aggregates that are disrupted by the addition of Tween 20 into the sample.

We then assessed the interaction of these compounds with AMA1 using techniques developed in our own search for inhibitors of the AMA1–RON2 interaction.[29] The aggregation behaviour of **5a** and **5b**, described above, precluded the use of ligand-detected NMR strategies, so only **5c** was evaluated in the Carr–Purcell–Meiboom–Gill (CPMG) binding assay (Fig. 4). This method utilises the CPMG spin-lock filter to eliminate signals of the rapidly relaxing AMA1 protein and bound ligands, with the signals of free compounds less affected.[34] CPMG spectra for detergent-free samples containing 80 μM **5c** in the presence and absence of 10 μM 3D7 *Pf*AMA1 were first acquired with 0 and 200 ms spin-relaxation filters. Thereafter, R1 and RON2 peptides were added into the samples and the same set of CPMG spectra was acquired to evaluate the binding activities of **5c** at the AMA1 hydrophobic cleft. The signal intensities and chemical shifts of a number of peaks for **5c** were affected by the presence of AMA1 (Fig. 4A). These effects are exemplified using the proton resonance of **5c** at around 7.35 ppm in Fig. 4B. The signals were broadened and shifted slightly upfield in the presence of AMA1 and, when a 200 ms spin-lock filter was applied, the signal in the presence of AMA1 relaxed significantly faster than the signal in the absence of protein (R_2 relaxation rates of 8.2 and 1.4 s^{-1}, respectively). These results are indicative of **5c** binding to AMA1.

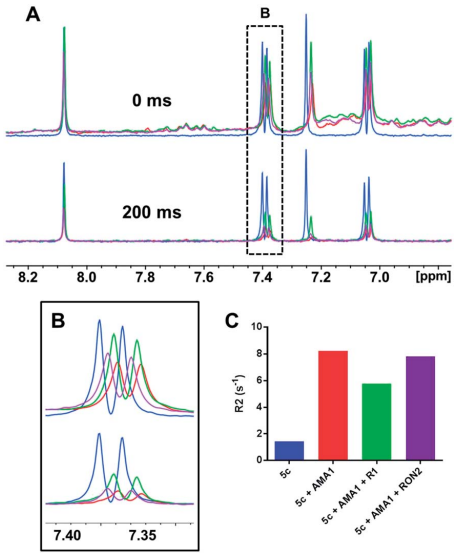
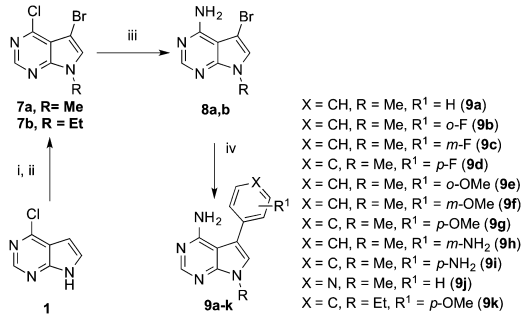


Fig. 4 CPMG binding assay. (A and B) All CPMG spectra were acquired with both 0 (top spectra) and 200 ms (bottom spectra) spin-relaxation filters in the absence of detergent. Blue and red spectra were results for samples containing **5c** in the presence or absence of 3D7 *Pf*AMA1, respectively. Green and purple spectra correspond to samples containing **5c** and 3D7 *Pf*AMA1 with the additions of R1 and RON2 peptides, respectively. (C) Transverse relaxation rate (R2) for different samples used in the binding studies.

When R1 and RON2 peptides were added to the samples containing both AMA1 and **5c**, the compound peaks shifted slightly downfield towards the chemical shift of the free compound. Also, partial restoration of the free compound line shape and relaxation rate was observed in both samples with the peptides added. These observations are consistent with competition between **5c** and R1 and RON2 peptides for binding sites on AMA1. The presence of the competing peptides reduced the fraction of bound **5c**, which in turn decreased the average transverse relaxation rate of the compound, and thus increased **5c** signals after the CPMG period (Fig. 4B and C). Similar results were observed when the same set of experiments was conducted in the presence of Tween 20 (Fig. S1†). This result confirms that **5c** in its monomeric form is able to bind AMA1.

Finally, we interrogated these compounds, including their synthetic precursors, by SPR, to estimate the binding affinities to AMA1 using methods described previously.[29] Not unexpectedly, the precursor molecule fragments **1–3**, showed minimal interaction with AMA1 at concentrations up to 200 μM. Once the 4-amino or 4-dimethylamino group was introduced, we saw evidence of super-stoichiometric binding, in the form of responses that exceeded the maximal response expected for these compounds and that failed to saturate, even at the highest concentrations studied. These issues were particularly acute in the case of **5a** and **5b**, and precluded any estimate of the affinity of these molecules with AMA1. For **5c**, we observed unambiguous over-binding only at higher concentrations. However, the response observed at lower concentrations is inconsistent with an affinity for AMA1 tighter than ~1 mM, while the IC_{50} for this compound reported by Srinivasan *et al.* is 9.8 μM.[26] This discrepancy suggests that this series of compounds exert their effects on host cell invasion by *P. falciparum* merozoites by some mechanism other than direct inhibition of AMA1, or by an additional mechanism of action.

To address the shortcomings of the three molecules, as outlined above, a series of pyrrolo[2,3-*d*]pyrimidine-4-amines was synthesised in an attempt to improve solubility whilst maintaining or improving affinity for AMA1 (Scheme 1). This was achieved by alkylation, bromination, amination and Suzuki coupling. We envisaged that replacement of the 7-cyclopentyl with a methyl group should aid in solubility, and we also explored substitutions on the 5-aryl group while retaining the 4-amino group. This scaffold featuring the *N*-methyl substitution is the core of a known protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK) inhibitor that has been selected as a preclinical candidate for tumour inhibition.[35,36] The compounds produced had clog *P* values of 0.93–2.35, which, in contrast to the clog *P* value of 3.35 for **5c**, provided soluble compounds for study. Aggregation propensities of **9a–k** were evaluated using 1D ^{1}H NMR spectroscopy (Fig. S2†) as described earlier for **5a–c**. The 7-methyl and 7-ethyl series exhibited substantially reduced aggregation behaviour, when compared to the 7-cyclopentyl series (**5a–c**), with or without the addition of Tween 20, demonstrating our success in eliminating some of the unwanted behaviour of **5a–c**.



Scheme 1 Reagents and conditions: (i) NaH, DMF, 0 °C, 30 min for **6a**, MeI and for **6b**, EtI, 60 °C, 4 h; (ii) NBS, DCM, 25 °C; (iii) NH_4OH, *i*-PrOH, 100 °C, 40 h; (iv) R^1-$B(OH)_2$, $PdCl_2(PPh_3)_2$, THF/1 M Na_2CO_3, 100 °C, 2 h.

All pyrrolo[2,3-*d*]pyrimidines (**5a–5c** and **9a–9k**) were analysed for binding to AMA1 using SPR, utilising similar conditions to those employed by Srinivasan *et al.*[26] In general, the sensorgrams showed weak binding affinities, with K_D values ≥1 mM. There were no visible precipitates in the samples and all compounds were soluble in the buffer conditions employed. From this small library of compounds, however, no meaningful structure–activity relationships (SAR) could be inferred and no increased affinity for AMA1 was observed, suggesting that these compounds are not potent inhibitors of AMA1 interactions. Finally, **9a–9k** were tested to ascertain their ability to inhibit parasite growth in erythrocytes of *P. falciparum* (Fig. S3†).[37] Compounds **5c** and **9a–k** showed substantial inhibitory activity in this assay, although this activity did not correlate with their affinity for AMA1 by SPR (Fig. S3†) . **5a** and **5b** demonstrated IC_{50} values > 1 mM, compared to the reported values of 6 and 30 μM, respectively.[26] The IC_{50} value for **5c** was 63 μM, in comparison to the reported value of 9.8 μM and the best analogue, **9c**, had an IC_{50} value < 31 μM. This represents still further evidence in support of an alternative mechanism of anti-malarial action by these compounds.

Discussion

Our results suggest that these compounds in their current form are not suitable for development as AMA1 inhibitors given their apparent low binding affinity to AMA1, sub-optimal potency, relative insolubility and tendency to form aggregates. Our findings suggest that these compounds may exhibit an alternative or additional mechanism of anti-malarial action. Compound **5b** was identified initially as a Src-family kinase inhibitor[38] and molecules containing the pyrrolo[2,3-*d*]pyrimidine scaffold have been shown to be involved in a number of therapeutically relevant areas such as Huntington's disease[39] and acute myeloid leukaemia,[40] where they are believed to act *via* a Src-family kinase mechanism. Srinivasan *et al.* demonstrated that a structurally different Src-inhibitor did not block invasion.[26] However, this does not preclude the possibility of the pyrrolo[2,3-*d*]pyrimidine inhibitors acting as kinase inhibitors in this context. A structurally similar framework, incorporating an extra nitrogen at the 6-position, namely the pyrazolo[3,4-*d*]-pyrimidin-4-amine scaffold with comparable N1- and 3-position modifications, has been implicated in a related apicomplexan species, *Toxoplasma gondii*, as a calcium-dependent protein kinase 1 (CDPK1) inhibitor.[41,42] CDPK is known to control microneme secretion and consequently block invasion in *T. gondii*, as well as *Plasmodium falciparum*.[43] Therefore, it is conceivable that these pyrrolo[2,3-*d*]pyrimidines could be acting *via* a CDPK mechanism, or in addition to AMA1-mediated inhibition. Colloidal aggregation of compounds is well known to contribute to promiscuity in high-throughput screens and in a range of other assays.[30–33,44] The observed aggregation properties of **5a** and **5b**, may be responsible for their inhibitory effects,[26] rather than being specific inhibitors of the AMA1–RON2 interaction.[44] Compound **5c** seems to be genuinely binding AMA1, but with insufficient affinity (using recombinant AMA1) to explain its reported activity. It remains possible that the binding affinity to native AMA1 is different from that measured using recombinant AMA1, although we note that these two preparations afford essentially identical results for genuine AMA1 inhibitors such as the peptide R1. Our attempts to produce more soluble analogues (**9a–k**) with increased activity did not generate any meaningful SAR.

Conclusion

Candidate pyrrolo[2,3-*d*]pyrimidine-4-amines (**5a–c**) appear to be moderately potent inhibitors of erythrocyte invasion by *P. falciparum* merozoites that prevent the formation of both the moving junction and the AMA1–RON complex.[26] These effects may not be mediated primarily by direct stoichiometric interaction with AMA1; inhibitory activity could include off-target mechanisms, which may be related to their tendency to form colloidal aggregates in aqueous solution, and/or their activity as kinase inhibitors. Molecules with lower clog *P* values were synthesised (**9a–k**), but failed to achieve greater inhibitory activity. Furthermore, this series of compounds carries significant physicochemical liabilities that are likely to impede their further development.

Acknowledgements

The authors would like to thank Brad Sleebs for fruitful discussions on this project, Samuel Hughes for preliminary work on the synthesis, and Jason Dang for obtaining HRMS data. This work was supported in part by an Australian National Health and Medical Research Council (NHMRC) project grant (1025150) and program grant (637406). R.S.N. and J.G.B acknowledge fellowship support from the NHMRC. The Burnet Institute is supported by the NHMRC Independent Research Institutes Infrastructure Support Scheme, and Operational Infrastructure Support from the State Government of Victoria, Australia.

References

1 World Health Organization World Malaria Report, 2013, pp. 1–284.
2 L. H. Miller, D. I. Baruch, K. Marsh and O. K. Doumbo, *Nature*, 2002, **415**, 673–679.
3 A. P. Phyo, S. Nkhoma, K. Stepniewska, E. A. Ashley, S. Nair, R. McGready, C. ler Moo, S. Al-Saai, A. M. Dondorp, K. M. Lwin, P. Singhasivanon, N. P. J. Day, N. J. White, T. J. C. Anderson and F. Nosten, *Lancet*, 2012, **379**, 1960–1966.
4 M. P. Anthony, J. N. Burrows, S. Duparc, J. J. Moehrle and T. N. C. Wells, *Malar. J.*, 2012, **11**, 316.
5 M. A. Biamonte, J. Wanner and K. G. Le Roch, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 2829–2843.
6 D. Richard, C. A. MacRaild, D. T. Riglar, J.-A. Chan, M. Foley, J. Baum, S. A. Ralph, R. S. Norton and A. F. Cowman, *J. Biol. Chem.*, 2010, **285**, 14815–14822.
7 M. Lamarque, S. Besteiro, J. Papoin, M. Roques, B. Vulliez-Le Normand, J. Morlon-Guyot, J.-F. Dubremetz, S. Fauquenoy, S. Tomavo, B. W. Faber, C. H. Kocken, A. W. Thomas, M. J. Boulanger, G. A. Bentley and M. Lebrun, *PLoS Pathog.*, 2011, **7**, e1001276.
8 T. Bai, M. Becker, A. Gupta, P. Strike, V. J. Murphy, R. F. Anders and A. H. Batchelor, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 12736–12741.
9 J. C. Pizarro, B. Vulliez-Le Normand, M.-L. Chesne-Seck, C. R. Collins, C. Withers-Martinez, F. Hackett, M. J. Blackman, B. W. Faber, E. J. Remarque, C. H. M. Kocken, A. W. Thomas and G. A. Bentley, *Science*, 2005, **308**, 408–411.
10 M. L. Tonkin, M. Roques, M. H. Lamarque, M. Pugnière, D. Douguet, J. Crawford, M. Lebrun and M. J. Boulanger, *Science*, 2011, **333**, 463–467.
11 S. Dutta, L. S. Dlugosz, D. R. Drew, X. Ge, D. Ababacar, Y. I. Rovira, J. K. Moch, M. Shi, C. A. Long, M. Foley, J. G. Beeson, R. F. Anders, K. Miura, J. D. Haynes and A. H. Batchelor, *PLoS Pathog.*, 2013, **9**, e1003840.
12 K. S. Harris, J. L. Casey, A. M. Coley, R. Masciantonio, J. K. Sabo, D. W. Keizer, E. F. Lee, A. McMahon, R. F. Anders and M. Foley, *Infect. Immun.*, 2005, **73**, 6981–6989.
13 B. Vulliez-Le Normand, M. L. Tonkin, M. H. Lamarque, S. Langer, S. Hoos, M. Roques, F. A. Saul, B. W. Faber, G. A. Bentley, M. J. Boulanger and M. Lebrun, *PLoS Pathog.*, 2012, **8**, e1002755.
14 D. Y. Bargieri, N. Andenmatten, V. Lagal, S. Thiberge, J. A. Whitelaw, I. Tardieux, M. Meissner and R. Ménard, *Nat. Commun.*, 2013, **4**, 2552.
15 D. Giovannini, S. Späth, C. Lacroix, A. Perazzi, D. Bargieri, V. Lagal, C. Lebugle, A. Combe, S. Thiberge, P. Baldacci, I. Tardieux and R. Ménard, *Cell Host Microbe*, 2011, **10**, 591–602.
16 M. J. Boyle, D. W. Wilson, J. S. Richards, D. T. Riglar, K. K. A. Tetteh, D. J. Conway, S. A. Ralph, J. Baum and J. G. Beeson, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 14378–14383.
17 K. S. Harris, J. L. Casey, A. M. Coley, J. A. Karas, J. K. Sabo, Y. Y. Tan, R. S. Norton, A. B. Hughes, D. Scanlon and M. Foley, *J. Biol. Chem.*, 2009, **284**, 9361–9371.
18 K. A. Henderson, V. A. Streltsov, A. M. Coley, O. Dolezal, P. J. Hudson, A. H. Batchelor, A. Gupta, T. Bai, V. J. Murphy, R. F. Anders, M. Foley and S. D. Nuttall, *Structure*, 2007, **15**, 1452–1466.
19 A. Yap, M. F. Azevedo, P. R. Gilson, G. E. Weiss, M. T. O'Neill, D. W. Wilson, B. S. Crabb and A. F. Cowman, *Cell. Microbiol.*, 2014, **16**, 642–656.
20 N. I. Proellocks, R. L. Coppel and K. L. Waller, *Trends Parasitol.*, 2010, **26**, 297–304.
21 A. Cortés, M. Mellombo, I. Mueller, A. Benet, J. C. Reeder and R. F. Anders, *Infect. Immun.*, 2003, **71**, 1416–1426.
22 S. D. Polley, W. Chokejindachai and D. J. Conway, *Genetics*, 2003, **165**, 555–561.
23 C. A. MacRaild, R. F. Anders, M. Foley and R. S. Norton, *Curr. Top. Med. Chem.*, 2011, **11**, 2039–2047.
24 G. Santos and N. V. Torres, *PLoS One*, 2013, **8**, e59968.
25 C. P. Sanchez, A. Dave, W. D. Stein and M. Lanzer, *Int. J. Parasitol.*, 2010, **40**, 1109–1118.
26 P. Srinivasan, A. Yasgar, D. K. Luci, W. L. Beatty, X. Hu, J. Andersen, D. L. Narum, J. K. Moch, H. Sun, J. D. Haynes, D. J. Maloney, A. Jadhav, A. Simeonov and L. H. Miller, *Nat. Commun.*, 2013, **4**, 2261.
27 C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv. Drug Delivery Rev.*, 1997, **23**, 3–25.
28 P. N. Mortenson and C. W. Murray, *J. Comput.-Aided Mol. Des.*, 2011, **25**, 663–667.
29 S. S. Lim, C. O. Debono, C. A. MacRaild, I. R. Chandrashekaran, O. Dolezal, R. F. Anders, S. M. Devine, P. J. Scammells and R. S. Norton, *Aust. J. Chem.*, 2013, **66**, 1530.
30 B. J. Davis and D. A. Erlanson, *Bioorg. Med. Chem. Lett.*, 2013, **23**, 2844–2852.
31 S. R. LaPlante, R. Carson, J. Gillard, N. Aubry, R. Coulombe, S. Bordeleau, P. Bonneau, M. Little, J. O'Meara and P. L. Beaulieu, *J. Med. Chem.*, 2013, **56**, 5142–5150.
32 S. R. LaPlante, N. Aubry, G. Bolger, P. Bonneau, R. Carson, R. Coulombe, C. Sturino and P. L. Beaulieu, *J. Med. Chem.*, 2013, **56**, 7073–7083.
33 A. Vom, S. Headey, G. Wang, B. Capuano, E. Yuriev, M. J. Scanlon and J. S. Simpson, *Aust. J. Chem.*, 2013, **66**, 1518.
34 P. J. Hajduk, G. Sheppard, D. G. Nettesheim, E. T. Olejniczak, S. B. Shuker, R. P. Meadows, D. H. Steinman, G. M. Carrera, P. A. Marcotte, J. Severin, K. Walter, H. Smith, E. Gubbins, R. Simmer, T. F. Holzman, D. W. Morgan, S. K. Davidsen, J. B. Summers and S. W. Fesik, *J. Am. Chem. Soc.*, 1997, **119**, 5818–5827.
35 J. M. Axten, J. R. Medina, Y. Feng, A. Shu, S. P. Romeril, S. W. Grant, W. H. H. Li, D. A. Heerding, E. Minthorn, T. Mencken, C. Atkins, Q. Liu, S. Rabindran, R. Kumar, X. Hong, A. Goetz, T. Stanley, J. D. Taylor, S. D. Sigethy, G. H. Tomberlin, A. M. Hassell, K. M. Kahler,

L. M. Shewchuk and R. T. Gampe, *J. Med. Chem.*, 2012, **55**, 7193–7207.
36 J. M. Axten, S. P. Romeril, A. Shu, J. Ralph, J. R. Medina, Y. Feng, W. H. H. Li, S. W. Grant, D. A. Heerding, E. Minthorn, T. Mencken, N. Gaul, A. Goetz, T. Stanley, A. M. Hassell, R. T. Gampe, C. Atkins and R. Kumar, *ACS Med. Chem. Lett.*, 2013, **4**, 964–968.
37 D. R. Drew, A. N. Hodder, D. W. Wilson, M. Foley, I. Mueller, P. M. Siba, A. E. Dent, A. F. Cowman and J. G. Beeson, *PLoS One*, 2012, **7**, e51023.
38 L. D. Arnold, D. J. Calderwood, R. W. Dixon, D. N. Johnston, J. S. Kamens, R. Munschauer, P. Rafferty and S. E. Ratnofsky, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 2167–2170.
39 N. Zhang, B. Li, I. Al-Ramahi, X. Cong, J. M. Held, E. Kim, J. Botas, B. W. Gibson and L. M. Ellerby, *J. Biol. Chem.*, 2012, **287**, 21204–21213.
40 Y. Saito, H. Yuki, M. Kuratani, Y. Hashizume, S. Takagi, T. Honma, A. Tanaka, M. Shirouzu, J. Mikuni, N. Handa, I. Ogahara, A. Sone, Y. Najima, Y. Tomabechi, M. Wakiyama, N. Uchida, M. Tomizawa-Murasawa, A. Kaneko, S. Tanaka, N. Suzuki, H. Kajita, Y. Aoki, O. Ohara, L. D. Shultz, T. Fukami, T. Goto, S. Taniguchi, S. Yokoyama and F. Ishikawa, *Sci. Transl. Med.*, 2013, **5**, 181.
41 R. C. Murphy, K. K. Ojo, E. T. Larson, A. Castellanos-Gonzalez, B. G. K. Perera, K. R. Keyloun, J. E. Kim, J. G. Bhandari, N. R. Muller, C. L. M. J. Verlinde, A. C. White Jr, E. A. Merritt, W. C. Van Voorhis and D. J. Maly, *ACS Med. Chem. Lett.*, 2010, **1**, 331–335.
42 E. T. Larson, K. K. Ojo, R. C. Murphy, S. M. Johnson, Z. Zhang, J. E. Kim, D. J. Leibly, A. M. W. Fox, M. C. Reid, E. J. Dale, B. G. K. Perera, J. Kim, S. N. Hewitt, W. G. J. Hol, C. L. M. J. Verlinde, E. Fan, W. C. Van Voorhis, D. J. Maly and E. A. Merritt, *J. Med. Chem.*, 2012, **55**, 2803–2810.
43 A. Bansal, S. Singh, K. R. More, D. Hans, K. Nangalia, M. Yogavel, A. Sharma and C. E. Chitnis, *J. Biol. Chem.*, 2013, **288**, 1590–1602.
44 J. Seidler, S. L. McGovern, T. N. Doman and B. K. Shoichet, *J. Med. Chem.*, 2003, **46**, 4477–4486.

A.5 Structure and Dynamics of Apical Membrane Antigen 1 from *Plasmodium falciparum* FVO

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

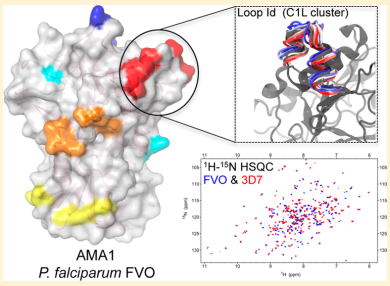
Structure and Dynamics of Apical Membrane Antigen 1 from *Plasmodium falciparum* FVO

San Sui Lim,[†] Wei Yang,[‡] Bankala Krishnarjuna,[†] Komagal Kannan Sivaraman,[‡] Indu R. Chandrashekaran,[†] Itamar Kass,[‡,§] Christopher A. MacRaild,[†] Shane M. Devine,[†] Cael O. Debono,[†] Robin F. Anders,[†,∥] Martin J. Scanlon,[†,⊥] Peter J. Scammells,[†] Raymond S. Norton,*[,†] and Sheena McGowan*[,‡]

[†]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia
[‡]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia
[§]Victorian Life Sciences Computation Initiative Life Sciences Computation Centre, Monash University, Clayton, Victoria 3800, Australia
[∥]Department of Biochemistry, La Trobe Institute for Molecular Science, La Trobe University, Bundoora, Victoria 3086, Australia
[⊥]Centre of Excellence for Coherent X-ray Science, Monash University, Parkville, Victoria 3052, Australia

Supporting Information

ABSTRACT: Apical membrane antigen 1 (AMA1) interacts with RON2 to form a protein complex that plays a key role in the invasion of host cells by malaria parasites. Blocking this protein−protein interaction represents a potential route to controlling malaria and related parasitic diseases, but the polymorphic nature of AMA1 has proven to be a major challenge to vaccine-induced antibodies and peptide inhibitors exerting strain-transcending inhibitory effects. Here we present the X-ray crystal structure of AMA1 domains I and II from *Plasmodium falciparum* strain FVO. We compare our new structure to those of AMA1 from *P. falciparum* 3D7 and *Plasmodium vivax*. A combination of normalized *B* factor analysis and computational methods has been used to investigate the flexibility of the domain I loops and how this correlates with their roles in determining the strain specificity of human antibody responses and inhibitory peptides. We also investigated the domain II loop, a key region involved in inhibitor binding, by comparison of multiple AMA1 crystal structures. Collectively, these results provide valuable insights that should contribute to the design of strain-transcending agents targeting *P. falciparum* AMA1.



Malaria is one of the most widespread infections, with more than 40% of the global population at risk of contracting the disease.[1,2] Each year, there are approximately 250 million clinical cases of malaria that result in more than 600000 deaths worldwide.[2] The majority of these deaths are due to *Plasmodium falciparum* infections occurring in young children in sub-Saharan Africa.[2] Although much less likely to cause death, *Plasmodium vivax* infections also contribute to a substantial malaria burden across the globe, with 70−80 million cases occurring annually.[3] Although current artemisinin combination therapies have been highly effective against *Plasmodium* parasites, signs of resistance have already emerged.[4] There is an urgent need to combat this threat using therapeutic agents that act against a broad range of parasite strains, especially those that have become resistant to available therapies.

Apical membrane antigen 1 (AMA1) forms part of the moving junction complex essential for erythrocyte invasion by *Plasmodium* merozoites, and ligands that disrupt AMA1 function inhibit the growth *in vitro* of *P. falciparum* asexual blood stages.[5−9] Further, a conditional knockdown of *Pf*AMA1 severely impaired the parasite's ability to invade red cells,[10] and a complete gene knockout is not viable in *P. falciparum*.[11] AMA1 is a type I integral membrane protein,[12] the extracellular region of which consists of three domains based on the connectivities of its eight intramolecular disulfide bonds: an N-terminal domain I, a central domain II, and a C-terminal domain III.[13,14] Antibodies to AMA1 can block red cell invasion by *P. falciparum in vitro* and contribute to the adaptive immune response that partially protects exposed individuals against malaria. AMA1 has been a leading candidate for inclusion in a vaccine against *P. falciparum*,[15] but in a Phase IIb clinical trial in Mali, a 3D7 *Pf*AMA1 vaccine provided protection against only a subset of *P. falciparum* AMA1 genotypes, reflecting the extensive sequence polymorphisms in this antigen.[16] A bivalent 3D7 and FVO *Pf*AMA1 vaccine, also tested in Phase IIb trials, failed to protect because of the poor immunogenicity of the

Received: September 25, 2014
Revised: October 30, 2014
Published: October 31, 2014

   dx.doi.org/10.1021/bi5012089 | *Biochemistry* 2014, 53, 7310−7320

alum formulation.[17] Currently, preclinical studies with multivalent vaccines (four to six AMA1 alleles) show promise, inducing a more broadly cross-reactive antibody response.[18,19]

The two most extensively studied forms of *Pf*AMA1, 3D7 and FVO, have 24 amino acid differences, and these polymorphic residues have been grouped into domain I−III clusters based on their spatial proximity on the X-ray crystal structure of 3D7 *Pf*AMA1.[20] The domain I cluster is the most important of the three in mediating escape from inhibitory antibodies and was further classified into subclusters C1−C3. Within C1, the region termed C1L is particularly important for immune escape.[18,21] In the preclinical studies of *Pf*AMA1 vaccines, it was noted that antisera to 3D7 *Pf*AMA1 were more strain-specific than antisera to FVO *Pf*AMA1.

AMA1 has a hydrophobic cleft that is the site of interactions with its protein-binding partner RON2.[9,22] The cleft is surrounded by six loops from domain I (loops Ia−If) and an extended loop from domain II (DII loop).[23] The DII loop appears to contain a strain-transcending epitope as the monoclonal antibody 4G2, which binds to the base of the DII loop, exhibits strain-independent inhibition of *P. falciparum*.[5] We and others have proposed that small molecules targeted to the hydrophobic cleft may interrupt the AMA1−RON2 protein−protein interaction and provide a route to novel therapeutics in the form of protein−protein interaction inhibitors.[24−26] However, polymorphic regions C1 and C3 surrounding one end of the cleft were shown to restrict the cross-reactivity of inhibitory antibodies and peptides such as IgNAR,[7] 1F9,[27,28] and R1.[6,9,29] It has been postulated that sequence variation in the highly polymorphic C1L region (of the C1 subcluster) may result in local secondary structure changes.[15,21] In particular, the presence of a two-turn helix in the Id loop has been questioned for the FVO allele because of the presence of a glycine residue at position 197.[15]

To answer fundamental questions regarding the impact of sequence diversity on AMA1, we have determined the first X-ray crystal structure of *Pf*AMA1 from FVO. We show through X-ray crystallography, all-atom molecular dynamics, and nuclear magnetic resonance (NMR) spectroscopy that the sequence divergence does not result in structural changes that account for the strain-specific effects documented for inhibitory antibodies and peptides.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Domains I and II of the ectodomains of FVO and 3D7 *Pf*AMA1 (residues 104−438) were produced according to the protocol described by Lim et al.[24] except that the hexa-His tag was removed using TEV protease. The pure His-tagged proteins were dialyzed using membrane tubing (Spectra/Por 3, 3.5 kDa MWCO) against a 100-fold volume of 50 mM Tris (pH 8.0) under constant stirring at 4 °C overnight, and then TEV protease was added to the sample at a 1:30 ratio (w/w) and cleavage allowed to proceed at 4 °C for 48 h with gentle agitation. The mixture was filtered (0.2 μm) and loaded onto a 5 mL CHT I ceramic hydroxyapatite column (Bio-Rad). The cleaved protein was eluted using a linear gradient of 10 to 150 mM phosphate buffers (Na_2HPO_4 and $NaH_2PO_4{\cdot}H_2O$) (pH 7.4) over 15 column volumes. The pooled fractions were concentrated and buffer exchanged into 20 mM Tris (pH 8.0) using an Amicon Ultra-4 centrifugal unit with an Ultracel-10 membrane (Millipore).

^{1}H NMR Spectroscopy and Size-Exclusion Chromatography. FVO *Pf*AMA1 purified from hydroxyapatite chromatography was buffer exchanged into 20 mM phosphate buffer (Na_2HPO_4 and $NaH_2PO_4{\cdot}H_2O$) and 50 mM NaCl (pH 7.4) containing 10% ^2H_2O using a PD-10 desalting column (GE Healthcare). The sample was subsequently concentrated as described above to a final protein concentration of 50 μM. Part of this final product was used for ^{1}H NMR. A ^{1}H-detected pulse program incorporating the excitation sculpting scheme for water suppression was employed to characterize FVO *Pf*AMA1. A total of 128 scans and 16K data points were acquired at 600 MHz on a Bruker Avance III spectrometer at 35 °C. The data were processed in Topspin 3.2 using an exponential multiplication function with 2 Hz line broadening. The water signal was used to reference the ^{1}H NMR spectrum. The final product was loaded onto a Superdex 200 10/30 GL column and eluted isocratically with 20 mM phosphate buffer (Na_2HPO_4 and $NaH_2PO_4{\cdot}H_2O$) and 50 mM NaCl (pH 7.4) at a flow rate of 0.5 mL/min.

^2H_2O *Escherichia coli* Adaptation. Fifty microliters of a competent *E. coli* BL21(DE3) glycerol stock, previously frozen at −80 °C, was thawed on ice for 10 min prior to adding 1 μL of plasmid carrying expression vector pPROEX HTb (Novagen) with FVO or 3D7 *Pf*AMA1$_{[104-438]}$ sequences. The mixture was left on ice for a further 30 min and then in water at 42 °C for 45 s. One milliliter of Luria broth (LB) was added to each sample, and the culture was incubated at 37 °C while being constantly shaken at 225 rpm. After 45 min, 50 μL of culture was spread over a LB plate containing 50% (v/v) ^2H_2O and 100 μg/mL ampicillin and then incubated at 37 °C overnight. A single colony of the freshly transformed cells was inoculated into 10 mL of LB medium with 50% (v/v) ^2H_2O and 100 μg/mL ampicillin. The culture was grown for 24 h at 37 °C while being constantly shaken at 225 rpm. The cell mixture was spread on a culture plate, and subsequently, a single colony of cells was incubated in growth medium as described above, except that the LB plate and medium prepared with 75% (v/v) ^2H_2O were used instead. The final cell culture containing 75% (v/v) ^2H_2O was stored at −80 °C with 20% (v/v) glycerol.

Isotopically Labeled AMA1. A scrape of the glycerol stock of ^2H_2O-adapted *E. coli* was inoculated into LB medium with 75% (v/v) ^2H_2O and 100 μg/mL ampicillin. The culture was incubated overnight at 37 °C while being constantly shaken at 225 rpm. The overnight culture was then centrifuged at 1500*g* for 15 min. The supernatant was decanted, and the cell pellets were resuspended in 2 volumes of optimized minimal medium[30] prepared with 100% (v/v) ^2H_2O, 1 g/L $^{15}NH_4Cl$, and 8 g/L [^{13}C]glucose. The cells were allowed to grow for 3 h while being shaken at 37 °C before being induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 24 h. The protein was then purified as described above.

^{1}H−^{15}N HSQC and Three-Dimensional (3D) HNCO Experiments. ^{2}H-, ^{15}N-, and ^{13}C-labeled 3D7 and FVO *Pf*AMA1 were dissolved at concentrations of 300 and 75 μM, respectively, in 20 mM phosphate buffer (Na_2HPO_4 and $NaH_2PO_4{\cdot}H_2O$) (pH 7.0) containing 50 mM L-arginine, 50 mM L-glutamic acid, 0.2% (w/v) protease inhibitor cocktail (Roche), 0.01% (w/v) sodium azide, and 10% (v/v) ^2H_2O. The FVO *Pf*AMA1 spectrum was acquired on a Bruker Avance III 600 MHz spectrometer at 35 °C. A spectrum of 3D7 *Pf*AMA1 was acquired on a Bruker Avance III 800 MHz spectrometer at 30 °C. Both the ^{1}H−^{15}N HSQC and 3D HNCO experiments

were conducted using pulse sequences with transverse relaxation-optimized spectroscopy (TROSY) effects.[31,32] $^{1}H-^{15}N$ HSQC spectra were acquired with 64 scans at 2048 and 256 data points for the ^{1}H and ^{15}N dimensions, respectively. A total of 32 scans was recorded for the 3D HNCO experiments, with 2048, 128, and 128 data points for the ^{1}H, ^{15}N, and ^{13}C dimensions, respectively. Both the direct and indirect dimensions of $^{1}H-^{15}N$ HSQC and 3D HNCO data were processed using a QSINE window with a phase shift of 2. Linear prediction using 32 coefficients was applied to all indirect dimensions.

R1−FVO *Pf*AMA1 Interactions. A Biacore T200 biosensor was employed to measure the interaction between recombinant FVO *Pf*AMA1 DI and DII and the R1 peptide (GL Biochem). Surface plasmon resonance (SPR) experiments were conducted essentially as described previously[24] except that dimethyl sulfoxide was not included in the running buffer. Approximately 8000 RU of protein was coupled in a single flow cell (1000 RU = 1 ng of protein/mm^2). The binding of R1 peptide[6,29] to FVO *Pf*AMA1 was evaluated using a 2-fold serial dilution ranging in concentration from 15.6 to 500 μM.

Sequence Alignment and Analysis. All sequence alignments and analyses were performed using the UniProt online tool (http://www.uniprot.org/). Accession numbers for AMA1 sequences used in this study are as follows: UniProt entry Q9TY48 for FVO *Pf*AMA1 and UniProt entry Q7KQK5 for 3D7 *Pf*AMA1. The residues defining the AMA1 hydrophobic cleft and polymorphic sites were obtained from published literature.[7,23,33−35]

Crystallization, X-ray Data Collection, Structure Determination, and Refinement. Crystallization conditions for FVO *Pf*AMA1 were identified following a robotic broad screen using the IndexHT (Hampton Research) and JCSGPlus (Molecular Dimensions) crystal screens. Optimization of a single initial hit from the Index screen used the hanging drop vapor diffusion method, with a 1:1 (v/v) ratio of protein to mother liquor (well volume of 0.5 mL). Small, stacked crystals appeared after three months in 25% (v/v) polyethylene glycol 3350, 0.1 M HEPES (pH 7.5), and 0.2 M $MgCl_2$. A single crystal was separated from the stacked cluster and cryoprotected by the addition of 10% glycerol prior to data collection.

3D7 *Pf*AMA1 crystals were grown in 12−15% (v/v) polyethylene glycol 3350, 0.02 M MES (pH 6.0), and 10 mM $MnCl_2$ as detailed in ref 23. 3D7 *Pf*AMA1 crystals were dehydrated overnight in a reservoir solution with an increased level [35% (v/v)] of polyethylene glycol 3350 before cryostabilization in 38% (v/v) polyethylene glycol 3350, 0.088 M MES (pH 6.0), and 44 mM $MnCl_2$ for 6−8 h prior to data collection. For crystals used to test soaking solvents, 5% (v/v) methanol or Milli-Q water was added to the stabilization solution.

Data were collected at 100 K for all crystals using the Australian Synchrotron micro crystallography MX2 beamline 3ID1. Diffraction images were processed using XDS[36] and AIMLESS[37] from the CCP4 suite.[38] Five percent of each data set was flagged for calculation of R_{Free}[39] with neither a sigma nor a low-resolution cutoff applied to the data. A summary of data collection statistics is provided in Table S1 of the Supporting Information.

Structure determination proceeded using the Molecular Replacement method and the program PHASER.[40] A search model for FVO *Pf*AMA1 was constructed by removing the solvent and flexible loops from a crystal structure of 3D7 *Pf*AMA1 [Protein Data Bank (PDB) entry 1Z40]. A single clear peak in both the rotation and translation functions was evident and packed well within the asymmetric unit. Together with the unbiased features in the initial electron density maps, the correctness of the molecular replacement solution was confirmed. All subsequent model building and structural validation for FVO and 3D7 *Pf*AMA1 structures was conducted using Phenix[41,42] and COOT.[43] Solvent molecules were added only if they had acceptable hydrogen bonding geometry contacts of 2.5−3.5 Å with protein atoms or with existing solvent and were in good $2F_o - F_c$ and $F_o - F_c$ electron density. Hydrogen bonds (excluding water-mediated bonds) and salt bridges were calculated using PDBePISA.[44] The coordinates and structure factors are available from the Protein Data Bank (entries 4R1A, 4R19, 4R1B, and 4R1C). Raw data and images are available from TARDIS[45] (www.tardis.edu.au).

***B* Factor Analysis.** The *B* factors obtained from PDB files cannot be used directly, because the values may be on different scales because of the application of different refinement procedures.[46] To compare the *B* factors from different structures, the values were normalized as described by Parthasarathy et al.[47] The Cα *B* factor values were extracted from FVO *Pf*AMA1 (PDB entry 4R1A) and *P. vivax* AMA1 (PDB entry 1W81) as well as the four 3D7 *Pf*AMA1 crystal structures (PDB entries 1Z40, 4R19, 4R1B, and 4R1C), and normalized using the equation $B_{normalized} = (C\alpha\ B\ factor - B_{mean})/\sigma(B)$, where B_{mean} and $\sigma(B)$ are the mean value and standard deviation, respectively, of the distribution of observed thermal factors. The average values and standard deviations of normalized *B* factors were calculated for α-helical and β-sheet regions, loops Ia−If, and the DII loop. Average values were not calculated for regions in which more than half of the amino acid sequence was missing in the crystal structures.

Molecular Dynamics Simulations. The dynamics of three different AMA1 structures were studied. Atomic coordinates of *P. vivax* AMA1 (PDB entry 1W81, residues 49−383) and 3D7 *Pf*AMA1 (PDB entry 1Z40, residues 104−438) were obtained from the Protein Data Bank. Coordinates for FVO *Pf*AMA1 (residues 104−438) were obtained in this study. Missing atoms and residues were modeled using MOE 2012.10.[48] Each protein was solvated in a water cubic box consisting of TIP3P water molecules[49] with Na^+ ions added to neutralize any charge. The minimum distance from the surface of the protein to any face of the water box was set to 12 Å for each simulation.

All-atom molecular dynamics (MD) simulations were performed using the NAMD 2.9 MD package[50] on the IBM Blue Gene/Q supercomputer of the Victorian Life Sciences Computation Initiative (VLSCI). Proteins were defined by the newly published and tested AMBER force field, FF12SB.[51−53] Equilibration was performed in three stages. First, potential steric clashes in the initial configurations were relieved with 50000 steps of energy minimization. Initial velocities for each system were assigned randomly according to a Maxwell−Boltzmann distribution at 100 K. Each system was then heated to 300 K over 0.1 ns, with the protein harmonically restrained (10 kcal mol^{-1} Å^{-2}) under the canonical ensemble (*NVT*) conditions. Following this, each system was simulated for an additional 0.1 ns under the isothermal−isobaric ensemble (*NPT*) conditions with all heavy protein atoms harmonically restrained (10 kcal mol^{-1} Å^{-2}). Thereafter, each system was subjected to 250 ns of free simulation.

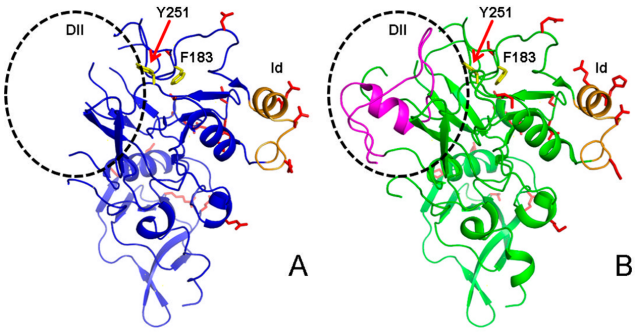


Figure 1. X-ray crystal structures of (A) FVO (PDB entry 4R1A) and (B) 3D7 (PDB entry 1Z40) *Pf*AMA1. Residues that vary between FVO and 3D7 *Pf*AMA1 proteins are shown as red sticks. The Id loops are colored orange in both structures. The DII loop is indicated by the dotted circle in both panels, and the structure of the DII loop in 1Z40 is colored magenta (B). Two residues critical for the AMA1−RON2 interaction (Phe183 and Tyr251) are shown as yellow sticks in both structures. The hydrophobic cleft runs across the top of the molecule in this view, from the Id loop at one end to the DII loop at the other.

For all simulations, an integration time of 2 fs was used and the nonbonded cutoff length was set at 1 nm. All simulations were conducted at a constant temperature (300 K) and pressure (1 atm), using a Langevin damping coefficient of 0.5 fs^{-1}. For each simulation system, periodic boundary conditions (*PBC*) were used together with the particle mesh Ewald (*PME*) method for electrostatics interactions.[54] Electrostatics and VDW nonbonded forces were cut off at 1 nm. For each protein, three trajectories were run in parallel, differing only in the distribution of their initial velocities. System conformations were saved every 10 ps for subsequent analysis.

All the analyses were performed using the GROMACS 4.0.7 simulation software package.[55] Prior to MD analyses, translational and rotational motions were eliminated by superposition of each frame onto the initial conformation. The root-mean-square deviations (rmsds) of the backbone heavy atoms in each system were calculated relative to their corresponding initial minimized structures. Backbone root-mean-square fluctuations (rmsfs) were calculated for the productive phase (50−250 ns) of each simulation. All images were created by VMD version 1.9.1[56] or PyMOL version 1.3r2 (Schrodinger, LLC, 2010, The PyMOL Molecular Graphics System).

Electrostatic Surface Potential Calculations. Protein electrostatic potentials were calculated using APBS version 1.3.[57] Atom electrostatic charges were taken from the FF12SB force field. The electrostatic potential was visualized using PyMOL 1.3r2 (Schrodinger, LLC) with positive potential in blue and negative potential in red over the range of $-3k_bT/e_c$ to $3k_bT/e_c$, where k_b is the Boltzmann constant, T is the temperature (set to 300 K), and e_c is the electron charge.

Effect of Mutations on the Stability of the R1−3D7 *Pf*AMA1 Complex. The contributions of specific mutations to the overall thermodynamic stability of the R1−3D7 *Pf*AMA1 complex structure[9] were estimated *in silico* with FoldX using default settings.[58] The reported Gibbs free energies are the differences between those of wild-type and mutated 3D7 *Pf*AMA1 in the context of the complex.

RESULTS

X-ray Crystal Structure of FVO *Pf*AMA1. To generate protein crystals of FVO *Pf*AMA1, a construct of FVO domains I and II (DI + II) equivalent to that of 3D7 *Pf*AMA1 was produced.[23] The quality of our recombinant FVO *Pf*AMA1 protein was assessed using ^{1}H NMR spectroscopy and size-exclusion chromatography (Figure S1 of the Supporting Information). Overall, good signal dispersion was observed in the ^{1}H NMR spectrum, with methyl proton signals at −0.48 and −0.52 ppm as well as amide proton signals beyond 9 ppm (Figure S1A of the Supporting Information); the spectrum was consistent with a single folded product. During size-exclusion chromatography, FVO *Pf*AMA1 eluted as a single peak consistent with a monomeric form of the protein (38 kDa) (Figure S1B of the Supporting Information).

The X-ray crystal structure of FVO *Pf*AMA1 was determined to 2.0 Å with final R and R_{free} values of 19.5 and 25.5%, respectively (Figure 1A and Table S1 of the Supporting Information). FVO *Pf*AMA1 crystallized with one molecule in the asymmetric unit in space group $C2_1$. Seven α-helical and 16 β-sheet regions were identified in the final FVO *Pf*AMA1 structure. Similar to the 3D7 *Pf*AMA1 (Figure 1B, PDB entry 1Z40) and *P. vivax* AMA1 structures (PDB entry 1W81), both DI and DII of FVO *Pf*AMA1 formed PAN folds that consist of a two-turn α-helix packed against a five-stranded β-sheet. The two PAN folds pack against each other to form the protein core, as seen in the 3D7 *Pf*AMA1[23] (PDB entry 1Z40; 0.27 Å rmsd over 219 Cα atoms) and *P. vivax*[59] AMA1 (PDB entry 1Z40; 0.29 Å rmsd over 197 Cα atoms) structures. The DI + II sequences of 3D7 and FVO *Pf*AMA1 have 21 amino acid residue differences [sharing 94% sequence identity (Figure S2 of the Supporting Information)].

All residue differences found in the C1 (residues 187, 190, 196, 197, 200, 204, 206, and 225) and C2 (residues 242, 243, 282, 283, and 285) clusters can be observed in our FVO *Pf*AMA1 structure (Figure S2 of the Supporting Information). Structural analysis of regions in the vicinity of the C1L (residues 196, 197, 200, 204, and 206), C2, and DII clusters, as well as residues 167 and 300, did not reveal any significant structural differences between FVO and 3D7 *Pf*AMA1 DI + II (Figure S3 of the Supporting Information). Structural differences were also not found in regions around residues 187 and 190 within the C1 cluster. Of the regions in the proximity of residue 225, residues 226−232 are disordered in FVO *Pf*AMA1 but ordered in 3D7 *Pf*AMA1. Structural comparisons could not be performed for C3 because residues 175 and 267, which

constitute this cluster, are disordered in both structures. Nine of the 12 residues that define the hydrophobic cleft could be observed in our FVO *Pf*AMA1 structure (V169, L176, F183, I190, Y202, V208, M224, Y251, and I252);[7] the three remaining hydrophobic residues (M273 from DI and L357 and F367 from DII) are disordered in the FVO *Pf*AMA1 structure.

The structure of FVO *Pf*AMA1 showed numerous disordered loops in both domains. In DI, the disordered residues not observed in the density (160−163, 173−176, 226−232, and 258−273) correspond to loops Ib and If and part of the Ie loop (Figure 1A). There are 38 residues missing in DII [351−388 (Figure 1A)], which correspond to most of the loop DII structure found in the 3D7 *Pf*AMA1 structure (Figure 1B).

Flexibility of the Loops in Different Forms of AMA1. To ascertain whether the sequence polymorphisms might provide a structure-based "escape" route from the host immune response and inhibitor binding, we analyzed the loop flexibility of our FVO *Pf*AMA1 structure in comparison to those of the published 3D7 *Pf*AMA1 (PDB entry 1Z40) and *P. vivax* AMA1 (PDB entry 1W81) structures. We compared normalized *B* factors from all Cα atoms from each AMA1. The normalized *B* values were expressed in units of standard deviations about the mean Cα *B* factor for the corresponding structure; therefore, regions that are more rigid in a protein would have low normalized *B* factors, whereas flexible regions would have high normalized *B* values. This analysis showed that loops Ib and If (polymorphic cluster C3) are highly flexible regions in all three structures (Figure 2 and Table S2 of the Supporting Information). Loops Ia, Ic (residues 187 and 190 of the C1 cluster), and Ie (residue 225 of the C1 cluster) are mobile in FVO *Pf*AMA1 and *P. vivax* AMA1 but exhibit limited mobility in 3D7 *Pf*AMA1. The Id loop (C1L cluster) appears to be more rigid in FVO *Pf*AMA1 than in 3D7 and *P. vivax* AMA1, but inferring the biological relevance of this result is difficult because of the presence of extensive crystal contacts made by residues within the Id loop of FVO *Pf*AMA1 (Figure S4 of the Supporting Information; see below).

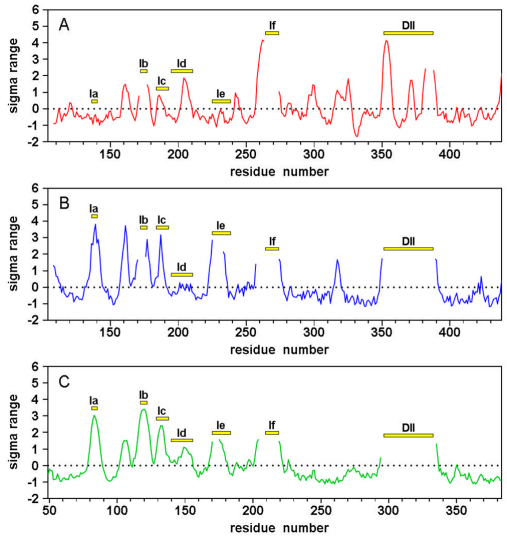


Figure 2. Flexibility of Ia−If and DII loops (identified with yellow bars). Normalized Cα *B* factors of (A) 3D7 *Pf*AMA1 (PDB entry 1Z40), (B) FVO *Pf*AMA1 (PDB entry 4R1A), and (C) *P. vivax* AMA1 (PDB entry 1W81). Residues 49−383 for the *P. vivax* AMA1 sequence are equivalent to residues 104−438 in the FVO and 3D7 *Pf*AMA1 sequences.

In an effort to further assess the flexibility of the three AMA1 proteins, we undertook all-atom molecular dynamics simulations. Throughout our MD simulations, all systems were found to be stable following an initial structural rearrangement that took place early in the simulations (30−50 ns). Therefore, all subsequent analyses were conducted for the last 200 ns of each simulation (productive stage). Calculated rmsd values for the productive stage of MD simulations [3D7 *Pf*AMA1, 0.19 ± 0.04 nm; FVO *Pf*AMA1, 0.20 ± 0.03 nm; and *P. vivax* AMA1, 0.30 ± 0.05 nm (Figure S5 of the Supporting Information)] indicated that all systems were stable and that the *P. vivax* protein was slightly more flexible than the other proteins [particularly in the region of loop Ia (Figure S6 of the Supporting Information)]. The rmsf results (Figure S6 of the Supporting Information) show that the fluctuation patterns of the three proteins were similar, with peak fluctuations occurring in the same loops of each AMA1 structure (loops Ib, Ic, Ie, and If), which is consistent with the *B* factor analysis. The Ie loop appears to be more flexible in FVO *Pf*AMA1 than in 3D7 *Pf*AMA1. This observation may explain why the Ie loop is disordered in the FVO *Pf*AMA1 crystal structure but not in 3D7 *Pf*AMA1. On the basis of the rmsf results, the Id loop is flexible in all three AMA1 structures, which supports our earlier conclusions that the rigidity of the Id loop (Figure 2B) seen in FVO *Pf*AMA1 was a consequence of crystal contacts.

Crystallization artifacts, such as crystal contacts, can often complicate structural comparisons and implications drawn from static structures. The presence of stabilizing crystal contacts was identified in the original 3D7 *Pf*AMA1 structure by Bai et al.[23] In the search for suitable crystallization conditions of 3D7 *Pf*AMA1 for our fragment screening campaign,[24] we noticed that such crystal contacts could be modulated by soaking crystals in different solvents. This allowed us to examine the effect of these contacts on the conformation and/or flexibility of the surface loops in 3D7 *Pf*AMA1. We reproduced the original crystal conditions of Bai et al.,[23] determined the 1.8 Å X-ray crystal structure (PDB entry 4R19), and showed that it was identical to the published structure (PDB entry 1Z40; 0.118 Å rmsd over 275 Cα atoms). Under these conditions, 3D7 *Pf*AMA1 crystallized in a $P3_1$ space group with two molecules per asymmetric unit and showed crystal contacts that potentially stabilize the Ia−If and DII loops surrounding the hydrophobic cleft (Table S3 of the Supporting Information). Subjecting the crystals to 5% methanol (PDB entry 4R1C) or water (PDB entry 4R1B) during their stabilization immediately prior to data collection produced changes in the space group and unit cell dimensions (Table S1 of the Supporting Information), and these two new 3D7 *Pf*AMA1 structures showed different degrees of flexibility within the Ia, Ic, and Ie loops relative to the published 3D7 *Pf*AMA1 structure (Table S2 of the Supporting Information). The 3D7 *Pf*AMA1 structures from the Milli-Q water- and methanol-treated crystals have higher average normalized *B* factors for the Ia, Ic, and Ie loops compared to those of the 1Z40 structure (Figures 3 and 4). The Ie loop of FVO *Pf*AMA1, which is disordered, is ordered in all the 3D7 *Pf*AMA1 structures.

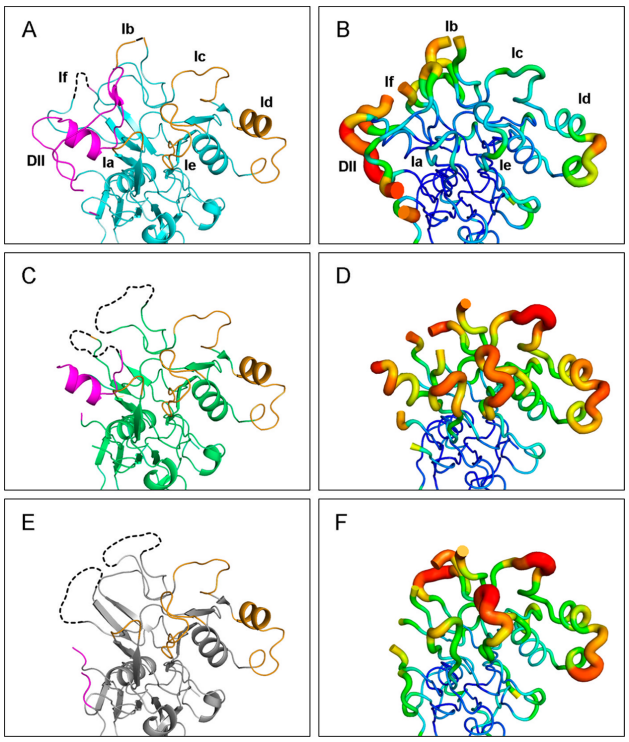


Figure 3. Cartoon model (left) and *B* factor putty (right) of the X-ray crystal structures of 3D7 *Pf*AMA1: (A and B) 3D7 crystal structure from ref 23, (C and D) 3D7 crystal treated with Milli-Q water (PDB entry 4R1B), and (E and F) 3D7 crystal treated with methanol (PDB entry 4R1C).

However, it is difficult to determine if there is a real difference in the flexibility of the Ie loop as crystal packing is found in all the 3D7 *Pf*AMA1 structures (Table S3 of the Supporting Information). The Ib and If loops are disordered in both the water- and methanol-treated 3D7 *Pf*AMA1 crystals, indicating that these regions are highly mobile in the protein (Figures 3 and 4 and Table S2 of the Supporting Information).

Conformational Flexibility of the DII Loop. The DII loop in both *P. vivax* AMA1 and our FVO *Pf*AMA1 (Figure 1A) structure is disordered. This is in contrast to the published 3D7 *Pf*AMA1 structure in which only five residues are missing from the DII loop (residues 383−387) (Figure 1B). However, there was missing density for DII loop residues 370−387 in our Milli-Q water-treated 3D7 *Pf*AMA1 structure and residues 351−387 in the methanol-treated 3D7 *Pf*AMA1 structure, similar to our FVO *Pf*AMA1 structure (Figures 1B and 3). MD simulations showed that, although the N- and C-termini of the DII loop are highly mobile for FVO *Pf*AMA1, 3D7 *Pf*AMA1, and *P. vivax* AMA1 (Figure S6 of the Supporting Information), large conformational changes or movements of the α-helix at the center of the DII loop were not observed in any of these proteins. This implies that, despite its flexibility, the DII loop undergoes slow conformational exchange, beyond the time scale sampled in our MD simulations.

The original 3D7 *Pf*AMA1 structure (PDB entry 1Z40) provides support for the DII loop being ordered as a consequence of crystal contacts.[23] Our investigation of these contacts found that there were 9 and 13 residues from neighboring molecules close to (<4.0 Å) the DII loop of reference 3D7 *Pf*AMA1 chains A and E, respectively (Figure S7A of the Supporting Information). In particular, we found that both Glu354 (Oε1) and His356 (NH) of chain E formed hydrogen bonds with the main and side chains of Ser423 of a symmetry-related molecule (Figure S7A of the Supporting Information). The DII loop of chain A was stabilized by polar interactions between Lys363 (Nζ) and Asp317 of a symmetry-related molecule. In our FVO structure, there are no crystal contacts close to the DII loop, possibly allowing the loop to populate different positions within the crystal lattice (Figure S7B of the Supporting Information).

NMR studies of FVO and 3D7 *Pf*AMA1 were undertaken to further assess their flexibility in solution. Both 2D $^{1}H-^{15}N$ HSQC and 3D HNCO spectra were acquired. Of the 316 amide backbone NMR resonances expected in both forms of *Pf*AMA1, only 261 and 250 peaks were identified in the HSQC spectra of FVO and 3D7 *Pf*AMA1, respectively (Figure S8 of the Supporting Information). The discrepancies between the expected and observed resonances are due to regions of the protein that undergo conformational exchange on the intermediate time scale (microseconds to milliseconds).[60] NMR signals from regions with such motion are often broadened beyond detection in multidimensional NMR experiments. In the HNCO spectra, there were fewer resonances in 3D7 *Pf*AMA1 than in FVO *Pf*AMA1 [248 and

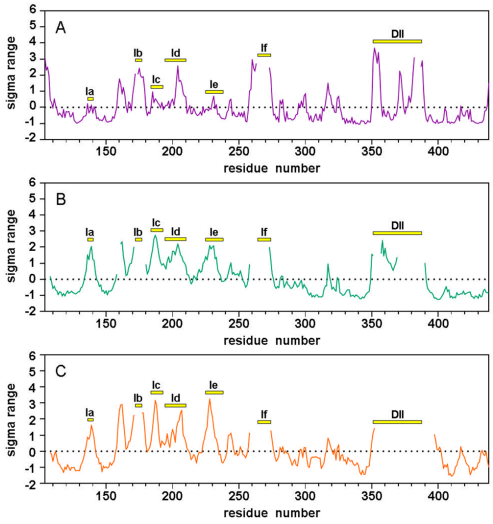


Figure 4. Flexibility of Ia−If and DII loops (identified with yellow bars) in 3D7 *Pf*AMA1. Normalized Cα *B* factors of (A) the 3D7 original condition,[23] (B) the 3D7 crystal treated with Milli-Q water, and (C) the 3D7 crystal treated with methanol.

216 peaks in FVO and 3D7 *Pf*AMA1, respectively (Figure S9 of the Supporting Information)]. This difference suggests that 3D7 *Pf*AMA1 has a slightly greater number of backbone resonances that are broadened by intermediate conformational exchange than FVO *Pf*AMA1. This is in contrast to what is predicted from comparison of the published 3D7 *Pf*AMA1 structure (PDB entry 1Z40), which is highly ordered throughout the molecule, with our crystal structures of both FVO *Pf*AMA1 and 3D7 *Pf*AMA1 in different solvents. In summary, these data indicate that in solution both FVO *Pf*AMA1 and 3D7 *Pf*AMA1 contain significant regions of disordered structure, some of which are undergoing conformational exchange on an intermediate time scale that produces substantial broadening of NMR resonances.[60]

Mapping the Strain Variation of the Inhibitory Peptide R1. The R1 peptide,[6,29] identified by phage display, inhibits red cell invasion by merozoites of 3D7 *P. falciparum* and related strains with a 50% inhibitory concentration (IC_{50}) of ~0.1 μM. X-ray crystal structures[23,25] show that R1 contacts three polymorphic residues (Tyr175, Met224, and Ile225) in 3D7 *Pf*AMA1, with residues at positions 175 and 225 being important determinants of R1 strain specificity. Substitution of these residues as in W2mef (I225E) or HB3 (Y175D and I225N) significantly reduced the peptide's inhibitory effect.[9] FVO *Pf*AMA1 also has Y175D and I225N substitutions (Figure S2 of the Supporting Information). Accordingly, we found that R1 binds weakly to FVO *Pf*AMA1, with an estimated K_d of $\geq$500 μM (Figure S10 of the Supporting Information).

To investigate why R1 binds so weakly to FVO *Pf*AMA1, we estimated the effect of mutations on binding of R1 by FVO *Pf*AMA1. The crystal structure of the 3D7 *Pf*AMA1−R1 complex (PDB entry 3SRJ) was used as a template, and four FVO *Pf*AMA1 sequence variations found in the R1-binding cleft (Y175D, M190I, D204N, and I225N) were generated by FoldX 3.0[61,62] to mimic an apparent FVO *Pf*AMA1−R1 complex-binding structure (Figure 5). Single mutations were also generated to compare the individual effects on the binding energy. Analysis of point mutations shows that the changes at positions 175 and 225 (Y to D and I to N, respectively) had the largest effect on stability, with increases in free energy of 6 and

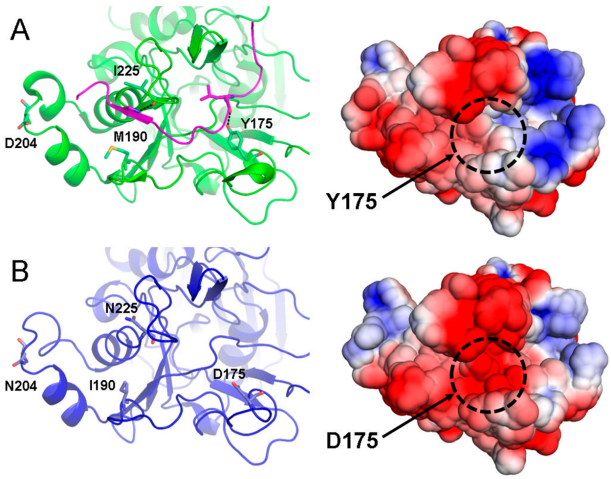


Figure 5. Comparison between the R1−3D7 *Pf*AMA1 complex and free FVO *Pf*AMA1. (A) R1−3D7 *Pf*AMA1 crystal structure (PDB entry 3SRJ). (B) FVO *Pf*AMA1 model. Electrostatic potentials mapped on the solvent accessible surface for the R1-bound *Pf*AMA1 structures are shown on the right. Sequence variations between 3D7 and FVO *Pf*AMA1 are shown as sticks. R1 is shown as a magenta ribbon. Protein surfaces are color-coded according to electrostatic potential gradient, where positively and negatively charged areas are colored blue and red (iso-values from $3k_bT/e_c$ to $-3k_bT/e_c$), respectively.

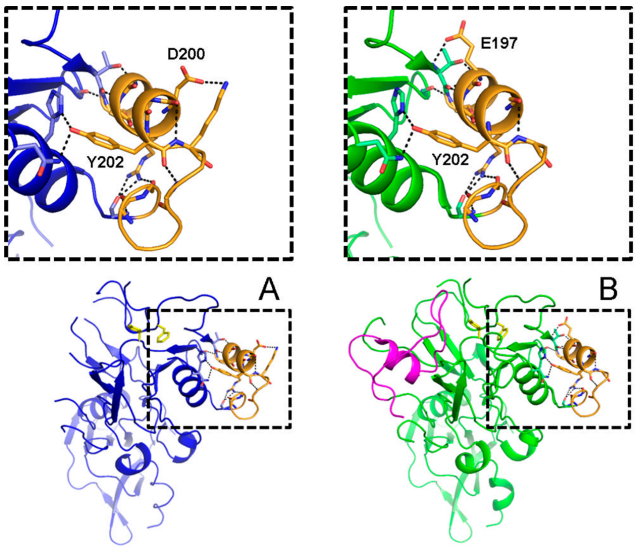


Figure 6. Cartoon showing polar interactions that stabilize the Id loop (orange) in (A) FVO (blue) and (B) 3D7 (green) *Pf*AMA1. The DII loop is colored magenta. Residues Phe183 and Tyr251 are shown as yellow sticks.

3 kcal/mol, respectively (Table S4 of the Supporting Information). These energy changes are presumably due to the loss of two hydrogen bonds between R1 and AMA1 at these two positions (Figure 5). The M190I change resulted in a small increase in free energy (1 kcal/mol), suggesting that the increase in cleft volume and hydrophobicity resulting from the M190I change has an effect on R1 binding. Distal to the R1 interaction sites, the D204N change has no effect on the complex. Taken together, the four changes were estimated to contribute to a total increase in ΔG of 10 kcal/mol and significant changes in the hydrophobic cleft, creating an unfavorable environment for the R1 peptide (Table S4 of the Supporting Information).

DISCUSSION

AMA1 is implicated in the invasion of host cells by malaria parasites as well as other apicomplexan parasites.[22,25,63,64] Sequence comparison of FVO and 3D7 *Pf*AMA1 domains I and II identifies 21 amino acid differences that occur exclusively at the polymorphic face of AMA1.[23,33,35] However, despite earlier hypotheses,[15,21] our FVO structure shows that these changes do not influence the overall fold of AMA1. The Ia−If loops that surround most of the hydrophobic cleft in AMA1 all display some level of disorder in the electron density from the FVO *Pf*AMA1, 3D7 *Pf*AMA1, and *P. vivax* structures. This apparent loop flexibility coupled with the polymorphic residues found in most of these regions provides AMA1 with an effective means of restricting cross-strain inhibitory activities of various ligands such as the R1 peptide and poses a challenge to efforts to design a vaccine or therapeutic agents effective against a broad range of strains and species of *Plasmodium* parasites.

Polymorphic residues within loop Id (C1L cluster) are important in mediating escape against AMA1 antibodies induced by *P. falciparum* infections or in vaccine trials.[23,25,65] Residue 197 appears to be one of the most important residues in AMA1 responsible for immune escape.[65,66] In this study, we have shown that residue changes, including a Gly at position 197, do not result in structural changes in this region. This includes the two-turn helix within C1L. Residues 196, 204, and 206 are not engaged in any polar interactions and are unlikely to be important in stabilizing the structure of the Id loop (Table S5 of the Supporting Information). The main chain atoms of residues 197 and 200 in both strains form polar interactions with main chain atoms of Thr194 and Lys203, respectively. Substitution of Glu with Gly at position 197 in FVO *Pf*AMA1 prevents the polar interaction between Thr194 [O] and Glu197 [Oγ1] observed in 3D7. The H200D change at position 200 results in an additional side chain electrostatic interaction between Asp200 [Oδ1] and Lys203 [Nζ] in the FVO form of AMA1. The extensive pattern of hydrogen bonds stabilizing the Id loop is largely conserved in both strains (Figure 6 and Table S5 of the Supporting Information). These interactions appear to stabilize the structure and permit radical changes at polymorphic sites in the Id loop without causing significant conformational changes. This implies that the polymorphic nature of this region does not affect the structure of the protein and that immune escape arises largely from changes in properties of individual side chains.

Unlike the published structures of 3D7 *Pf*AMA1 and one 1F9−3D7 *Pf*AMA1 complex (PDB entries 1Z40 and 2Q8A, respectively), in which the conserved DII loop is partially ordered, the DII loop of FVO *Pf*AMA1 is completely disordered. This is similar to the case in another crystal form of the 1F9−3D7 *Pf*AMA1 complex (PDB entry 2Q8B)[27] and the *P. vivax* AMA1 structure (PDB entry 1W81).[59] The flexibility of the DII loop is linked to a conformational change that allows the protein to interact with its protein-binding partner, RON2 (PDB entry 3ZWZ).[27] By sequence and structural superposition, it appears that the intraprotein contacts involved in stabilizing the local secondary structure of the DII loop are maintained in FVO *Pf*AMA1. The differences observed between the crystal structures of the two

forms of AMA1 reflect different individual conformational states captured under the different crystallization conditions rather than an inherently greater level of disorder in the FVO *Pf*AMA1 DII loop. This view is further supported by our 3D7 *Pf*AMA1 structures, in which different extents of crystal contacts gave rise to different degrees of order in the DII loop.

In conclusion, our results show that the overall structure, including the flexible nature of the DII loop, is conserved between the FVO and 3D7 forms of AMA1. The interacting interfaces between DI and the DII loop consist of invariant residues across all *P. falciparum* strains. The structural conservation of DI and DII appears likely to be conserved across all allelic forms of AMA1 and hence represents an attractive site for strain-transcending therapeutic interventions. Given that DII loop displacement is associated with formation of the AMA1−RON2 complex, it is conceivable that stabilizing the DII loop in its ordered state would be inhibitory to AMA1 function.[9] Recently, we have undertaken a fragment-based screening campaign against AMA1 to identify chemical scaffolds capable of inhibiting the protein−protein interactions.[24] The structure described here will help guide the design of small molecule inhibitors of AMA1 with broad strain specificity.

■ ASSOCIATED CONTENT

Supporting Information

^{1}H NMR and size-exclusion chromatography of FVO *Pf*AMA1 (Figure S1), sequence alignment of 3D7 and FVO *Pf*AMA1 (Figure S2), 3D7 and FVO *Pf*AMA1 sequence variations in the D1 and D2 polymorphic clusters (Figure S3), crystal packing of the FVO *Pf*AMA1 Id loop (Figure S4), rmsds of FVO *Pf*AMA1, 3D7 *Pf*AMA1, and *P. vivax* AMA1 (Figure S5), rmsfs of FVO *Pf*AMA1, 3D7 *Pf*AMA1, and *P. vivax* AMA1 (Figure S6), crystal packing against the *Pf*AMA1 DII loop (Figure S7), ^{1}H−^{15}N HSQC spectra of FVO and 3D7 *Pf*AMA1 (Figure S8), ^{1}H−^{13}C projection of HNCO spectra of FVO and 3D7 *Pf*AMA1 (Figure S9), SPR of R1−FVO *Pf*AMA1 interaction (Figure S10), X-ray data collection and refinement statistics (Table S1), average values of normalized *B* factors for Ia−If and DII loops (Table S2), crystal packing in AMA1 crystal structures (Table S3), FoldX-calculated energies (Table S4), and polar interactions involved in stabilizing the AMA1 Id loops (Table S5). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: ray.norton@monash.edu. Telephone: +613 9903 9167. Fax: +613 9903 9582.

*E-mail: sheena.mcgowan@monash.edu. Telephone: +613 9902 9309. Fax: +613 9902 9500.

Funding

S.M. is an Australian Research Council Future Fellow (FT100100690). R.S.N. acknowledges fellowship support from the National Health and Medical Research Council of Australia. This work was supported by the National Health and Medical Research Council (Project Grant 1025150). This research was supported by the Victorian Life Sciences Computation Initiative Life Sciences Computation Centre (RAS Grant VR0316), a collaboration among Melbourne, Monash, and La Trobe Universities and an initiative of the Victorian Government of Australia.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Nyssa Drinkwater for technical assistance. We thank the Australian Synchrotron for beam time and technical assistance.

■ ABBREVIATIONS

AMA1, apical membrane antigen 1; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; *Pf*, *P. falciparum*; TROSY, transverse relaxation-optimized spectroscopy; MD, molecular dynamics; SPR, surface plasmon resonance; MWCO, molecular weight cutoff; TEV, tobacco etch virus; LB, Luria broth; IPTG, isopropyl β-D-1-thiogalactopyranoside; rmsd, root-mean-square deviation; rmsf, root-mean-square fluctuation; PEG, polyethylene glycol; MES, 2-(*N*-morpholino)ethanesulfonic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; PME, particle mesh Ewald.

■ REFERENCES

(1) White, N. J., Dondorp, A. M., Faiz, A., Mishra, S., and Hien, T. T. (2012) New global estimates of malaria deaths. *Lancet 380*, 559−560.

(2) *World Malaria Report* (2013) World Health Organization, Geneva.

(3) Mendis, K., Sina, B. J., Marchesini, P., and Carter, R. (2001) The neglected burden of *Plasmodium vivax* malaria. *Am. J. Trop. Med. Hyg. 64*, 97−106.

(4) Duffy, P. E., and Sibley, C. H. (2005) Are we losing artemisinin combination therapy already? *Lancet 366*, 1908−1909.

(5) Collins, C. R., Withers-Martinez, C., Bentley, G. A., Batchelor, A. H., Thomas, A. W., and Blackman, M. J. (2007) Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1. *J. Biol. Chem. 282*, 7431−7441.

(6) Harris, K. S., Casey, J. L., Coley, A. M., Karas, J. A., Sabo, J. K., Tan, Y. Y., Dolezal, O., Norton, R. S., Hughes, A. B., Scanlon, D., and Foley, M. (2009) Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning. *J. Biol. Chem. 284*, 9361−9371.

(7) Henderson, K. A., Streltsov, V. A., Coley, A. M., Dolezal, O., Hudson, P. J., Batchelor, A. H., Gupta, A., Bai, T., Murphy, V. J., Anders, R. F., Foley, M., and Nuttall, S. D. (2007) Structure of an IgNAR-AMA1 complex: Targeting a conserved hydrophobic cleft broadens malarial strain recognition. *Structure 15*, 1452−1466.

(8) Li, F., Dluzewski, A., Coley, A. M., Thomas, A., Tilley, L., Anders, R. F., and Foley, M. (2002) Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes. *J. Biol. Chem. 277*, 50303−50310.

(9) Vulliez-Le Normand, B., Tonkin, M. L., Lamarque, M. H., Langer, S., Hoos, S., Roques, M., Saul, F. A., Faber, B. W., Bentley, G. A., Boulanger, M. J., and Lebrun, M. (2012) Structural and functional insights into the malaria parasite moving junction complex. *PLoS Pathog. 8*, e1002755.

(10) Yap, A., Azevedo, M. F., Gilson, P. R., Weiss, G. E., O'Neill, M. T., Wilson, D. W., Crabb, B. S., and Cowman, A. F. (2014) Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites. *Cell. Microbiol. 16*, 642−656.

(11) Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S., and Cowman, A. F. (2000) Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species. *Mol. Microbiol. 38*, 706−718.

(12) Waters, A. P., Thomas, A. W., Deans, J. A., Mitchell, G. H., Hudson, D. E., Miller, L. H., McCutchan, T. F., and Cohen, S. (1990) A merozoite receptor protein from *Plasmodium knowlesi* is highly

conserved and distributed throughout *Plasmodium*. *J. Biol. Chem. 265*, 17974−17979.
(13) Hodder, A. N., Crewther, P. E., Matthew, M. L., Reid, G. E., Moritz, R. L., Simpson, R. J., and Anders, R. F. (1996) The disulfide bond structure of *Plasmodium* apical membrane antigen-1. *J. Biol. Chem. 271*, 29446−29452.
(14) Nair, M., Hinds, M. G., Coley, A. M., Hodder, A. N., Foley, M., Anders, R. F., and Norton, R. S. (2002) Structure of domain III of the blood-stage malaria vaccine candidate, *Plasmodium falciparum* apical membrane antigen 1 (AMA1). *J. Mol. Biol. 322*, 741−753.
(15) Anders, R. F., Adda, C. G., Foley, M., and Norton, R. S. (2010) Recombinant protein vaccines against the asexual blood stages of *Plasmodium falciparum*. *Hum. Vaccines 6*, 39−53.
(16) Thera, M. A., Doumbo, O. K., Coulibaly, D., Laurens, M. B., Ouattara, A., Kone, A. K., Guindo, A. B., Traore, K., Traore, I., Kouriba, B., Diallo, D. A., Diarra, I., Daou, M., Dolo, A., Tolo, Y., Sissoko, M. S., Niangaly, A., Sissoko, M., Takala-Harrison, S., Lyke, K. E., Wu, Y., Blackwelder, W. C., Godeaux, O., Vekemans, J., Dubois, M. C., Ballou, W. R., Cohen, J., Thompson, D., Dube, T., Soisson, L., Diggs, C. L., House, B., Lanar, D. E., Dutta, S., Heppner, D. G., Jr., and Plowe, C. V. (2011) A field trial to assess a blood-stage malaria vaccine. *N. Engl. J. Med. 365*, 1004−1013.
(17) Ouattara, A., Mu, J. B., Takala-Harrison, S., Saye, R., Sagara, I., Dicko, A., Niangaly, A., Duan, J. H., Ellis, R. D., Miller, L. H., Su, X. Z., Plowe, C. V., and Doumbo, O. K. (2010) Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine. *Malar. J. 9*, 175.
(18) Dutta, S., Dlugosz, L. S., Drew, D. R., Ge, X., Ababacar, D., Rovira, Y. I., Moch, J. K., Shi, M., Long, C. A., Foley, M., Beeson, J. G., Anders, R. F., Miura, K., Haynes, J. D., and Batchelor, A. H. (2013) Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1. *PLoS Pathog. 9*, e1003840.
(19) Miura, K., Herrera, R., Diouf, A., Zhou, H., Mu, J., Hu, Z., MacDonald, N. J., Reiter, K., Nguyen, V., Shimp, R. L., Jr., Singh, K., Narum, D. L., Long, C. A., and Miller, L. H. (2013) Overcoming allelic specificity by immunization with five allelic forms of *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun. 81*, 1491−1501.
(20) Dutta, S., Lee, S. Y., Batchelor, A. H., and Lanar, D. E. (2007) Structural basis of antigenic escape of a malaria vaccine candidate. *Proc. Natl. Acad. Sci. U.S.A. 104*, 12488−12493.
(21) Harris, K. S., Adda, C. G., Khore, M., Drew, D. R., Valentini-Gatt, A., Fowkes, F. J., Beeson, J. G., Dutta, S., Anders, R. F., and Foley, M. (2014) Immunodampening to overcome diversity in the malarial vaccine candidate apical membrane antigen 1. *Infect. Immun.*, DOI: 10.1128/IAI.02061-14.
(22) Tonkin, M. L., Crawford, J., Lebrun, M. L., and Boulanger, M. J. (2013) *Babesia divergens* and *Neospora caninum* apical membrane antigen 1 structures reveal selectivity and plasticity in apicomplexan parasite host cell invasion. *Protein Sci. 22*, 114−127.
(23) Bai, T., Becker, M., Gupta, A., Strike, P., Murphy, V. J., Anders, R. F., and Batchelor, A. H. (2005) Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U.S.A. 102*, 12736−12741.
(24) Lim, S. S., Debono, C. O., MacRaild, C. A., Chandrashekaran, I. R., Dolezal, O., Anders, R. F., Simpson, J. S., Scanlon, M. J., Devine, S. M., Scammells, P. J., and Norton, R. S. (2013) Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening. *Aust. J. Chem. 66*, 1530−1536.
(25) MacRaild, C. A., Anders, R. F., Foley, M., and Norton, R. S. (2011) Apical membrane antigen 1 as an anti-malarial drug target. *Curr. Top. Med. Chem. 11*, 2039−2047.
(26) Srinivasan, P., Yasgar, A., Luci, D. K., Beatty, W. L., Hu, X., Andersen, J., Narum, D. L., Moch, J. K., Sun, H., Haynes, J. D., Maloney, D. J., Jadhav, A., Simeonov, A., and Miller, L. H. (2013) Disrupting malaria parasite AMA1−RON2 interaction with a small molecule prevents erythrocyte invasion. *Nat. Commun. 4*, 2261.
(27) Coley, A. M., Gupta, A., Murphy, V. J., Bai, T., Kim, H., Foley, M., Anders, R. F., and Batchelor, A. H. (2007) Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody. *PLoS Pathog. 3*, 1308−1319.
(28) Coley, A. M., Parisi, K., Masciantonio, R., Hoeck, J., Casey, J. L., Murphy, V. J., Harris, K. S., Batchelor, A. H., Anders, R. F., and Foley, M. (2006) The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody. *Infect. Immun. 74*, 2628−2636.
(29) Harris, K. S., Casey, J. L., Coley, A. M., Masciantonio, R., Sabo, J. K., Keizer, D. W., Lee, E. F., McMahon, A., Norton, R. S., Anders, R. F., and Foley, M. (2005) Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1. *Infect. Immun. 73*, 6981−6989.
(30) Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., and Li, Q. (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci. 18*, 936−948.
(31) Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U.S.A. 95*, 13585−13590.
(32) Riek, R., Pervushin, K., and Wüthrich, K. (2000) TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends Biochem. Sci. 25*, 462−468.
(33) Chesne-Seck, M. L., Pizarro, J. C., Vulliez-Le Normand, B., Collins, C. R., Blackman, M. J., Faber, B. W., Remarque, E. J., Kocken, C. H., Thomas, A. W., and Bentley, G. A. (2005) Structural comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites. *Mol. Biochem. Parasitol. 144*, 55−67.
(34) Gunasekera, A. M., Wickramarachchi, T., Neafsey, D. E., Ganguli, I., Perera, L., Premaratne, P. H., Hartl, D., Handunnetti, S. M., Udagama-Randeniya, P. V., and Wirth, D. F. (2007) Genetic diversity and selection at the *Plasmodium vivax* apical membrane antigen-1 (*Pv*AMA-1) locus in a Sri Lankan population. *Mol. Biol. Evol. 24*, 939−947.
(35) Remarque, E. J., Faber, B. W., Kocken, C. H., and Thomas, A. W. (2008) Apical membrane antigen 1: A malaria vaccine candidate in review. *Trends Parasitol. 24*, 74−84.
(36) Kabsch, W. (2010) Xds. *Acta Crystallogr. D66*, 125−132.
(37) Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D69*, 1204−1214.
(38) Collaborative Computational Project, No. 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760−763.
(39) Brunger, A. T. (1993) Assessment of phase accuracy by cross validation: The free *R* value. Methods and applications. *Acta Crystallogr. D49*, 24−36.
(40) McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr. D61*, 458−464.
(41) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213−221.
(42) Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D68*, 352−367.
(43) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.
(44) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol. 372*, 774−797.
(45) Androulakis, S., Schmidberger, J., Bate, M. A., DeGori, R., Beitz, A., Keong, C., Cameron, B., McGowan, S., Porter, C. J., Harrison, A., Hunter, J., Martin, J. L., Kobe, B., Dobson, R. C., Parker, M. W., Whisstock, J. C., Gray, J., Treloar, A., Groenewegen, D., Dickson, N.,

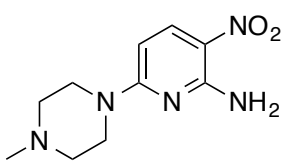
and Buckle, A. M. (2008) Federated repositories of X-ray diffraction images. *Acta Crystallogr. D64*, 810−814.
(46) Tronrud, D. E. (1996) Knowledge-based B-factor restraints for the refinement of proteins. *J. Appl. Crystallogr. 29*, 100−104.
(47) Parthasarathy, S., and Murthy, M. R. N. (1997) Analysis of temperature factor distribution in high-resolution protein structures. *Protein Sci. 6*, 2561−2567.
(48) *Molecular Operating Environment (MOE)* (2012) CCGI, Montreal, QC.
(49) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(50) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(51) Case, D., Darden, T., Cheatham, T., III, Simmerling, C., Wang, J., Duke, R., Luo, R., Walker, R., Zhang, W., and Merz, K. (2012) *AMBER 12*, University of California, San Francisco.
(52) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf. 65*, 712−725.
(53) Wickstrom, L., Okur, A., and Simmerling, C. (2009) Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. *Biophys. J. 97*, 853−856.
(54) Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577−8593.
(55) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput. 4*, 435−447.
(56) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.
(57) Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 10037−10041.
(58) Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: An online force field. *Nucleic Acids Res. 33*, W382−W388.
(59) Pizarro, J. C., Vulliez-Le Normand, B., Chesne-Seck, M. L., Collins, C. R., Withers-Martinez, C., Hackett, F., Blackman, M. J., Faber, B. W., Remarque, E. J., Kocken, C. H., Thomas, A. W., and Bentley, G. A. (2005) Crystal structure of the malaria vaccine candidate apical membrane antigen 1. *Science 308*, 408−411.
(60) Ge, X., MacRaild, C. A., Devine, S. M., Debono, C. O., Wang, G., Scammells, P. J., Scanlon, M. J., Anders, R. F., Foley, M., and Norton, R. S. (2014) Ligand-induced conformational change of *Plasmodium falciparum* AMA1 detected using ^{19}F NMR. *J. Med. Chem. 57*, 6419−6427.
(61) Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol. 320*, 369−387.
(62) Schymkowitz, J. W. H., Rousseau, F., Martins, I. C., Ferkinghoff-Borg, J., Stricher, F., and Serrano, L. (2005) Prediction of water and metal binding sites and their affinities by using the Fold-X force field. *Proc. Natl. Acad. Sci. U.S.A. 102*, 10147−10152.
(63) Hehl, A. B., Lekutis, C., Grigg, M. E., Bradley, P. J., Dubremetz, J. F., Ortega-Barria, E., and Boothroyd, J. C. (2000) *Toxoplasma gondii* homologue of *Plasmodium* apical membrane antigen 1 is involved in invasion of host cells. *Infect. Immun. 68*, 7078−7086.
(64) Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon-Guyot, J., Dubremetz, J. F., Fauquenoy, S., Tomavo, S., Faber, B. W., Kocken, C. H., Thomas, A. W., Boulanger, M. J., Bentley, G. A., and Lebrun, M. (2011) The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites. *PLoS Pathog. 7*, e1001276.
(65) Takala, S. L., Coulibaly, D., Thera, M. A., Batchelor, A. H., Cummings, M. P., Escalante, A. A., Ouattara, A., Traore, K., Niangaly, A., Djimde, A. A., Doumbo, O. K., and Plowe, C. V. (2009) Extreme polymorphism in a vaccine antigen and risk of clinical malaria: Implications for vaccine development. *Sci. Transl. Med. 1*, 2ra5.
(66) Ouattara, A., Takala-Harrison, S., Thera, M. A., Coulibaly, D., Niangaly, A., Saye, R., Tolo, Y., Dutta, S., Heppner, D. G., Soisson, L., Diggs, C. L., Vekemans, J., Cohen, J., Blackwelder, W. C., Dube, T., Laurens, M. B., Doumbo, O. K., and Plowe, C. V. (2013) Molecular basis of allele-specific efficacy of a blood-stage malaria vaccine: Vaccine development implications. *J. Infect. Dis. 207*, 511−519.

Appendix B – Additional Scaffold Investigation

These compounds were synthesized as a separate investigation from the MIPS fragment library investigation of *Pf*AMA1. All intermediates/analogues were derived from a separate paper, Ramachandran *et al.* that identified *N*-aryl-2-aminobenzimdazoles as novel hits against the asexual blood stage of *Plasmodium falciparum*. Although no target or mode of action was disclosed within the paper, our objective was to use resources available for the at Monash University for the AMA1 project to screen compounds disclosed within the paper against 3D7 and FVO *Pf*AMA1 to determine if AMA1 was a target for the inhibitory activity of the scaffold. The synthesized compounds were not active and therefore not evaluated further. A separate compound numbering system will be employed herein.

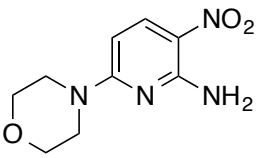
The paper:
Ramachandran, S.; Hameed P, S.; Srivastava, A.; Shanbhag, G.; Morayya, S.; Rautela, N.; Awasthy, D.; Kavanagh, S.; Bharath, S.; Reddy, J.; Panduga, V.; Prabhakar, K. R.; Saralaya, R.; Nanduri, R.; Raichurkar, A.; Menasinakai, S.; Achar, V.; Jiménez-Díaz, M. B.; Martínez, M. S.; Angulo-Barturen, I.; Ferrer, S.; Sanz, L. M.; Gamo, F. J.; Duffy, S.; Avery, V. M.; Waterson, D.; Lee, M. C. S.; Coburn-Flynn, O.; Fidock, D. A.; Iyer, P. S.; Narayanan, S.; Hosagrahara, V.; Sambandamurthy, V. K. *N*-aryl-2-aminobenzimidazoles: novel, efficacious, antimalarial lead compounds. *Journal of Medicinal Chemistry* **2014**, *57*, 6642–6652.

*Synthesis of 6-(4-methylpiperazin-1-yl)-3-nitropyridin-2-amine (**1**)*



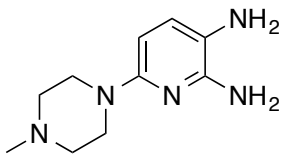
To a solution of 2-amino-6-chloro-3-nitropyridine (500 mg, 2.88 mmol) in anhydrous DMF (3 mL), was added 1-methylpiperazine (433 mg, 480 μL, 4.32 mmol). A nitrogen atmosphere was flowed over the reaction mixture before K_2CO_3 (1.1 g, 8.64 mmol) was added. The resultant reddish mixture was heated for 16 h at 120 °C. Once complete, the reaction was quenched with H_2O (100 mL) and extracted using EtOAc (2 x 40 mL). The combined organic layers were washed with brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. Title compound was afforded once dried (632 mg, 93 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 8.04 (d, 1H, J = 9.5 Hz, Ar), 8.02-7.85 (m, 1H, NH), 7.85-7.63 (m, 1H, NH), 6.34 (d, 1H, J = 9.5 Hz, Ar), 3.70 (d, 4H, J = 0.4 Hz, CH_2), 2.35 (dd, 4H, J = 6.6 Hz, 3.5 Hz, CH_2), 2.20 (s, 3H, CH_3).

*Synthesis of 6-morpholino-3-nitropyridin-2-amine (**2**)*



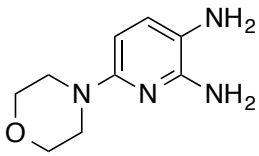
To a solution of 2-amino-6-chloro-3-nitropyridine (2 g, 11.52 mmol) in anhydrous DMF (10 mL), was added morpholine (1.5 g, 1.51 mL, 17.28 mmol). A nitrogen atmosphere was flowed over the reaction mixture before K_2CO_3 (4.7 g, 34.57 mmol) was added. The resultant reddish mixture was heated for 24 h at 120 °C. Once complete, the reaction was quenched with H_2O (300 mL) and extracted using EtOAc (2 x 100 mL). The combined organic layers were washed with brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. Title compound was afforded once dried (2.2 g, 87 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 8.07 (d, 1H, J = 9.5 Hz, Ar), 7.98 (s, 1H, NH), 7.72 (s, 1H, NH), 6.32 (d, 1H, J = 9.5 Hz, Ar), 3.68-3.64 (m, 8H, CH_2).

*Synthesis of 6-(4-methylpiperazin-1-yl)pyridine-2,3-diamine (**3**)*



1 (2 g, 8.42 mmol) was dissolved in MeOH: DCM (40 %) (120 mL) and stirred at rt. N_2 was flowed over the reaction mixture before Pd on carbon (10 %) (89 mg, 0.84 mmol) was added. Once the Pd on carbon was added, the mixture was sealed and flushed three consecutive times with both vacuum and H_2 respectively. The reaction was stirred for 16 h at rt. Once complete, the reaction mixture was filtered using celite and washed with MeOH (25 mL). The resultant green extract gave a black solid (1.8 g) and was purified via column chromatography using (DCM: MeOH: TFA = 89: 10: 1) to afford a black amorphous solid (1.2 g, 72 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 6.69 (d, 1H, J = 8.0 Hz, Ar), 5.86 (d, 1H, J = 8.0 Hz, Ar), 5.23 (s, 2H, NH_2), 3.42 (dq, 2H, J = 4.4 Hz, 1.4 Hz, NH_2), 3.36-3.34 (m, 4H, CH_2), 2.93 (s, 4H, CH_2), 2.57 (s, 3H, CH_3).

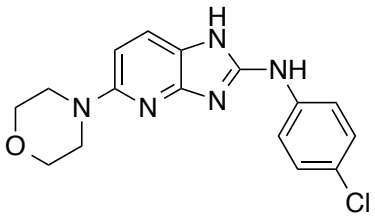
*Synthesis of 6-morpholinopyridine-2,3-diamine (**4**)*



2 (2 g, 8.91 mmol) was dissolved in MeOH: DCM (10 %) (120 mL) and stirred at rt. N_2 was flowed over the reaction mixture before Pd on carbon (10 %) (95 mg, 0.89 mmol) was added. Once the Pd on carbon was added, the mixture was sealed and flushed three consecutive times with both vacuum and H_2 respectively. The reaction was stirred for 16 h at rt. Once complete, the reaction mixture was filtered using celite and washed with MeOH (25 mL). The resultant green extract gave a black solid (1.8 g) and purified via column chromatography using (DCM: MeOH: TFA = 89: 10: 1) to afford a black amorphous solid

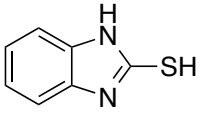
(492 mg, 28 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 6.67 (d, 1H, *J* = 8.0 Hz, Ar), 5.79 (d, 1H, *J* = 8.0 Hz, Ar), 5.15 (s, 2H, NH_2), 4.01 (t, 2H, *J* = 7.1 Hz, NH_2), 3.65 (t, 4H, *J* = 4.8 Hz, CH_2), 3.10 (t, 4H, *J* = 4.8 Hz, CH_2).

*Synthesis of N-(4-chlorophenyl)-5-morpholino-1H-imidazo[4,5-b]pyridin-2-amine (**5**)*



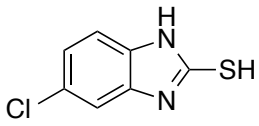
To a solution of **4** (150 mg, 0.77 mmol) in THF (5 mL), EDC.HCl (296 mg, 1.54 mmol) was added and allowed to dissolve at rt. 4-chlorophenyl isothiocyanate (144 mg, 0.84 mmol) was added to the solution and the mixture was sealed in a microwave vial and heated for 2 h at 100 °C. After 2 h, the reaction mixture was added drop-wise to cold H_2O. A brown precipitate forms and was collected via vacuum filtration to yield a brown powder (107 mg). The resultant solid was purified via column chromatography using (DCM: MeOH = 95: 5) to afford title compound (38 mg, 15 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.08 (m, 1H, NH), 9.53 (d, 1H, *J* = 122.1 Hz, NH), 7.82-7.78 (m, 2H, Ar), 7.49 (dd, 1H, *J* = 42.9 Hz, 8.5 Hz, Ar), 7.37-7.31 (m, 2H, Ar), 6.52 (dd, 1H, *J* = 26.1 Hz, 8.5 Hz, Ar), 3.76-3.69 (m, 4H, CH_2), 3.38-3.35 (m, 4H, CH_2); ^{13}C NMR (101 MHz; DMSO-d_6): δ 156.7, 148.6, 148.5, 129.4, 128.5, 124.7, 118.7, 118.4, 99.8, 66.1, 66.0, 46.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 330.1036; Observed $[M+H]^+$ = 330.1109.

*Synthesis of 1H-benzo[d]imidazole-2-thiol (**6**)*



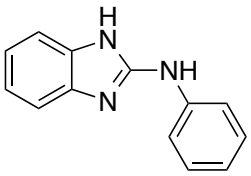
o-Phenylenediamine (350 mg, 3.23 mmol) was dissolved in a mixture of (EtOH: H_2O = 6: 1). Potassium ethyl xanthogenate (779 mg, 4.85 mmol) was added to the stirred solution and solution was refluxed for 16 h at 75 °C. After the 16 h, the reaction was allowed to cool to rt and reduced *in vacuo*. A white precipitate forms when solvent reduces. The rest of the mixture was added drop-wise to cold H_2O to precipitate remaining product in solution. The resultant solid was collected via vacuum filtration to afford title compound (342 mg, 70 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.51 (s, 1H, SH), 12.51 (s, 1H, NH), 7.15-7.13 (m, 2H, Ar), 7.12-7.10 (m, 2H, Ar).

*Synthesis of 5-chloro-1H-benzo[d]imidazole-2-thiol (**7**)*



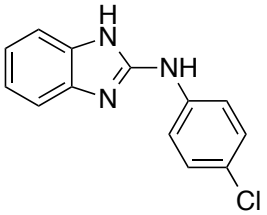
4-Chloro-1,2-phenylenediamine (1 g, 7.01 mmol) was dissolved in a mixture of (EtOH: H_2O = 6: 1) (30 mL). Potassium ethyl xanthogenate (1.6 g, 10.51 mmol) was added to the stirred solution and solution was refluxed for 16 h at 75 °C. After the 16 h, the reaction was allowed to cool to rt and was added to cold H_2O (stirred) drop-wise. A reddish precipitate (533 mg) was afforded and collected via vacuum filtration. The resultant filtrate was acidified using (AcOH: H_2O = 1: 1) (6 mL) yielding title compound (594 mg). In total, two crops of compound were extracted (1.1 g, 86 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 12.68 (s, 1H, SH), 12.66 (s, 1H, NH), 7.16-7.11 (m, 3H, Ar).

*Synthesis of N-phenyl-1H-benzo[d]imidazol-2-amine (**8**)*



o-Phenylenediamine (250 mg, 2.31 mmol) was dissolved in THF (3 mL). Once completely dissolved, phenyl isothiocyanate (344 mg, 304 µL, 2.54 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (887 mg, 4.62 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction mixture was added to cold H_2O drop-wise to yield a brownish amorphous solid. The resultant solid was dissolved in a mixture of EtOAc (2 x 20 mL) and H_2O (10 mL) and extracted. The organic layers were combined, washed with brine (10 mL), dried over $MgSO_4$ and reduced *in vacuo*. The afforded brown solid (578 mg) was purified via column chromatography using (petroleum spirits: EtOAc = 2: 1) to give title compound (261 mg, 54 %). R_f: 0.4. ^{1}H NMR (400 MHz; DMSO-d_6): δ 10.88 (s, 1H, NH), 9.38 (s, 1H, NH), 7.77-7.74 (m, 2H, Ar), 7.35-7.27 (m, 4H, Ar), 6.98 (dd, 2H, J = 8.7 Hz, 6.5 Hz, Ar), 6.94-6.90 (m, 1H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 150.5, 140.8, 132.9, 128.7, 120.5, 120.0, 117.0, 114.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 210.0955; Observed $[M+H]^+$ = 210.1027.

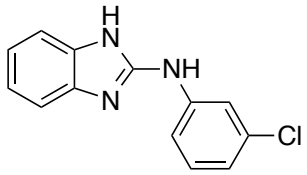
*Synthesis of N-(4-chlorophenyl)-1H-benzo[d]imidazol-2-amine (**9**)*



o-Phenylenediamine (250 mg, 2.31 mmol) was dissolved in THF (3 mL). Once completely dissolved, 4-chlorophenyl isothiocyanate (432 mg, 2.54 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (887 mg, 4.62 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction mixture was added to cold H_2O drop-wise to yield a light

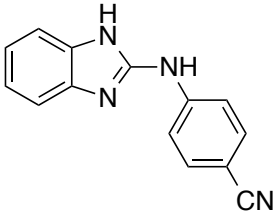
green-coloured oil. The resultant oil was dissolved in a mixture of EtOAc (2 x 20 mL) and H_2O (10 mL) and extracted. The organic layers were combined, washed with brine (10 mL), dried over $MgSO_4$ and reduced *in vacuo*. The afforded brown solid (639 mg) was recrystallized using DCM to give title compound (337 mg, 60 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 10.97 (s, 1H, NH), 9.58 (s, 1H, NH), 7.82-7.80 (m, 2H, Ar), 7.34 (d, 4H, *J* = 9.0 Hz, Ar), 6.99 (dd, 2H, *J* = 5.8 Hz, 3.1 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 150.1, 139.8, 128.5, 123.8, 120.4, 118.4, 115.9, 109.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 244.0565; Observed $[M+H]^+$ = 244.0639.

*Synthesis of N-(3-chlorophenyl)-1H-benzo[d]imidazol-2-amine (**10**)*



o-Phenylenediamine (250 mg, 2.31 mmol) was dissolved in THF (3 mL). Once completely dissolved, 3-chlorophenyl isothiocyanate (432 mg, 334 µL, 2.54 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (887 mg, 4.62 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction mixture was added to cold H_2O drop-wise yielding a brownish amorphous solid (673 mg). The solid was collected via vacuum filtration and recrystallized using DCM affording title compound (312 mg, 56 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.04 (s, 1H, NH), 9.67 (s, 1H, NH), 8.09 (t, 1H, *J* = 2.0 Hz, Ar), 7.56 (ddd, 1H, *J* = 8.2 Hz, 2.1 Hz, 0.9 Hz, Ar), 7.31 (t , 3H, *J* = 8.1 Hz, Ar), 7.03-6.99 (m, 2H, Ar), 6.95 (ddd, 1H, *J* = 7.9 Hz, 2.0 Hz, 0.9 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 149.8, 142.6, 142.4, 133.2, 130.3, 120.5, 120.0, 116.2, 115.4, 109.6. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 244.0565; Observed $[M+H]^+$ = 244.0639.

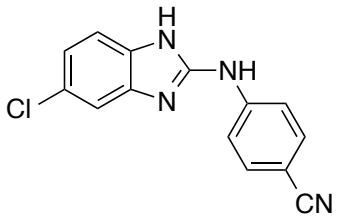
*Synthesis of 4-((1H-benzo[d]imidazol-2-yl)amino)benzonitrile (**11**)*



o-Phenylenediamine (500 mg, 4.62 mmol) was dissolved in THF (6 mL). Once completely dissolved, 4-cyanophenyl isothiocyanate (815 mg, 5.08 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (1.7 g, 9.25 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction mixture was added to cold H_2O drop-wise to yield a tan powder (962 mg), the solid was collected via vacuum filtration. The resultant solid was suspended in ether (10 mL) and filtered via vacuum filtration to afford title compound (449 mg, 41 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.19 (s, 1H, NH), 10.08 (s, 1H,

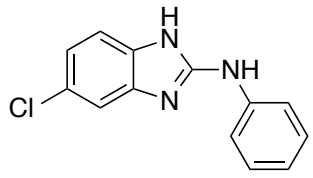
NH), 7.96 (d, 2H, *J* = 8.9 Hz, Ar), 7.77-7.74 (d, 2H, *J* = 8.9 Hz, Ar), 7.37 (d, 2H, *J* = 26.7 Hz, Ar), 7.04 (dd, 2H, *J* = 5.9 Hz, 2.8 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 149.1, 145.0, 133.2, 120.7, 120.4, 119.7, 116.8, 109.9, 101.5. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 235.0902; Observed $[M+H]^+$ = 235.0976.

*Synthesis of 4-((5-chloro-1H-benzo[d]imidazol-2-yl)amino)benzonitrile (**12**)*



4-Chloro-1,2-phenylenediamine (500 mg, 3.50 mmol) was dissolved in THF (6 mL). Once completely dissolved, 4-cyanophenyl isothiocyanate (618 mg, 3.85 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (1.3 g, 7.01 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction mixture was added to cold H_2O drop-wise to yield a purple solid. The resultant solid was dissolved using EtOAc (30 mL). The organic layer was combined with the aqueous layer and extracted. The organic layer was washed using brine (15 mL), dried over $MgSO_4$ and reduced *in vacuo* to yield title compound (777 mg, 82 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.32 (d, 1H, *J* = 20.7 Hz, NH), 10.21 (d, 1H, *J* = 9.6 Hz, NH), 7.94 (d, 2H, *J* = 8.8 Hz, Ar), 7.76 (d, 2H, *J* = 8.9 Hz, Ar), 7.45-7.32 (m, 2H, Ar), 7.06 (d, 1H, *J* = 8.3 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 144.6, 133.2, 120.8, 120.2, 119.5, 117.3, 117.0, 115.8, 110.9, 109.8, 101.9. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 269.0509; Observed $[M+H]^+$ = 269.0582.

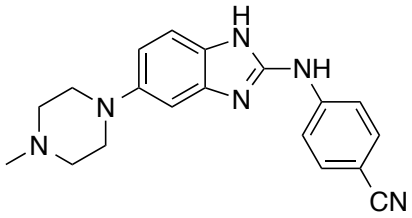
*Synthesis of 5-chloro-N-phenyl-1H-benzo[d]imidazol-2-amine (**13**)*



4-Chloro-1,2-phenylenediamine (250 mg, 1.75 mmol) was dissolved in THF (3 mL). Once completely dissolved, phenyl isothiocyanate (261 mg, 230 µL, 1.92 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (672 mg, 3.50 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction was diluted using H_2O (50 mL) and extracted using EtOAc (2 x 30 mL). The organic layers were combined and washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant solid was recrystallized using DCM to afford title compound (209 mg, 49 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.01 (s, 1H, NH), 9.54 (s, 1H, NH), 7.74-7.72 (m, 2H, Ar), 7.33-7.29 (m, 4H, Ar), 6.99 (dd, 1H, *J* = 8.3 Hz, 1.8 Hz, Ar), 6.96-6.92 (m, 1H, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 151.8, 151.3, 144.3, 140.4,

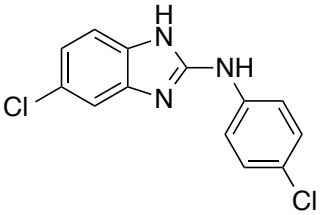
131.6, 128.8, 124.6, 120.9, 119.3, 117.3, 115.2. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 244.0563; Observed $[M+H]^+$ = 244.0637.

*Syntheis of 4-((5-(4-methylpiperazin-1-yl)-1H-benzo[d]imidazol-2-yl)amino)benzonitrile (**14**)*



To a solution of **3** (382 mg, 1.85 mmol) in THF (6 mL), EDC.HCl (711 mg, 3.71 mmol) was added and allowed to dissolve at rt. 4-cyanophenyl isothiocyanate (327 mg, 2.04 mmol) was added to the solution and the mixture was sealed in a microwave vial and heated for 16 h at 75 °C. After 16 h, the reaction mixture was diluted using H_2O (100 mL) and extracted using EtOAc (2 x 20 mL). The organic layers were combined and washed with brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The afforded red-brown solid (425 mg) was recrystallized in DCM to yield title compound (200 mg, 33 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 10.94 (s, 1H, NH), 9.98 (d, 1H, *J* = 16.3 Hz, NH), 7.90 (d, 2H, *J* = 8.2 Hz, Ar), 7.72 (d, 2H, *J* = 8.6 Hz, Ar), 7.28-7.16 (m, 1H, Ar), 6.99-6.88 (m, 1H, Ar), 6.77-6.75 (m, 1H, Ar), 3.05 (t, 4H, *J* = 4.7 Hz, CH_2), 2.50-2.46 (m, 4H, CH_2), 2.22 (s, 3H, CH_3); ^{13}C NMR (101 MHz; DMSO-d_6): δ 171.0, 167.6, 165.2, 165.1, 156.8, 145.1, 133.2, 116.7, 116.5, 116.5, 116.4, 98.0, 54.9, 50.3, 50.1, 45.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 333.1748; Observed $[M+H]^+$ = 333.1818.

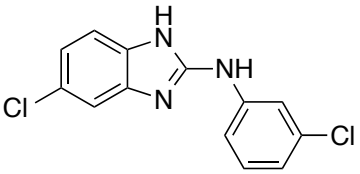
*Synthesis of 5-chloro-N-(4-chlorophenyl)-1H-benzo[d]imidazol-2-amine (**15**)*



4-Chloro-1,2-phenylenediamine (250 mg, 1.75 mmol) was dissolved in THF (3 mL). Once completely dissolved, 4-chlorophenyl isothiocyanate (327 mg, 1.92 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (672 mg, 3.50 mmol). The vial was sealed and heated for 16 h at 80 °C. Once complete, the reaction was diluted using H_2O (50 mL) and extracted using EtOAc (2 x 30 mL). The organic layers were combined and washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant reddish amorphous solid (582 mg) was recrystallized in DCM to afford title compound (248 mg, 51 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.11 (s, 1H, NH), 9.72 (s, 1H, NH), 7.79 (d, 2H, *J* = 8.9 Hz, Ar), 7.38-7.34 (m, 2H, Ar), 7.34-7.31 (m, 2H, Ar), 7.00 (d, 1H, *J* = 7.8 Hz, Ar); ^{13}C NMR (101

MHz; DMSO-d_6): δ 151.4, 150.9, 139.4, 133.6, 131.6, 128.6, 124.4, 120.4, 118.7, 116.8, 115.4. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 278.0168; Observed $[M+H]^+$ = 278.024.

*Synthesis of 5-chloro-N-(3-chlorophenyl)-1H-benzo[d]imidazol-2-amine (**16**)*



4-Chloro-1,2-phenylenediamine (250 mg, 1.75 mmol) was dissolved in THF (3 mL). Once completely dissolved, phenyl isothiocyanate (327 mg, 253 µL, 1.92 mmol) was added drop-wise to the reaction mixture followed by the addition of EDC.HCl (672 mg, 3.50 mmol). The vial was sealed and heated for 16 h at 75 °C. Once complete, the reaction was diluted using H_2O (50 mL) and extracted using EtOAc (2 x 30 mL). The organic layers were combined and washed using brine (20 mL), dried over $MgSO_4$ and reduced *in vacuo*. The resultant solid was purified via column chromatography using (petroleum spirits: EtOAc = 2: 1) to afford a golden oil (114 mg, 23 %). ^{1}H NMR (400 MHz; DMSO-d_6): δ 11.22 (s, 1H, NH), 9.81 (s, 1H, NH), 8.04 (t, 1H, *J* = 2.0 Hz, Ar), 7.55 (ddd, 1H, *J* = 8.2 Hz, 2.1 Hz, 0.9 Hz, Ar), 7.37 (s, 1H, Ar), 7.32 (t, 2H, *J* = 8.1 Hz, Ar), 7.02 (dd, 1H, *J* = 8.3 Hz, 2.0 Hz, Ar), 6.98 (ddd, 1H, *J* = 7.9 Hz, 2.0 Hz, 0.9 Hz, Ar); ^{13}C NMR (101 MHz; DMSO-d_6): δ 143.7, 141.2, 139.8, 136.4, 135.6, 131.0, 129.7, 123.3, 122.4, 115.2, 114.5, 112.0, 111.7. *m/z* (ESI-HRMS) Calc $[M+H]^+$ = 278.0168; Observed $[M+H]^+$ = 278.0236.