



## Open Source Drug Discovery: Highly Potent Antimalarial Compounds Derived from the Tres Cantos Arylpyrroles

Alice E. Williamson,<sup>†,‡</sup> Paul M. Ylloja,<sup>†,‡</sup> Murray N. Robertson,<sup>†,‡</sup> Yevgeniya Antonova-Koch,<sup>§</sup> Vicky Avery,<sup>||</sup> Jonathan B. Baell,<sup>⊥</sup> Harikrishna Batchu,<sup>#</sup> Sanjay Batra,<sup>#</sup> Jeremy N. Burrows,<sup>¶</sup> Soumya Bhattacharyya,<sup>#</sup> Felix Calderon,<sup>α</sup> Susan A. Charman,<sup>⊥</sup> Julie Clark,<sup>β</sup> Benigno Crespo,<sup>α</sup> Martin Dean,<sup>†</sup> Stefan L. Debbert,<sup>γ</sup> Michael Delves,<sup>δ</sup> Adelaide S. M. Dennis,<sup>ε</sup> Frederik Deroose,<sup>ζ</sup> Sandra Duffy,<sup>||</sup> Sabine Fletcher,<sup>||</sup> Guri Giaever,<sup>η</sup> Irene Hallyburton,<sup>θ</sup> Francisco-Javier Gamo,<sup>α</sup> Marinella Gebbia,<sup>η</sup> R. Kiplin Guy,<sup>β</sup> Zoe Hungerford,<sup>†</sup> Kiaran Kirk,<sup>ε</sup> Maria J. Lafuente-Monasterio,<sup>α</sup> Anna Lee,<sup>η</sup> Stephan Meister,<sup>§</sup> Corey Nislow,<sup>η</sup> John P. Overington,<sup>ι</sup> George Papadatos,<sup>ι</sup> Luc Patiny,<sup>κ</sup> James Pham,<sup>λ</sup> Stuart A. Ralph,<sup>λ</sup> Andrea Ruecker,<sup>δ</sup> Eileen Ryan,<sup>⊥</sup> Christopher Southan,<sup>μ</sup> Kunkum Srivastava,<sup>#</sup> Chris Swain,<sup>ν</sup> Matthew J. Tarnowski,<sup>†</sup> Patrick Thomson,<sup>ξ</sup> Peter Turner,<sup>†</sup> Iain M. Wallace,<sup>ι,ο</sup> Timothy N. C. Wells,<sup>¶</sup> Karen White,<sup>⊥</sup> Laura White,<sup>†</sup> Paul Willis,<sup>¶</sup> Elizabeth A. Winzeler,<sup>§</sup> Sergio Wittlin,<sup>π</sup> and Matthew H. Todd<sup>\*,†,‡</sup>

<sup>†</sup>School of Chemistry, The University of Sydney, Sydney, New South Wales 2006, Australia

<sup>§</sup>Department of Pediatrics, Pharmacology & Drug Development, University of California San Diego, 9500 Gilman Drive, La Jolla, California 92093, United States

<sup>||</sup>Discovery Biology, Eskitis Institute for Drug Discovery, Griffith University, Nathan, Queensland 4111, Australia

<sup>⊥</sup>Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia

<sup>#</sup>CSIR-Central Drug Research Institute, Sector 10, Jankipuram Extension, Sitapur Road, Lucknow, 226 031, India

<sup>¶</sup>Medicines for Malaria Venture, PO Box 1826, 20 rte de Pre-Bois, 1215 Geneva 15, Switzerland

<sup>α</sup>Tres Cantos Medicines Development Campus, Diseases of the Developing World, GlaxoSmithKline, Severo Ochoa 2, 28760 Tres Cantos, Spain

<sup>β</sup>Department of Chemical Biology & Therapeutics, St. Jude Children's Research Hospital, MS 1000, Room E9050, 262 Danny Thomas Place, Memphis, Tennessee 38105-3678, United States

<sup>γ</sup>Department of Chemistry, Lawrence University, 233 Steitz Science Hall, 711 East Boldt Way, Appleton, Wisconsin 54911, United States

<sup>δ</sup>Department of Life Sciences, Imperial College London, South Kensington, London SW7 2AZ, U.K.

<sup>ε</sup>Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia

<sup>ζ</sup>Asclepia Outsourcing Solutions, Damvalleistraat 49, B-9070 Destelbergen, Belgium

<sup>η</sup>Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College Street, Toronto, Ontario M5S 3E1, Canada

<sup>θ</sup>Drug Discovery Unit, Division of Biological Chemistry and Drug Discovery, University of Dundee, Dundee, DD1 5EH, U.K.

<sup>ι</sup>European Molecular Biology Laboratory—European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SD, U.K.

<sup>κ</sup>Institute of Chemical Sciences and Engineering (ISIC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015, Switzerland

<sup>λ</sup>Department of Biochemistry & Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, Victoria 3010, Australia

<sup>μ</sup>IUPHAR/BPS Guide to PHARMACOLOGY, Centre for Integrative Physiology, School of Biomedical Sciences, University of Edinburgh, Edinburgh, EH8 9XD, U.K.

<sup>ν</sup>Cambridge MedChem Consulting, 8 Mangers Lane, Duxford, Cambridge CB22 4RN, U.K.

<sup>ξ</sup>School of Chemistry, The University of Edinburgh, Joseph Black Building, West Mains Road, Edinburgh EH9 3JJ, U.K.

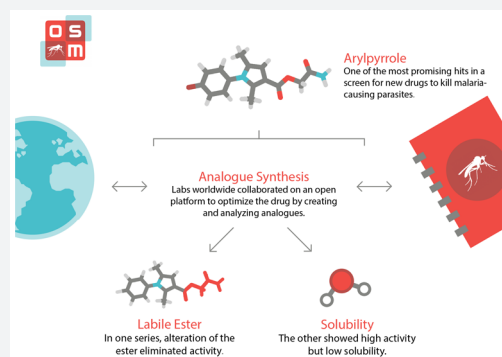
<sup>π</sup>Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland

Received: May 31, 2016

Published: September 14, 2016

## S Supporting Information

**ABSTRACT:** The development of new antimalarial compounds remains a pivotal part of the strategy for malaria elimination. Recent large-scale phenotypic screens have provided a wealth of potential starting points for hit-to-lead campaigns. One such public set is explored, employing an open source research mechanism in which all data and ideas were shared in real time, anyone was able to participate, and patents were not sought. One chemical subseries was found to exhibit oral activity but contained a labile ester that could not be replaced without loss of activity, and the original hit exhibited remarkable sensitivity to minor structural change. A second subseries displayed high potency, including activity within gametocyte and liver stage assays, but at the cost of low solubility. As an open source research project, unexplored avenues are clearly identified and may be explored further by the community; new findings may be cumulatively added to the present work.



## ■ INTRODUCTION

Malaria remains one of the world's most deadly diseases. There were an estimated 214 million cases of malaria in 2015, including around 438,000 deaths of which the majority, tragically, were young children.<sup>1</sup> Besides the threat to human health, there is significant economic and social impact on the affected communities with malaria costing Africa billions of dollars *per annum* in direct losses and even more when considering lost economic growth.<sup>2,3</sup>

The continual threat of drug resistance has led to the World Health Organization (WHO) recommending that all treatments should only be used in combination; artemisinin combination therapies (ACTs) comprising an artemisinin derivative and a 4-aminoquinoline or amino alcohol currently represent the front line. However, the inevitable reports of resistance or tolerance, in the form of increased parasite clearance times, have already appeared.<sup>4–6</sup> Loss of the artemisinin class of drugs is a terrifying scenario that requires urgent risk mitigation. Apart from the introduction of the ACTs, no viable new drug for malaria has entered the market in the past 15 years, and the recent results of the Mosquirix vaccine phase III trials showed 18–36% efficacy depending on patient age and other factors.<sup>7,8</sup> New chemical series that can replace and complement the ACTs are urgently needed and are being sought by a combination of academic and industrial groups, sometimes in collaboration with non-governmental organizations (NGOs).<sup>9–11</sup> Of particular interest are lead candidates with differentiated activity profiles, ideally targeting gametocyte or liver stage parasites in addition to blood stages.<sup>12–15</sup>

The ability of the pharmaceutical industry to provide new medicines cost-effectively is diminishing.<sup>16</sup> The industry acknowledges that lack of innovation is a problem.<sup>17</sup> Pharma is responsible for the creation of most marketed drugs, yet many of these are arguably not innovative; in contrast academia and the biotech industry generate more innovative leads, but many are orphan drugs.<sup>18</sup> Such challenges disproportionately affect research into new medicines for tropical diseases, which would inevitably generate a slim profit margin unlikely to recoup the necessary expenses of research and development.<sup>19</sup>

The weak status of many drug development pipelines in the pharmaceutical industry is driving the exploration of alternative models. The current model of drug discovery, whether in academia<sup>20</sup> or industry, can generally be characterized by secrecy and an underlying profit motive.<sup>21</sup> In the area of tropical diseases there are significant philanthropic efforts being

made by many companies in providing treatments<sup>22</sup> and engaging in drug development<sup>23</sup> but also in conducting not-for-profit research.<sup>24,25</sup> There have been calls for greater sharing of data in the NTD field,<sup>26,27</sup> including the development of patent pools<sup>28</sup> and new Product Development Partnerships.<sup>29</sup> Repositioning of existing drugs is seen as a possible general strategy for the development of new antimalarials, even though the challenges of such an approach are clear.<sup>30</sup> There has been much recent discussion of the need for “Open Innovation”, a term with a nebulous definition but typically describing a range of ideas from the sharing of data in a precompetitive environment through to competitions that allow organizations to bring in the best external ideas to complement in-house research but for which there is no requirement for any collaboration.<sup>31</sup>

A model that has been mooted,<sup>32–38</sup> but never properly implemented and evaluated, is drug discovery and development where all data and ideas are freely shared, there are no barriers to participation, and there are no patents—so-called “Open Source” Drug Discovery. The requirement for total sharing of data as well as workflows in drug discovery and development (i.e., the experimental science, as opposed to the software used in the project<sup>39</sup>) would mirror the same practices in open source software development—a model that has created robust and successful products in widespread use and formed the foundation of major industries as well as spawning for-profit open source software companies. It was shown recently that the opening up of a laboratory-based chemical research project to unrestricted participation by anyone accelerated the research because experts unknown to the core team were able to join the project and solve transient project needs.<sup>36</sup> That project involved the discovery of a new synthetic route to a known compound, specifically the active enantiomer of the drug of choice for the treatment of schistosomiasis, praziquantel.<sup>40</sup> The project benefited from the open sharing of chemical data and procedures on the Internet, i.e., using the Internet as a medium that facilitated collaboration and peer review where participants could influence the direction of the research before it occurred, rather than use the online content merely as an information resource. Several other initiatives have leveraged the advantage of open online discussion of chemical data and results generated by others.<sup>41–43</sup>

The present work extends this idea to the identification of novel bioactive compounds. A previous demonstration of an open drug discovery cycle was shown by the Usefulchem project,

- 1<sup>st</sup> Law: All data are open and all ideas are shared  
 2<sup>nd</sup> Law: Anyone can take part at any level  
 3<sup>rd</sup> Law: There will be no patents  
 4<sup>th</sup> Law: Suggestions are the best form of criticism  
 5<sup>th</sup> Law: Public discussion is much more valuable than private email  
 6<sup>th</sup> Law: An open project is bigger than, and is not owned by, any given lab

Figure 1. Six laws of open research governing the present project.

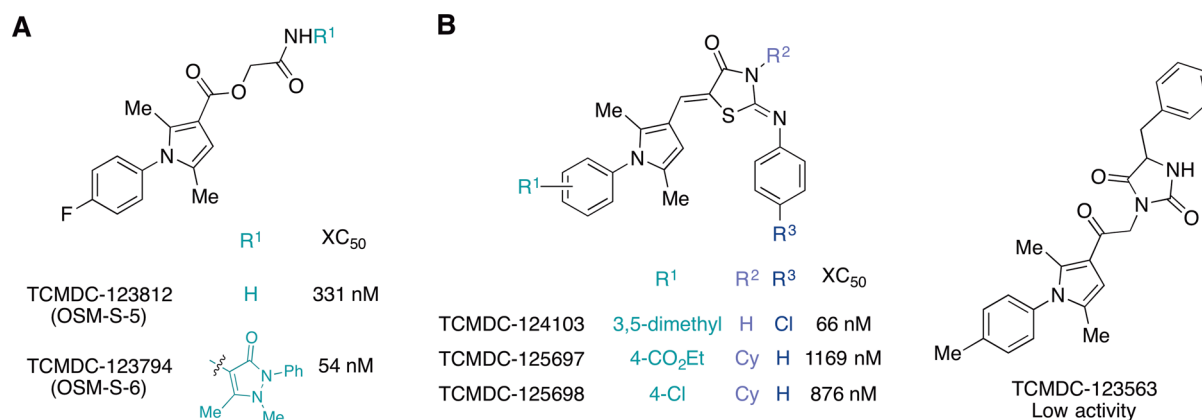


Figure 2. (A) Arylpyrrole hits from the TCAMS data set. (B) Selected “near neighbors” (NN) from the TCAMS data set. Activities quoted are vs 3D7. Not shown: related NN compounds TCMDC-124456 (680 nM) and TCMDC-125659 (237 nM). (XC<sub>50</sub> = approximate IC<sub>50</sub> value.<sup>53</sup>)

which found four micromolar hits from a small product library after *in silico* target prediction and docking.<sup>44</sup> The Open Source Drug Discovery project in India has carried out a crowdsourcing project for annotation of the *Mycobacterium tuberculosis* genome.<sup>45–47</sup> In the field of biotechnology the CAMBIA organization used patents to enforce a code of conduct based on the open sharing of technologies, in that experimental tools could be freely used, provided no further patents were taken to restrict the use of those tools by others.<sup>48</sup>

The “open source” moniker is not merely semantic and distinguishes such projects from other “open” ventures in several important regards,<sup>38</sup> described by the six laws that governed the operation of the present project (Figure 1, Figure S1).<sup>49</sup> Crucially, the research *process* (i.e., strategic discussions, issues involving doubt) takes place in the public domain. The Creative Commons license covering project content ensures free reuse of all content including for commercial purposes (CC-BY),<sup>50</sup> and uses a free or open source composite technical platform that has recently been reviewed.<sup>51,52</sup>

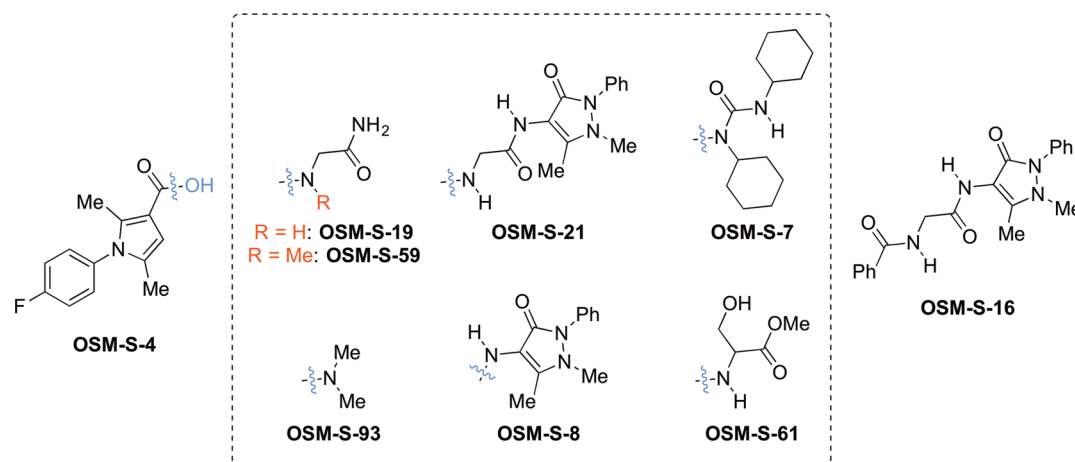
In recent years, teams led by Novartis, St. Jude’s Children’s Research Hospital, and GlaxoSmithKline (GSK) Tres Cantos have released large data sets of antimalarial compounds derived from phenotypic high throughput screening (HTS) campaigns.<sup>53–56</sup> The GSK Tres Cantos Antimalarial (TCAMS) data set contained 13,533 filtered hits which were subsequently prioritized and grouped to provide numerous starting points for other research groups;<sup>57</sup> all of these 47 scaffolds have been explored either internally by GSK, by independent groups, or in collaboration between GSK and academic groups. GSK themselves have published evaluations of two of these priority series, the cyclopropyl carboxamides<sup>58,59</sup> and an indoline series.<sup>60</sup> No further optimization work has been performed by GSK on these two series due to the inherent risks identified. The original GSK data set was used to identify starting points for the present campaign, resulting in the selection of TCMDC-123812 (OSM-S-5, Figure 2A) and its 4-aminoantipyrene

derivative TCMDC-123794 (OSM-S-6) because of their attractive physicochemical properties, such as low logP and molecular weight, coupled with promising bioactivity and therefore presumed high ligand efficiency. (Compound numbering in this paper is based on the original project numbering, rather than a renumbering for this paper, allowing simpler cross-correlation between this paper and the live Web sites. The convention used is Open Source Malaria (OSM)-first letter of city in which compound was first made (e.g., S = Sydney)-incremental integer. Batch numbers are included in internal project numbering schemes.) A number of compounds in the original TCAMS set<sup>53</sup> featured the arylpyrrole core with alternative head groups lacking the ester linkage (Figure 2B). TCMDC-123563 was discounted (Figure S2)<sup>61</sup> as it represented a singleton and contained an unfavorable ketone linkage. A cluster of related compounds, the 2-iminothiazolidin-4-ones (the “near neighbors”, NN), was targeted because the members (including TCMDC-124103, -125697, and -125698) possessed promising activities without the ester, and indicated tolerance to variation elsewhere in the structure. While this work was being written up for publication, Gilbert et al. published details of a series of pyrrolones (identified through an unpublished screen performed by the World Health Organization Special Programme for Research and Training in Tropical Diseases (WHO-TDR) but also identified in the Novartis antimalarial data set<sup>55</sup>) that have some structural similarities to the NN compounds (i.e., an arylpyrrole joined through a double bond to a different heterocycle), and comparisons will be drawn below with this series.<sup>62,63</sup>

It was important to resynthesize the original hit compounds (Figure 2A) to confirm their activity with authentic samples prior to beginning a hit to lead campaign where the NN compounds would act as a potential backup subseries.

## RESULTS AND DISCUSSION

**Resynthesis of Original Hits and Their Putative Prodrug Fragments.** The synthetic approaches to any molecule



**Figure 3.** Synthesized amide analogues. Synthetic details and structures of purchased compounds may be found in the [Supporting Information](#). All compounds were found to possess low activity ([Table SB2](#)).

in this paper may be found in full in the relevant electronic laboratory notebook<sup>64</sup> and are summarized in the [Supporting Information](#) (Chemical Protocols; figures with the prefix “SC” may be found therein).

The two original GSK hits (OSM-S-5 and -6) were successfully resynthesized via a novel pyrrole acid (OSM-S-4, [Figure SC1](#)) that was prepared via a Paal–Knorr cyclization of the relevant aniline and ethyl 2-acetyl-4-oxopentanoate.<sup>65,66</sup> This approach was found to be superior to an alternative method involving initial synthesis of the unfunctionalized *N*-arylpyrrole, followed by conversion to the corresponding aldehyde with a Vilsmeier–Haack reaction (a procedure that was improved through a community suggestion ([Figure S3](#))<sup>67</sup>) and then oxidation to a carboxylic acid, because the pyrrole aldehyde was found to be remarkably resilient to a range of oxidants. (An alternative route using a Friedel–Crafts acylation between the unsubstituted pyrrole and ethyl chloroformate, suggested in an e-mail from the community, gave only starting material in two attempts.)

To confirm the promise of the two starting points OSM-S-5 and -6, they were evaluated against 3D7 (drug-sensitive) and K1 (chloroquine resistant) strains of *Plasmodium falciparum* in a whole cell assay and against HEK-293 cells as a cytotoxicity marker ([Table SB1](#), Biological Protocols; tables with the prefix “SB” may be found in the [Supporting Information](#)). Biological data may be browsed in a static data set taken as a snapshot for this paper ([Data Sets S1](#) (Excel), [S2](#) (SDF)), online in a database constructed through periodic batch uploads,<sup>68</sup> or in a “living database” to which may be added future data;<sup>69</sup> the latter may be visualized in a web browser<sup>70</sup> using an open source system that was recently deployed in the Wikipedia Chemical Structure Explorer,<sup>71</sup> or the data may be downloaded and visualized offline with proprietary ([Figure S4](#))<sup>72</sup> or open source ([Figure S5](#))<sup>73</sup> tools). The evaluation was performed in three different institutions using different assays and widely employed controls; controls are important to assess reproducibility and interassay variability<sup>74</sup> but also to minimize possible bias in evaluating compounds where there are pre-existing data from other researchers already in the public domain ([Figure S6](#)).<sup>75</sup> The tests confirmed that the original compounds TCMDC-123812 (OSM-S-5) and TCMDC-123794 (OSM-S-6) are potent (300–500 nM range), although slightly less than previously reported (IC<sub>50</sub> of ca. 330 and 54 nM respectively from a colorimetric LDH assay over 72 h<sup>53</sup>) with low associated cytotoxicities

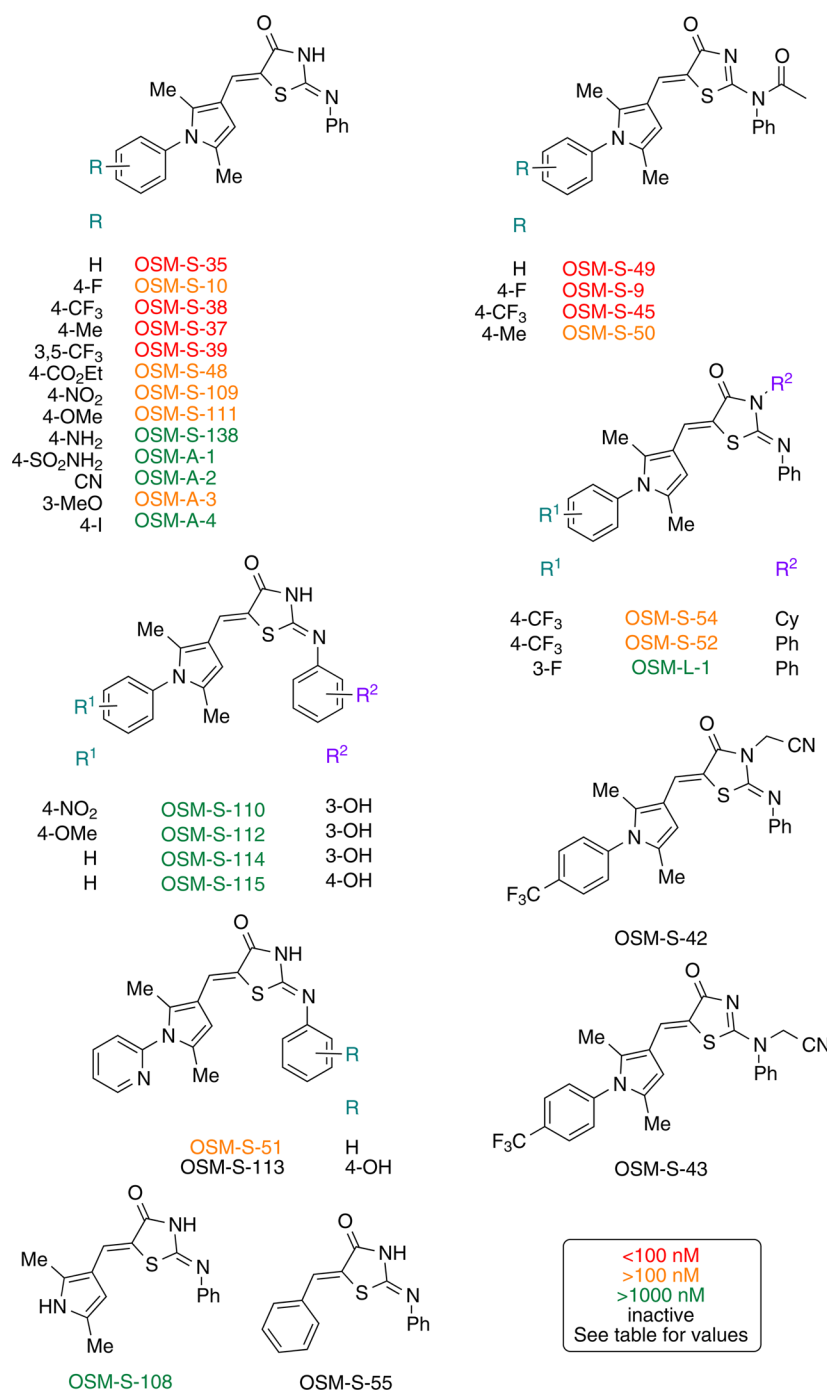
and similar efficacy against 3D7 and K1 strains. The stability of the ester linker under biologically relevant conditions was expected to be poor, so the aldehyde, ethyl ester, and carboxylic acid 4-fluoropyrroles made en route to these compounds were evaluated as potentially active fragments but were found to exhibit relatively low activity, suggesting that the parent compounds do not act as prodrugs in this way (similar inactivity was observed for the 4-H, 4-Me, and 4-CF<sub>3</sub> aldehydes, esters, and acid fragments) ([Data Sets S3–S5](#)).

**Amide Analogues of the Original Hits.** Replacement of the ester with a hydrolytically more stable amide was undertaken through synthesis of eight derivatives ([Figure 3](#)), most of which were obtained through SOCl<sub>2</sub>-mediated conversion of acid OSM-S-4 to the acid chloride. Compound OSM-S-16 served as a control lacking the pyrrole moiety, OSM-S-8 served as a truncated “des-glycinyl” analogue, and the importance of amide methylation was explored with compounds OSM-S-59 and -93. In addition, the six most relevant commercially available compounds were purchased ([Figure SC2](#)). (For more on strategies for selecting compounds for commercial acquisition, colloquially known as “SAR by catalog”, see the [Supporting Information](#) ([Text S2](#) and the component files referred to therein).)

The conversion of ester to amide resulted in essentially total loss of biological activity in all cases ([Table SB2](#)), even for the analogues involving minimal changes (OSM-S-19 and OSM-S-21), so the amide series was not explored further.

**The Alternative “Near Neighbor” 2-Iminothiazolidinones.** In parallel with the initial evaluation of the amide analogues, a number of NN analogues were synthesized ([Figure 4](#)), typically from a Vilsmeier–Haack oxidation of the relevant pyrrole to the corresponding aldehyde and then condensation with the appropriate 2-iminothiazolidin-4-one ([Figure SC3](#)).<sup>76–78</sup> While double bond geometry was undefined for the original hits, *Z*-geometry was established here by X-ray crystallography on four compounds and therefore assumed more broadly for the series ([Figure SC4](#)). Given the low predicted solubility of these NN compounds, several were generated with lower clogP values and submitted for biological testing (OSM-S-109 through -115, OSM-S-108, and OSM-S-138), alongside analogues synthesized and submitted by an independent undergraduate laboratory cohort: OSM-A-1 through -4 and resyntheses of OSM-S-37 and -111.

Many compounds in the NN series showed high potencies (shown schematically in [Figure 4](#); raw data in [Table SB3](#)), with



**Figure 4.** Near neighbor compounds biologically evaluated. Raw data may be found in [Table SB3](#).

several found to be more active than the original TCAMS hits. The compounds exhibited low associated cytotoxicity. The aryl component of the pyrrole moiety was moderately tolerant to changes (though small changes could result in large changes in activity, exemplified by the 4-F and 3-F isomers), while the thiazolidinone component was found to be more sensitive. Incorporation of cyclopentyl, phenyl, and acetyl components was tolerated, but the methylenitrile group was not. Replacement of the arylpyrrole moiety with a phenyl resulted in loss of activity.

While high potency was obtained with several members of the NN series, this was achieved at the cost of high calculated lipophilicity with many compounds in this series exhibiting

calculated logP values of 5 or more and correspondingly poor lipophilic efficiency ([Figures S15 and S16](#)).<sup>79</sup> Solubility was a challenge in several of the assays examining these compounds. Those compounds with more polar groups on the constituent aromatic rings suffered a drop in potency, but the pyridinyl analogue OSM-S-51 (calculated LogP = 1.8) provides a possible future line of inquiry for the community. A lipophilicity/potency trend was also generally seen in the pyrrolone series,<sup>62</sup> though potency was seen for several compounds containing a substituted piperidine ring in place of the *N*-aryl group.<sup>63</sup>

The open nature of the project enabled regular consultation with the wider medicinal chemistry community in real time, i.e., where community input could influence the direction of the

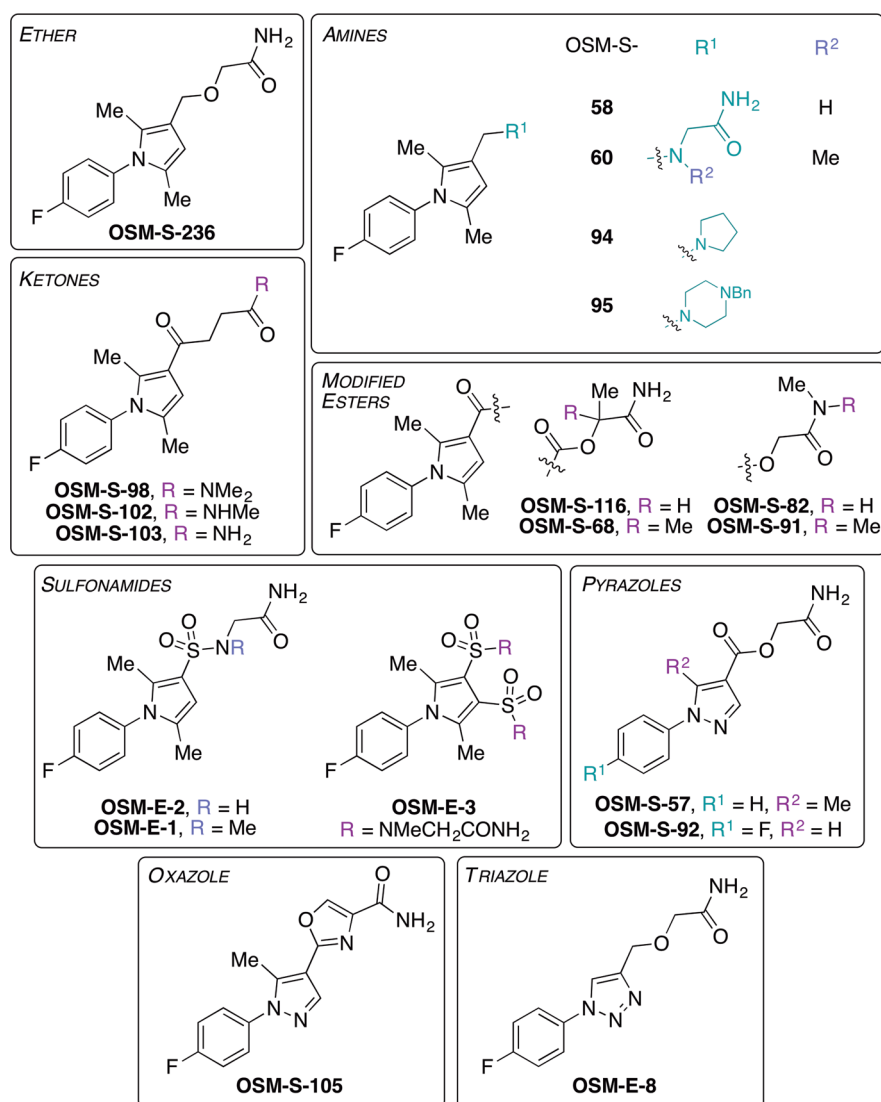


Figure 5. Analogues evaluated in the main arylpyrrole series.

research. An important contribution to the project from outside the core experimental team (Figure S17)<sup>80</sup> was discussion of whether the most potent members of the NN series were pan assay interference compounds (PAINS), i.e., compounds frequently appearing as hits in high throughput screens yet which do not exhibit a straightforward “drug-like” interaction with a biological target.<sup>81</sup> The OSM compounds were run through the KNIME PAINS filter,<sup>82,83</sup> and both the 2-imino-4-thiazolidinone and arylpyrrole components of these compounds were flagged as potential PAINS, with the proposed cause of the interference of the former being the thiazolidinone *exo*-double bond acting as a Michael acceptor, though it has been noted that this core is much more problematic in rhodanines than in their 2-imino counterparts.<sup>84</sup> The topic of PAINS has been the subject of extensive recent discussion in papers<sup>84–88</sup> and in online communities (Figures S18 and S19).<sup>89,90</sup> Although most concern has centered on rhodanines, any related structure could be problematic if it contains a potentially reactive conjugated *exo*-double bond. In the area of chemotherapeutic and antiparasitic agents, such motifs may still be present in viable leads.<sup>84</sup> The negative view of rhodanine derivatives in the medicinal chemistry community is generally derived from academic reports and patents where positive assay

hits have been reported without adequate evaluation of SAR or elucidation of the mode of action. A complicating factor, often not considered, is that PAINS are defined on the basis of results from target-based screens, where one would not link a cellular readout to the target if nonspecific protein reactivity were possible.<sup>84</sup> Conversely, a covalent modifier from a cellular screen may still be useful as a probe under some circumstances.<sup>91</sup> The present compounds may not be problematically reactive and may still be progression candidates for the following reasons: (1) The parent hit compounds (TCMDC-123812 (OSM-S-5) and -123974 (OSM-S-6), Figure 2) were shown not to be “promiscuous” frequent hitters in the original GSK data;<sup>53</sup> (2) the assay data described above show that the relevant 2-imino-4-thiazolidinone fragment was inactive on its own (OSM-S-55, Table SB3, entry 30), and (3) preliminary experimental controls were performed to assess the reactivity of the *exo* double bond (Figure SCS5) that showed no reactivity of the exocyclic double bond to hydrogenation or the addition of hydride or a thiol in model cases. Yet it was noted that these compounds not only are closely related to known PAINS but also fail by ALARM NMR filtering, which is designed to detect known protein-reactive cores.<sup>92–94</sup> Ultimately, the authorship team adopted varying positions. Overall, given the

poor physical properties of these compounds and the extensive additional work that would be needed to fully mitigate the series risks, and encouraged by one of us (J.B.B.) to move away from PAINS-like structures, the team decided not to further pursue this subseries.

Poor solubility with excessive lipophilicity may not just impart poor pharmacokinetic properties but also drive non-specific protein reactivity through hydrophobic burial that may not be picked up in *in vitro* assays. The emphasis for this series moved toward analogues that promised improved solubility. As with all key strategic decision points in this open source project, discussion of possible structures took place in an online public consultation.<sup>95</sup> The meeting recalibrated the project focus with selection of the next synthetic compounds and agreement on which commercially available analogues to purchase and evaluate (Figure S24).<sup>96</sup> This community consultation confirmed ethers, amines, sulfonamides, oxadiazoles, and substituted ester analogues of the original arylpyrrole hits as the most valuable targets (Figure S25).<sup>97</sup> Synthesis of a shortlist of such analogues was planned and undertaken by whomever wished to do so (half of the ten top-ranked synthetic shortlist were ultimately made; synthetic planning assistance was received *gratis* from the private sector (Supporting Information, Texts S6 and S7), and the most relevant commercially available compounds were purchased and evaluated. GSK assessed the proposed compounds and confirmed (publicly) that none of the molecules had previously been evaluated for antimalarial activity as part of the TCAMS screen.

**Analogues in the Arylpyrrole Series.** Several analogue series of the original arylpyrroles were synthesized and evaluated (Figure 5, raw data in each case may be found in the Supporting Information).

**Ethers.** One of the most promising proposed analogues was the ether OSM-S-236. A number of approaches have failed to generate the desired product, attributed to the instability of either the arylpyrrole alcohol starting material OSM-S-11 (which decomposed on silica and when stored under inert conditions at 2 °C) or (if formed) the desired product itself (Figure SC6). The synthesis of this compound was abandoned, given that similar side reactivity may be seen *in vivo*, although such reactivity could be mitigated through the use of more electron-deficient pyrroles.

**Amines.** Four amine analogues (OSM-S-58, -60, -94, and -95) were successfully prepared by reductive alkylation of the relevant pyrrole aldehyde, and three additional amine analogues were purchased (OSM-S-88 to -90, Figure SC7). All were found to be inactive. The synthesized structures were found to be prone to decomposition under ambient conditions.

**Modified Esters.** In an attempt to decrease the hydrolytic lability of the ester, methyl groups were introduced adjacent to this functional group (OSM-S-116 and -68) (Figure SC8). To assess the influence of introducing methyl groups to the terminal amide, analogues OSM-S-82 and OSM-S-91 were purchased. All compounds were found to exhibit low potency (Table SB4), further suggesting that minor structural changes to the potent TCAMS hit compounds reduce activity.

**Ketones.** Friedel–Crafts acylation of a pyrrole precursor with succinic anhydride and subsequent amidation provided three ketone analogues (OSM-S-98, -102, and -103, Figure SC9) that were found to exhibit essentially zero potency, with only OSM-S-103 showing any activity at higher concentrations.

**Sulfonamides.** Three sulfonamide analogues (OSM-E-1 to -3) were synthetically derived directly from the relevant arylpyrrole (Figure SC10) and were also found to lack potency.

**Pyrazoles.** Two pyrazole analogues (OSM-S-57 and -92) were synthesized (Figure SC11) and evaluated to assess the impact of modifying the arylpyrrole core. Both compounds were found to be inactive; the fluoro analogue of OSM-S-57 was thus not synthesized. A related heterocycle alteration was also found to be detrimental to the literature pyrrolone series.<sup>62</sup>

**Oxazoles.** The oxazole analogue OSM-S-105 was prepared from the carboxylic acid of the corresponding arylpyrazole, though the analogous sequence for the parent arylpyrrole series could not be completed (Figure SC12). This compound and several of the synthetic precursors were evaluated and all found to be inactive.

**Triazoles.** The triazole analogue OSM-E-8 was synthesized via a Cu(I)-catalyzed cycloaddition reaction (Figure SC13) and evaluated alongside four synthetic precursors and side products, but all compounds were found to be inactive.

**Synthetic Threads That Remain Open.** Several synthetic targets for this series remain open (Figure 6) such as the ether

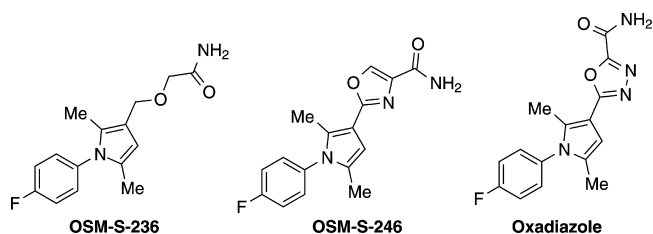


Figure 6. Three targets that remain open for synthetic inquiry.

compound OSM-S-236 (though it seems likely that this compound will be unstable) and the oxazole OSM-S-246. The oxadiazole shown was proposed, and preliminary experiments toward its synthesis were performed (see precursor OSM-S-269 in the project laboratory notebooks for further details).<sup>64</sup> The oxadiazole is a common replacement for carbonyl containing compounds in hit to lead campaigns,<sup>98</sup> and so it was reasoned that the inclusion of this heterocycle might have favorable consequences for drug metabolism, though a commercial oxadiazole analogue (OSM-S-85, Supporting Information) had been found to be inactive, leading to the synthetic effort toward the oxadiazole being downgraded. The tolerance of the NN set to the introduction of the pyridine in OSM-S-51 (Figure 4) could be explored further as a means to increase solubility in that cluster, and a pendant substituted piperidine was found to lead to several potent compounds in the pyrrolone series that possesses some structural similarities to the NN series.<sup>63</sup> Indeed variation of the aromatic group in the arylpyrrole series was not explored given the intractability of substituting the ester: given the tight SAR, low solubility, and poor metabolic stability observed for the series, the project viewed the probabilities of success as limited and so did not pursue these targets.

*In silico* pharmacophore modeling has to date proven ineffective at high-confidence analogue prediction, but this remains an open challenge<sup>99</sup> to which others may contribute given the data set available.<sup>68,69</sup> Some preliminary results suggesting a common feature map for the arylpyrrole and NN subseries were of particular interest and could be explored with the more substantial bioactivity data now available.

Along these lines, an automated bioisosteric replacement analysis<sup>100</sup> was performed for the NN series (Text S8, Data Set S19) focusing on replacement of the pyrrole phenyl substituent, and output suggestions were generated for potent compound OSM-S-35 (Figures S27–S30); these may be explored in the future.

All these structures are available for investigation by the community, building on the unsuccessful attempts detailed in the online laboratory notebooks. It is hoped that the sharing of negative synthetic data in this way (16 attempts in the case of the ether OSM-S-236) will lead to a faster completion of syntheses of analogues in the future since prior attempts are not “orphaned” in undisclosed or unpublished notebooks. Participants in future work may be physically located anywhere; they are requested (but not obligated) to operate open source (unrestricted sharing of all data and ideas) to avoid wasteful duplication of effort. The results, whether substantial or incremental, may be added to the series wiki.<sup>101</sup> However, it is important to acknowledge the limitations of the series identified to date, meaning further analogue synthesis undertaken by the community in the absence of better knowledge of the biological target is likely to be unproductive.

**Advanced Biological Evaluation.** To assess further the attractiveness of this class of compounds, attention was focused on the most potent of the newly discovered analogues and the original hits. Given that the compounds arose from a phenotypic assay, no mechanism of action (MoA) was known, and preliminary investigations described below were designed to probe this, with a view to minimizing any potential MoA overlap with other compounds already in development.

**Metabolic and Solubility Assays.** The two original GSK hits (OSM-S-5 and -6) and six NN compounds were evaluated for their kinetic solubility in phosphate buffer and their metabolic stability in human liver microsomes (Table SB5, Data Sets S20 and S21). Both the arylpyrrole esters showed good solubility, but showed degradation in microsomes even in the absence of cofactors for cytochrome P450 and glucuronidation enzymes (NADPH and UDPGA, respectively) suggesting degradation by nonspecific enzymes. The iminothiazolidinone (NN) compounds showed generally low rates of metabolic degradation but at a cost of very low solubility (a general feature also of the literature pyrrolone series<sup>62</sup> that displayed typically higher rates of metabolic clearance).

**Oral Efficacy in Mice.** Representative compounds from both subseries were evaluated in an oral *in vivo* mouse trial (Data Set S22). The original starting points OSM-S-5 and -6 (Figure 2) were chosen along with NN representative OSM-S-35 (Figure 4). All three compounds were found to be inactive *in vivo* in *Plasmodium berghei* ANKA infected mice at 50 mg/kg after 4 days po. It is plausible that at least some of the arylpyrrole ester compounds would be degraded by general hydrolysis during absorption, and this could explain their inactivity in the mouse model despite their more favorable (although still high) cLogP values. Analysis of the plasma samples from the trial with OSM-S-5 showed that the compound was indeed orally available, but concentrations in the blood were above the EC<sub>50</sub> for only approximately 4 h (Data Set S23). The same compound was evaluated for stability in human and mouse plasma and found to be susceptible to hydrolysis (*t*<sub>1/2</sub> 114 min), but it was stable in human plasma with no measurable loss after 240 min (Data Set S24). Esterase activity is higher in rodents than in other species,<sup>102</sup> as confirmed by using *p*-nitrophenol acetate as a control compound in this assay. By way of comparison, the literature pyrrolone series

had also exhibited low oral bioavailability that likely resulted from a combination of low solubility and significant metabolic clearance.<sup>62,63</sup>

A metabolite identification and glutathione trapping experiment was carried out on OSM-S-35 in the presence of metabolic activation (human microsomes, Data Set S25). A number of metabolites were detected, mainly oxygenated species (mono-, bis-, and trioxxygenated metabolites) with the predominant metabolite (based on peak area and assuming similar response factors for each metabolite) arising from likely hydroxylation of the pyrrole substituted benzene. In the presence of glutathione ethyl ester (GSH-EE), weak signals for adducts were observed for both parent compound and mono- and bisoxygenated metabolites. In the case of the parent compound, the GSH-EE adduct was observed both in the presence and in the absence of metabolic activation. Further characterization of the adducts was precluded by the very weak MS/MS spectra; however, the detection of adducts suggests that the formation of reactive species cannot be ruled out.

**hERG Liability.** There is increasing awareness of the importance for drug candidates of inhibition of hERG (the human *ether-à-go-go*-related gene ion channel), which is sensitive to blockade by many drug-like structures. Such blockage has led to a number of prominent postmarketing withdrawals.<sup>103</sup> Regulators are sensitive to any hERG activity, and a hERG counterscreen is often now run early in hit characterization and series prioritization. Compounds OSM-S-5 and OSM-S-35 were shown not to suffer from significant hERG activity (IC<sub>50</sub> > 33 μM vs 0.7 μM for control compound quinidine), implying that the original ester and NN class of compounds are not likely to exhibit undesirable cardiac side effects later in development (Data Set S26, Text S11).

**Late Stage Gametocyte Assay.** There are relatively few compounds effective against the gametocyte stage of the parasite,<sup>12,13</sup> though such compounds are important in the prevention of disease transmission. Four compounds found to be active in the asexual assay were evaluated against late stage (IV–V) gametocytes. The results (Table 1, Data Sets S27 and S28)

**Table 1. Potency vs Late Stage Gametocytes for Selected Compounds**

OSM-S-	late stage GAM <sup>a</sup> (nM)
5	75% at 60 μM
9	28
38	4
39	2.6
111 <sup>b</sup>	686

<sup>a</sup>Data Set S27. <sup>b</sup>Value for this compound taken from Data Set S28. Controls: see the Supporting Information.

indicated that the NN compounds OSM-S-38 and OSM-S-39 exhibited highly promising IC<sub>50</sub> values of 4 nM and 2.6 nM respectively, comparable to the activities of artemisinin and artesunate. Compound OSM-S-9 also showed good activity while one of the original hit compounds (OSM-S-5) exhibited low levels of activity. Several compounds evaluated were found to lead to an unusual parasite morphology that may be indicative of a slow-acting mechanism of action (Supporting Information, late stage gametocyte assays 1 and 2).

When the compounds were evaluated in a dual gamete formation assay (DGFA)<sup>104</sup> to evaluate separately the susceptibilities of male and female mature stage V gametocytes to both



the hit ester (OSM-S-5) and a selection of NN compounds, it was found that all compounds possessed low activity at 1  $\mu\text{M}$  against both sexes (Text S10). The discrepancy in the data arising from these two gametocyte assays may arise from the slightly different cell biology assayed (stage IV/V gametocytes vs stage V mature gametocytes) or most likely is a function of the relative compound exposure times (96 h vs 24 h).

**Liver Stage.** Liver stage activity is strategically important in antimalarial drug discovery because compounds that block development of exoerythrocytic parasite stages in hepatic cells often prove to have causal prophylactic activity in animal models.<sup>14,105</sup> Three compounds (original hit OSM-S-5 and the NN compounds OSM-S-38 and -111) were assessed for their activity against sporozoites in liver cells (vs atovaquone as positive control) and displayed varying potencies (Table 2;

**Table 2. Liver and Blood Stage Potencies for Selected Compounds**

OSM-S-	<i>P. berghei</i> IC <sub>50</sub> (nM)	cytotoxicity (HepG2, IC <sub>50</sub> , nM)	blood stage potency (3D7, nM)
5 <sup>a,b</sup>	24000, 14000	9800, 18000	502 <sup>c</sup>
38 <sup>a,b</sup>	19, 13	>50000	79 <sup>d</sup>
111 <sup>e</sup>	350	4400	147 <sup>d</sup>

<sup>a</sup>Data Set S30. <sup>b</sup>Data Set S31. <sup>c</sup>See Table SB1. <sup>d</sup>See Table SB3. <sup>e</sup>Data Set S29; Figure S31, liver stage potency curve.

Data Sets S29–31; Figure S31) that track with blood stage potency. Of particular note is the striking potency of the nontoxic compound OSM-S-38, which may at this level provide protection from malaria to people who have been treated with the compound after an infectious bite. Given the similarity observed in the possible mode of action (see next section), it is unclear why the arylpyrrole compound OSM-S-5 should exhibit such low levels of activity in the present assay vs the NN compounds. Possible explanations include differential solubility or stability; the level of cytotoxicity observed for OSM-S-5 suggests that the liver stage potency is probably just a generic toxic effect.

**Mode of Action.** As with most phenotypic hit to lead projects, the activity assays were carried out on whole parasites, giving a more realistic measure of activity than enzyme-based assays at the expense of an unknown mode of action. To predict the mode of action of the members of these series, eight representative active compounds (structures in the Supporting Information) across both subseries were evaluated in a yeast-based genetic sensitivity Hip Hop assay<sup>106,107</sup> that seeks mutations that result in enhanced compound potency, i.e., to identify any sensitive biological processes. The most potent compounds (OSM-S-39 and -51) showed enrichments for processes involved in chromatin remodeling and DNA repair. Indeed, inspection of the rank ordered genes for all compounds showed a preponderance of genes for diverse components of chromatin architecture, remodeling, and post-translational modification of chromatin proteins. These results are consistent with a perturbation of chromatin condensation with predicted follow-on effects on gene expression. Several groups have highlighted apparent plasticity of global gene expression in the malaria parasite which is manifested as a complex program of histone modification.<sup>108,109</sup> To attempt to gain additional insight into genome wide screens, we compared the data from several OSM compounds to a data set of 3200 similar genome wide screens to identify the 10 most similar profiles for each of

these OSM analogues. This provided clear concordance between these results and screens for other compounds, for OSM-S-39 (danthron, an anthraquinone derivative, as well as artemisinin), OSM-S-51 (the antineoplastics mitoxanthrone and mitomycin C), and OSM-S-9 (the antiarrhythmic agent DTBHQ). Taken together, these data point to a global dysregulation of chromatin architecture and suggest that further bioinformatics analysis both within the hip-hop data set and to orthologous data sets may be informative. The finding that the profile of the compounds is similar to that of known antimalarials (e.g., artemisinin) is provocative and warrants consideration. While it is formally possible that the predicted mechanism of select OSM compounds is similar to that of artemisinin, in our experience the degree of concordance we see between the OSM compounds and artemisinin suggests that the OSM compounds perturb a similar, but distinct, cellular target or pathway. This observation is consistent with our recent observation that, in a compendium of 3200 chemogenomic screens, the majority of compounds fell into  $\sim 45$  cellular response classes.<sup>106</sup> Our finding that different compounds from the same series have different profiles is also compelling and is consistent with previously published work demonstrating that single atom changes within a compound class can produce distinct cellular responses, which we attribute to small differences in compound structure resulting in significant fitness differences.<sup>110</sup> In order to derive experimental evidence for possible differences in mode of action, a parasite reduction ratio (PRR, generically “rate of killing”)<sup>111</sup> assay was run for six of these compounds (across both arylpyrrole and NN series) even though this assay is not a direct measure of differences in gene expression. The results (Text S1, Figure S37) suggest a common mechanism of action between the subseries and one that is distinct from artemisinin, which has previously exhibited a substantially faster killing profile. It remains a possibility that the *in vivo* target and/or impact on gene expression of the compounds in the arylpyrrole/near neighbor series are distinct and yet they still share a common mode of action, but this was not investigated further here.

The potency of the near neighbor set led to an *in silico* target prediction study being performed on OSM-S-39 (Text S11). The method employed a naive Bayes statistical model identifying molecular structural features of small molecules with protein targets, using the ChEMBL database. Similar statistical approaches based on known activities of compounds have been successfully applied to identify targets of antitubercular compounds.<sup>112–114</sup> A total of 1,287 proteins were included for which at least 50 active compounds were known with activities  $< 10 \mu\text{M}$ , with the other compounds in the database comprising the inactives. After the OSM compounds were scored against each target, the scores obtained were standardized by comparison with scores obtained through comparison with a random set of  $> 10\text{K}$  compounds. The prediction list was culled for proteins relevant to malaria (3D7 proteome), and the analysis gave as the most likely candidate targets, in order of significance, carboxy-terminal domain RNA polymerase II polypeptide A small phosphatase 1 (Q8I3U9), dihydroorotate dehydrogenase (DHODH, Q08210), SUMO-activating enzyme subunit 2 (Q8I553) and 1 (Q8IHS2), and cyclin-dependent kinase 1 (P61075). To kick-start the process of exploring these predictions, this compound and 24 others representative of both the hit compounds were evaluated in an experimental assay for DHODH inhibition (Data Set S36; Text S14) vs two positive control compounds (TCMDC-125840 and -123822)

known to inhibit this enzyme.<sup>115</sup> None of the compounds exhibited any activity, strongly suggesting that this is not the target for either subseries of compounds. A line of inquiry remaining open is the equivalent assays against the other targets identified in the *in silico* prediction.

Involvement of PfATP4, a putative plasma membrane ion pump, in the mechanism of action was ruled out. The relevant experiments measure the effect of a compound on the cytosolic  $[Na^+]_i$  ( $[Na^+]_i$ ) in isolated (trophozoite-stage) parasites pre-loaded with the  $Na^+$ -sensitive fluorescent indicator SBFI. PfATP4 is important since it is the proposed target of the antimalarial compound KAE609<sup>116</sup> that has recently successfully completed phase 2 trials,<sup>117</sup> as well as that of a pyrazoleamide<sup>118</sup> a dihydroisoquinolone ((+)-SJ-733),<sup>119</sup> various aminopyrazoles,<sup>120</sup> 28 of the 400 potent antiparasitodal compounds comprising the MMV Malaria Box,<sup>121</sup> and a triazolopyrazine series being investigated by the Open Source Malaria Consortium.<sup>122</sup> All of these compounds cause an immediate increase in parasite  $[Na^+]_i$  on addition to isolated parasites. Here, several compounds across both subseries were tested at 1  $\mu M$  for their effect on  $[Na^+]_i$  in saponin-isolated SBFI-loaded trophozoite stage 3D7 parasites.<sup>116</sup> In each case there was no significant effect on  $[Na^+]_i$ , consistent with PfATP4 not being the relevant biological target of this compound class (Data Set S37; Figure S38).

## CONCLUSIONS

The public deposition of novel antimalarial hits from phenotypic whole-cell assays has had a significant effect on worldwide

antimalarial drug discovery by providing an embarrassment of riches for the early stages of discovery. The plethora of alternative structures available for investigation has ultimately led to the series described in this paper being “parked” in favor of other possible avenues of inquiry. Interestingly the “stop” decision was straightforward to make in part because the decisions taken communally in the project had to be justifiable to all onlookers. The time taken to reach the stop decision in this case was probably slightly longer than would be expected from a traditionally structured project because certain contributions were not paid for or grant-supported, necessitating a lower priority than core interests of the contributing laboratories.

Both subseries investigated have members with major strengths, including potency with low molecular weight and significant late stage gametocyte and liver-stage activity coupled with low toxicity and low levels of activity in the hERG assay. Indeed, OSM-S-5 is bioavailable. Many of the most obvious structural changes to the hits, in several cases changes of a single atom, led to total loss, rather than moderation, of biological activity, known colloquially as “activity cliffs” (Figure 7). In the case of the ester-containing hit OSM-S-5, the ester was likely the key metabolic liability but could not be replaced with other common isosteres without loss of activity, yet (in the mouse model, at least) was found to be too short-lived in plasma to be taken on further. Other minor changes remote from the ester were also found to result in a precipitous decrease in potency. The “near neighbor” analogue set displayed impressive potency coupled with low cytotoxicity but

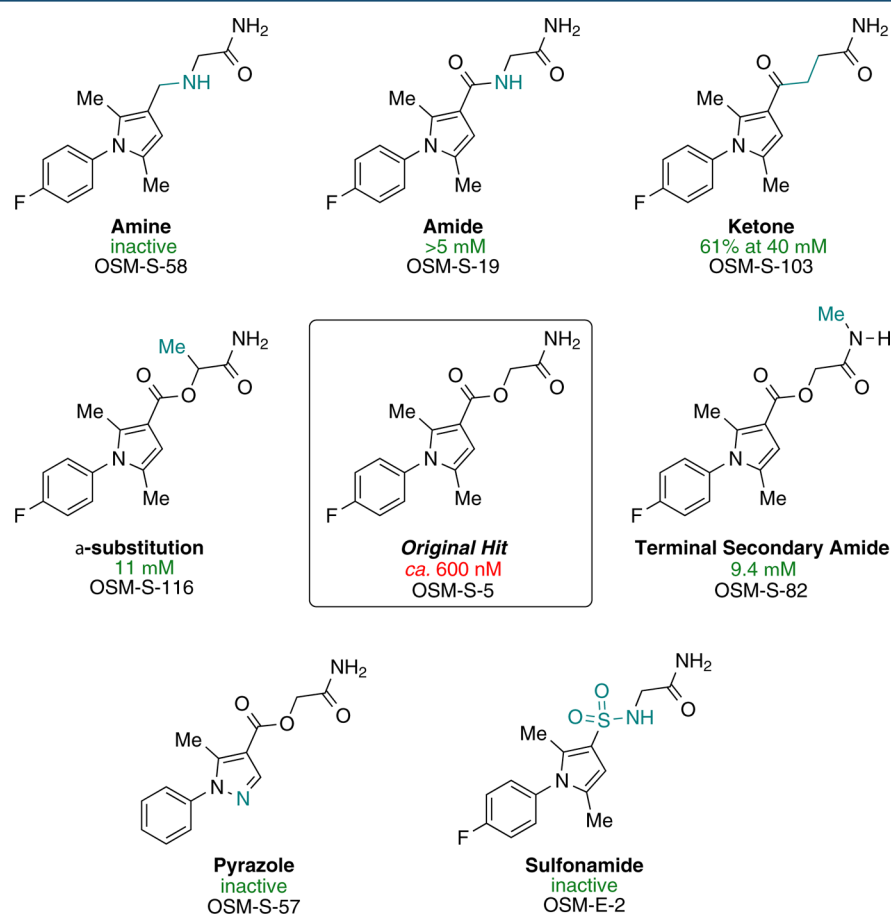
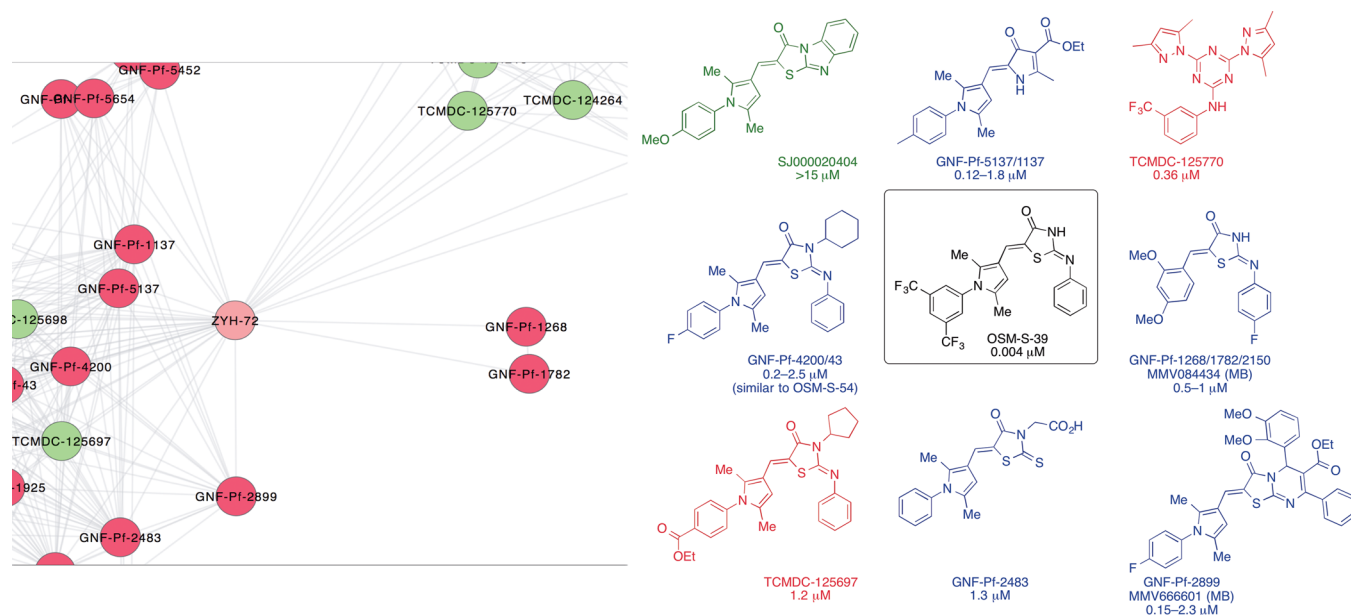


Figure 7. Sensitivity of the initial hit OSM-S-5 to minor structural changes.



**Figure 8.** Closest neighbors of OSM-S-39. (Left) Portion of a network similarity map generated for OSM-S-39 (coded as batch ZYH-72 in the figure) using methods described in the Supporting Information. (Right) The structures (stereochemistry assumed) and potencies (3D7, range if multiple values reported) of the most similar compounds in the ChEMBL database (v13) (key to compound sources: red = GSK TCAMS, blue = Novartis, green = St. Jude's).

alongside low solubility that could not easily be engineered out through side chain modifications.

The emergence of large amounts of open data in the field of drug discovery for malaria makes it straightforward to search for new ways forward in a project where a stop decision has been made. A similarity network map was generated for the most potent compound identified (OSM-S-39, Data Set S38, Figure S39) that identifies those compounds most similar in structure known to the open databases (Figure 8). Some of these are now represented in the MMV Malaria Box,<sup>123</sup> making simpler their investigation by other groups. A portion of this map is represented, alongside the structures and potencies of the closest neighbors in the map. While it is clear from this analysis that OSM-S-39 remains the most potent compound identified in this class, exploration of the neighboring structures in such open data sets may identify a way forward for this series that could suggest unexplored strategies to increase solubility with a realistic chance of maintaining potency, such as the potential scaffold-hop to the triazine TCMDC-125770. (One of the compounds shown from the Novartis screen, GNF-Pf-5137/1137, is from the same series as the published pyroloones.<sup>62</sup>) The recommended way forward, rather than further analogue synthesis in the present series, is to pursue clarity in the mechanisms of action of such series using, for example, generation of resistance coupled with genomic sequencing, or pull-down studies. Subsequent screening might then establish better starting points from related structures (informed by what has already been tried via comparison with network maps) against the relevant target.

One of the unique features of this project, the open source research method, ensures that the unexplored lines of inquiry remain open alongside the attendant data posted online that makes it straightforward for others to resume any portion of the research project as fully fledged participants, with access to both positive and negative data, details of all procedures as they were carried out (to aid reproducibility), and anecdotal insight into project loose ends that are easy to explore.

The machine-readability of the present project (for example the use of cheminformatic strings in the online electronic lab notebook) permits an unusually straightforward link between a high throughput screening result in a public database and a “live” research project that has investigated that compound.

Contributions to this project arose from disparate sources: (i) core government/NGO/foundation support, via both direct grant support and in-kind contributions (it is likely that direct support from public/philanthropic funds will always be needed as part of any open source drug discovery campaign); (ii) additional compound synthesis by undergraduate or postgraduate university students, either individually or as part of crowdsourced classes; (iii) expertise and advice emerging from the wider online community, both solicited and spontaneous; and (iv) additional laboratory contributions from other specialist centers, such as academic laboratories with block grants or pharmaceutical companies. These sources have considerable potential and can be mobilized further. For example, the time of researchers in major organizations (pharmaceutical companies, universities, CROs) could be contributed via policies that allow staff to work part-time on their own projects mimicking the *pro bono* system of the major law firms; such policies have been sporadically adopted (Figure S40)<sup>124</sup> and are a component of the GSK Tres Cantos research campus that kick-started and then supported the work described in this paper. In cases where resources are not forthcoming, direct payment (e.g., generated through crowdsourcing) could be awarded via competition, as occurred in this paper to create the graphical abstract. Each situation in the diverse base of support for an open source project will be different, and may be driven by appropriate selfless and selfish motives (e.g., contribution to a public health project vs securing publication authorship). Besides the transparency and completeness of the data arising from the research, this experiment with open source drug discovery made clear some of the advantages of this way of working: the breadth of people and expertise arising from “external” sources that were contributed rather than actively solicited, such as successful

compound synthesis by federated groups of graduate and undergraduate students, or the analysis of data obtained by others. Allowing participation by “strangers” introduces costs, such as the time required by a “core” team to maintain the project status, data, and methodology to allow meaningful contributions (in particular a real-time data set making clear the series SAR). We predict that tools and practices will arise to make easier some of these challenges, as open source methods in drug discovery and development become more widely adopted.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acscentsci.6b00086](https://doi.org/10.1021/acscentsci.6b00086).

Complete electronic laboratory notebooks, images of NMR spectra for novel compounds, and archived snapshots of OSM wiki pages are available from The University of Sydney eScholarship Repository at <http://hdl.handle.net/2123/14132>, <http://hdl.handle.net/2123/14123>, and <http://hdl.handle.net/2123/15389>, respectively.

Chemical and biological protocols, other text-based files, and screengrabs (PDF)

All molecules (XLS)

All molecules in SDF format (ZIP)

Potency data from Ralph assay 1 (XLS)

Potency data from GSK assay 1 (XLS)

Potency data from Avery assay 1 (XLS)

Similarity network data for Tres Cantos series in Cytoscape format (ZIP)

Map of purchasable compounds around OSM-S-35 in Cytoscape format (ZIP)

Potency data from Ralph assay 2 (ZIP)

Potency data from GSK assay 2 (XLS)

Potency data from Avery assay 2 (XLS)

Potency data from Avery assay 3a (PDF)

Potency data from Avery assay 3b (XLS)

X-ray structural information for OSM-S-35, -42, -54, and -9 (ZIP)

Potency data from Avery assay 5 (XLS)

Potency data from Guy assay (XLS)

Potency data from Avery assay 3c (XLS)

Potency data from Dundee assay 1 (PDF)

Potency data from Dundee assay 2 (PDF)

Bioisostere analysis in Cytoscape format (ZIP)

Solubility and microsomal stability data for OSM-S-5, -6, -9, -10, -37, -38, -39, and -54 (ZIP)

Solubility and microsomal stability data for OSM-S-111 (XLS)

Data from oral *in vivo* *P. berghei* mouse trial for compounds OSM-S-5, -6, and -35 (XLS)

Pharmacokinetic data from oral *in vivo* *P. berghei* mouse trial (XLS)

Human and mouse plasma stability data for OSM-S-5 (XLS)

Metabolite identification assay data for OSM-S-35 in human liver microsomes (XLS)

hERG assay data for compounds OSM-S-5 and -35 (XLS)

Avery late stage gametocyte assay 1 (XLS)

Avery late stage gametocyte assay 2 (XLS)

UCSD liver stage assay 1 (XLS)

UCSD liver stage assay 2 (XLS)

UCSD liver stage assay 3 (XLS)

Nislow gene set enrichment analysis data (XLS)

Nislow gene set enrichment analysis data, Spotfire format (ZIP)

Nislow gene set enrichment analysis data (PDF)

Nislow fitness defect scores for all deletion strains (XLS)

Data from GSK DHODH assay (XLS)

Data from Kirk ion regulation assay (XLS)

Similarity map surrounding OSM-S-39 in Cytoscape format (ZIP)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [matthew.todd@sydney.edu.au](mailto:matthew.todd@sydney.edu.au), [@mattoddchem](mailto:@mattoddchem).

### Present Address

<sup>†</sup>I.M.W.: Merck Research Laboratories, 470 Atlantic Ave, Boston, MA 02210, United States.

### Author Contributions

<sup>‡</sup>The authorship list for this paper is alphabetical aside from the bookenders: the corresponding author is shown last, and the first three authors performed the majority of the experimental work.

### Funding

Medicines for Malaria Venture (MMV), the Australian Research Council (LP120100552, LP120200557), the Australian National Health and Medical Research Council (Project Grant 1042272). G.P. and J.P.O. were supported by EMBL Member States and MMV. The characterization by NMR spectroscopy of compounds from Lawrence University was made possible by the National Science Foundation (CHE-0923473). C.S. was supported by Wellcome Trust Biomedical Resources Grant 099156/Z/12/Z.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We would like to thank the Labtrove team at the University of Southampton led by Professor Jeremy Frey, OpenWetWare for hosting the OSM wiki, and publicly available platforms such as Twitter, G+, Facebook, and Github that have helped in the wider dissemination required for live research projects. We thank Michael Robins for ongoing technical support of the online project components. OSM-A-1 through -A-4 were synthesized by undergraduate organic chemistry students at Lawrence University; in particular, the work of Shay Albrecht, Margaret Brickner, Patrick Doughty, Joshua Graber, Sarah Gunby, Briana Harter, Alexander Hurlburt, Jake Johengen, Sarah Prophet, Greta Schmitt, Chad Skaer, and Konstantinos Vlachos is acknowledged. Dr. Colin Campbell (The University of Edinburgh) is thanked for provision of laboratory resources and time, as is the University of Edinburgh for funding (to P.T.). The Australian Red Cross Blood Service is thanked for the provision of blood (to K.K. and V.A.). The authors wish to thank Prof. David Fidock (Columbia University, NY) for the provision of the NF54 transgenic parasite expressing GFP linked to the early expressed gametocyte specific protein, Pfs16. Miscellaneous project inputs were gratefully received from William Jackson (Creative Chemistry), Egon Willighagen (Maastricht University), Martine Keenan (Epichem), Darren Dressen (Los Altos High School), Robert Snell (The University

of Cambridge, now Charterhouse School) and Joie Garfunkel (Merck). For their generous time and insights, we thank the other (sometimes anonymous) online contributors. We thank Viputheshwar Sitaraman (Draw Science) for creating the graphical abstract as part of a competition for this purpose. We thank Ginger Taylor for creating The Synaptic Leap website, which hosted much of the early activity in the project. We would like to thank student volunteers (Min Kyung Chong, Jun Ki Hong, Martina Yousif, Sebastien Dath) for help with online data entry.

## REFERENCES

- (1) WHO. World Malaria Report 2015, <http://www.who.int/malaria/publications/world-malaria-report-2015/en/>, accessed March 15, 2016.
- (2) Snow, R. W.; Craig, M.; Newton, C.; Steketee, R. *The Public Health Burden of Plasmodium falciparum Malaria in Africa: Deriving the Numbers*; Fogarty International Centre, National Institutes of Health: Bethesda, MD, 2010, 1–81.
- (3) The Global Malaria Action Plan, Geneva, 2008, <http://www.rollbackmalaria.org/microsites/gmap/>, accessed Dec 8, 2015.
- (4) Carrara, V. L.; Zwang, J.; Ashley, E. A.; Price, R. N.; Stepniwska, K.; Barends, M.; et al. Changes in the treatment responses to artesunate-mefloquine on the northwestern border of Thailand during 13 years of continuous deployment. *PLoS One* 2009, 4, e4551.
- (5) Dondorp, A. M.; Nosten, F.; Yi, P.; Das, D.; Phyto, A. P.; Tarning, J.; et al. Artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* 2009, 361, 455–467.
- (6) Phyto, A. P.; Nkhoma, S.; Stepniwska, K.; Ashley, E. A.; Nair, S.; McGready, R.; et al. Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. *Lancet* 2012, 379, 1960–1966.
- (7) RTS,S Clinical Trials Partnership. Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomized, controlled trial. *Lancet* 2015, 386, 31–45.
- (8) Olotu, A.; Fegan, G.; Wambua, J.; Nyangweso, G.; Awuondo, K. O.; Leach, A.; et al. Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure. *N. Engl. J. Med.* 2013, 368, 1111–1120.
- (9) Wells, T. N. C.; van Huijsduijnen, R. H.; Van Voorhis, W. C. Malaria medicines: a glass half full? *Nat. Rev. Drug Discovery* 2015, 14, 424–442.
- (10) Flannery, E. L.; Chatterjee, A. K.; Winzeler, E. A. Antimalarial drug discovery – approaches and progress towards new medicines. *Nat. Rev. Microbiol.* 2013, 11, 849–862.
- (11) Burrows, J. N.; Elliott, R. L.; Kaneko, T.; Mowbray, C. E.; Waterson, D. The role of modern drug discovery in the fight against neglected and tropical diseases. *MedChemComm* 2014, 5, 688–700.
- (12) Peatey, C. L.; Leroy, D.; Gardiner, D. L.; Trenholme, K. R. Antimalarial drugs: how effective are they against *Plasmodium falciparum* gametocytes? *Malar. J.* 2012, 11, 34.
- (13) Duffy, S.; Avery, V. M. Identification of inhibitors of *Plasmodium falciparum* gametocyte development. *Malar. J.* 2013, 12, 408.
- (14) McNamara, C. W.; Lee, M. C. S.; Lim, C. S.; Lim, S. H.; Roland, J.; Nagle, A.; et al. Targeting *Plasmodium* PI(4)K to eliminate malaria. *Nature* 2013, 504, 248–253.
- (15) Dembélé, L.; Franetich, J.-F.; Lorthiois, A.; Gego, A.; Zeeman, A.-M.; Kocken, C. H. M.; et al. Persistence and activation of malaria hypnozoites in long-term primary hepatocyte cultures. *Nat. Med.* 2014, 20, 307–312.
- (16) Scannell, J. W.; Blanckley, A.; Boldon, H.; Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discovery* 2012, 11, 191–200.
- (17) Paul, S. M.; Mytelka, D. S.; Dunwiddie, C. T.; Persinger, C. C.; Munos, B. H.; Lindborg, S. R.; et al. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discovery* 2010, 9, 203–214.
- (18) Kneller, R. The importance of new companies for drug discovery: origins of a decade of new drugs. *Nat. Rev. Drug Discovery* 2010, 9, 867–882.
- (19) The 10/90 Report on Health Research 1999, *Global Forum for Health Research*, 1999, <http://www.css.ethz.ch/content/specialinterest/gess/cis/center-for-securities-studies/en/services/digital-library/publications/publication.html/20437>, accessed Aug 9, 2016.
- (20) Jorgensen, W. L. Challenges for academic drug discovery. *Angew. Chem., Int. Ed.* 2012, 51, 11680–11684.
- (21) Lezaun, J.; Montgomery, C. M. The pharmaceutical commons. Sharing and exclusion in global health drug development. *Sci. Technol. Hum. Val.* 2015, 40, 3–29.
- (22) Uniting to Combat Neglected Tropical Diseases, <http://www.unitingtocombatntds.org>, accessed Dec 8, 2015.
- (23) Institute of One World Health, now the PATH drug development program, <http://sites.path.org/drugdevelopment/>, accessed Dec 8, 2015.
- (24) Novartis Institutes for Biomedical Research, <https://www.nibr.com/our-research/institutes/novartis-institute-tropical-diseases>, accessed Dec 8, 2015.
- (25) Tres Cantos OpenLab Foundation, <http://www.openlabfoundation.org/>, accessed Dec 8, 2015.
- (26) Hopkins, A. L.; Witty, M. J.; Nwaka, S. Neglected diseases. Mission possible. *Nature* 2007, 449, 166–169.
- (27) Ortí, L.; Carbajo, R. J.; Pieper, U.; Eswar, N.; Maurer, S. M.; Rai, A. K.; et al. A kernel for the Tropical Disease Initiative. *Nat. Biotechnol.* 2009, 27, 320–321.
- (28) World Intellectual Property Office Re:Search, <http://www.wipo.int/research/en/>, accessed Dec 8, 2015.
- (29) Jakobsen, P. H.; Wang, M.-W.; Nwaka, S. Innovative partnerships for drug discovery against neglected diseases. *PLoS Neglected Trop. Dis.* 2011, 5, e1221.
- (30) Lotharius, J.; Gamo-Benito, F. J.; Angulo-Barturen, I.; Clark, J.; Connelly, M.; Ferrer-Bazaga, S.; et al. Repositioning: the fast track to new anti-malarial medicines? *Malar. J.* 2014, 13, 143.
- (31) Hunter, J.; Stephens, S. Is open innovation the way forward for big pharma? *Nat. Rev. Drug Discovery* 2010, 9, 87–88.
- (32) Maurer, S. M.; Rai, A.; Sali, A. Finding cures for tropical diseases: is open source an answer? *PLoS Med.* 2004, 1, e56.
- (33) Munos, B. Can open-source R&D reinvigorate drug research? *Nat. Rev. Drug Discovery* 2006, 5, 723–729.
- (34) Singh, S. India takes an open source approach to drug discovery. *Cell* 2008, 133, 201–203.
- (35) Munos, B. Can open-source drug R&D repower pharmaceutical innovation? *Clin. Pharmacol. Ther.* 2010, 87, 534–536.
- (36) Woelfle, M.; Olliaro, P.; Todd, M. H. Open science is a research accelerator. *Nat. Chem.* 2011, 3, 745–748.
- (37) Ortí, L.; Carbajo, R. J.; Pieper, U.; Eswar, N.; Maurer, S. M.; Rai, A. K.; et al. A kernel for open source drug discovery in tropical diseases. *PLoS Neglected Trop. Dis.* 2009, 3, e418.
- (38) Årdal, C.; Alstadsæter, A.; Rottingen, J. A. Common characteristics of open source software development and applicability for drug discovery: a systematic review. *Health Res. Policy Syst.* 2011, 9, 36.
- (39) DeLano, W. L. The case for open-source software in drug discovery. *Drug Discovery Today* 2005, 10, 213–217.
- (40) Woelfle, M.; Seerden, J.-P.; de Gooijer, J.; Pouwer, K.; Olliaro, P.; Todd, M. H. Resolution of praziquantel. *PLoS Neglected Trop. Dis.* 2011, 5, e1260.
- (41) Org Prep Daily, <https://orgprepdaily.wordpress.com/>, accessed Dec 8, 2015.
- (42) ChemSpider Synthetic Pages, <http://cssp.chemspider.com/>, accessed Dec 8, 2015.
- (43) Reaves, M. L.; Sinha, S.; Rabinowitz, J. D.; Kruglyak, L.; Redfield, R. J. Absence of detectable arsenate in DNA from arsenate-grown GFAJ-1 cells. *Science* 2012, 337, 470–473.
- (44) The UsefulChem Project, <http://usefulchem.wikispaces.com/>, accessed Dec 8, 2015.

- (45) Indian OSDD Project, <http://www.osdd.net/>, accessed Dec 8, 2015.
- (46) Anand, P.; Sankaran, S.; Mukherjee, S.; Yeturu, K.; Laskowski, R.; Bhardwaj, A.; et al. Structural annotation of *Mycobacterium tuberculosis* proteome. *PLoS One* **2011**, *6*, e27044.
- (47) Bhardwaj, A.; Scaria, V.; Raghava, G. P. S.; Lynn, A. M.; Chandra, N.; Banerjee, S.; et al. Open source drug discovery – a new paradigm of collaborative research in tuberculosis drug development. *Tuberculosis* **2011**, *91*, 479–486.
- (48) Jefferson, R. Science as social enterprise: The CAMBIA BiOS Initiative. *Innovations: Technology, Governance, Globalization* **2006**, *1*, 13–44.
- (49) Todd, M. H. Open source drug discovery for malaria. *Synaptic Leap* 2011, Node 343. <http://www.thesynapticleap.org/node/343>.
- (50) Carroll, M. W. Creative commons and the openness of open access. *N. Engl. J. Med.* **2013**, *368*, 789–791.
- (51) Robertson, M. N.; Ylloja, P. M.; Williamson, A. E.; Woelfle, M.; Robins, M.; Badiola, K. A.; et al. Open source drug discovery - a limited tutorial. *Parasitology* **2014**, *141*, 148–157.
- (52) Badiola, K. A.; Bird, C.; Brocklesby, W. S.; Casson, J.; Chapman, R. T.; Coles, S. J.; et al. Experiences with a researcher-centric ELN. *Chem. Sci.* **2015**, *6*, 1614–1629.
- (53) Gamo, F.-J.; Sanz, L. M.; Vidal, J.; de Cozar, C.; Alvarez, E.; Lavandera, J.-L.; et al. Thousands of chemical starting points for antimalarial lead identification. *Nature* **2010**, *465*, 305–310.
- (54) Guiguemde, W. A.; Shelat, A. A.; Bouck, D.; Duffy, S.; Crowther, G. J.; Davis, P. H.; et al. Chemical genetics of *Plasmodium falciparum*. *Nature* **2010**, *465*, 311–315.
- (55) Plouffe, D.; Brinker, A.; McNamara, C.; Henson, K.; Kato, N.; Kuhlen, K.; et al. *In silico* activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9059–9064.
- (56) ChEMBL Neglected Tropical Disease Archive, <http://www.ebi.ac.uk/chemblntd>.
- (57) Calderón, F.; Barros, D.; Bueno, J. M.; Coterón, J. M.; Fernández, E.; Gamo, F. J.; et al. An invitation to open innovation in malaria drug discovery: 47 quality starting points from the TCAMS. *ACS Med. Chem. Lett.* **2011**, *2*, 741–746.
- (58) Rueda, L.; Castellote, I.; Castro-Pichel, J.; Chaparro, M. J.; de la Rosa, J. C.; Garcia-Perez, A.; et al. Cyclopropyl carboxamides: a new oral antimalarial series derived from the Tres Cantos anti-malarial set (TCAMS). *ACS Med. Chem. Lett.* **2011**, *2*, 840–844.
- (59) Sanz, L. M.; Jiménez-Díaz, M. B.; Crespo, B.; De-Cozar, C.; Almela, M. J.; Angulo-Barturen, I.; et al. Cyclopropyl carboxamides, a chemically novel class of antimalarial agents identified in a phenotypic screen. *Antimicrob. Agents Chemother.* **2011**, *55*, 5740–5745.
- (60) Calderón, F.; Vidal-Mas, J.; Burrows, J.; de la Rosa, J. C.; Jiménez-Díaz, M. B.; Mulet, T.; et al. A divergent SAR study allows optimization of a potent 5-HT<sub>2c</sub> inhibitor to a promising antimalarial scaffold. *ACS Med. Chem. Lett.* **2012**, *3*, 373–377.
- (61) Todd, M. H. Known near neighbors of initial Tres Cantos leads. *Synaptic Leap* 2011, Node 349. <http://www.thesynapticleap.org/node/349>.
- (62) Murugesan, D.; Mital, A.; Kaiser, M.; Shackelford, D. M.; Morizzi, J.; Katneni, K.; et al. Discovery and structure-activity relationships of pyrrolone antimalarials. *J. Med. Chem.* **2013**, *56*, 2975–2990.
- (63) Murugesan, D.; Kaiser, M.; White, K. L.; Norval, S.; Riley, J.; Wyatt, P. G.; et al. Structure-activity relationship studies of pyrrolone antimalarial agents. *ChemMedChem* **2013**, *8*, 1537–1544.
- (64) Relevant Open Source Malaria electronic laboratory notebooks, <http://malaria.ouexperiment.org/>, available offline at <http://hdl.handle.net/2123/14132>.
- (65) Gu, X. P.; Nishida, N.; Ikeda, I.; Okahara, M. 2-(Chloromethyl)-3,5-dioxahex-1-ene. An effective acetylating reagent. *J. Org. Chem.* **1987**, *52*, 3192–3196.
- (66) Chiu, P.-K.; Sannes, M. P. The synthesis and chemistry of azolenines. Part 18. Preparation of 3-ethoxycarbonyl-3H-pyrroles via the Paal-Knorr reaction, and sigmatropic rearrangements involving competitive ester migrations to C-2, C-4 and N. *Tetrahedron* **1990**, *46*, 3439–3456.
- (67) Ylloja, P. M. Proposed resynthesis strategy for TCMD-123812 and TCMD-123794. *Synaptic Leap* 2011, Node 344. <http://www.thesynapticleap.org/node/344>.
- (68) Avery, V.; Baell, J.; Batra, S.; Burrows, J.; Bhattacharyya, S.; Calderon, F.; et al. *Open Source Malaria deposition 1: screening of compounds synthesized and sourced by the OSDDMalaria Consortium against Plasmodium falciparum and measurement of cytotoxicity.* 2013, <https://www.ebi.ac.uk/chembl/malaria/doc/inspect/CHEMBL2113921>. DOI: 10.6019/CHEMBL2113921.
- (69) Open Source Malaria Master Compound List, <http://tinyurl.com/OSM-Compounds>, accessed Nov 23, 2015.
- (70) ChemInfo Malaria Hub, Open Source Malaria Series 1, [http://malaria.cheminfo.org/Display\\_data.html?series=1](http://malaria.cheminfo.org/Display_data.html?series=1), accessed Nov 23, 2015.
- (71) Ertl, P.; Patiny, L.; Sander, T.; Rufener, C.; Zasso, M. Wikipedia chemical structure explorer: substructure and similarity searching of molecules from Wikipedia. *J. Cheminf.* **2015**, *7*, 10.
- (72) Swain, C. Importing Open Source Malaria project data, [http://macinchem.org/reviews/vortex/tut26/scripting\\_vortex26.php](http://macinchem.org/reviews/vortex/tut26/scripting_vortex26.php), **2015**, accessed Nov 23, 2015.
- (73) Swain, C. Accessing Open Source Malaria data using an iPython notebook, **2015**, <http://www.macinchem.org/reviews/osm/osmipython.php>, accessed Nov 23, 2015.
- (74) Avery, V. M.; Bashyam, S.; Burrows, J. N.; Duffy, S.; Papadatos, G.; Puthukkuti, S.; et al. Screening and hit evaluation of a chemical library against blood-stage *Plasmodium falciparum*. *Malar. J.* **2014**, *13*, 190.
- (75) Open Source Malaria Google+ Community, <https://plus.google.com/114702323662314783325/posts/Pma1Ddk2XCy>, **2012**, accessed Dec 8, 2015.
- (76) Bhandari, S. V.; Bothara, K. G.; Patil, A. A.; Chitre, T. S.; Sarkate, A. P.; Gore, S. T.; et al. Design, synthesis and pharmacological screening of novel antihypertensive agents using hybrid approach. *Bioorg. Med. Chem.* **2009**, *17*, 390–400.
- (77) Urbano, M.; Guerrero, M.; Velaparthi, S.; Crisp, M.; Chase, P.; Hodder, P.; et al. Discovery, synthesis and SAR analysis of novel selective small molecule S1P<sub>4</sub>-R agonists based on a (2Z,5Z)-5-((pyrrol-3-yl)methylene)-3-alkyl-2-(alkylimino)thiazolidin-4-one chemotype. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 6739–6745.
- (78) Singh, T.; Sharma, P. K.; Mondal, S. C.; Kumar, N. Difference in rate of reaction of some thiazolidinone derivatives with synthesis and spectral characterization. *J. Adv. Sci. Res.* **2011**, *2*, 58–61.
- (79) Hopkins, A. L.; Keserü, G. M.; Leeson, P. D.; Rees, D. C.; Reynolds, C. H. The role of ligand efficiency metrics in drug discovery. *Nat. Rev. Drug Discovery* **2014**, *13*, 105–121.
- (80) Todd, M. H. Biological results for first set of compounds. *Synaptic Leap* 2012, Node 367. <http://www.thesynapticleap.org/node/367>.
- (81) 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.
- (82) KNIME: <https://www.knime.org/>, accessed Dec 8, 2015.
- (83) Saubern, S.; Guha, R.; Baell, J. B. KNIME workflow to assess PAINS filters in SMARTS format. Comparison of RDKit and Indigo cheminformatics libraries. *Mol. Inf.* **2011**, *30*, 847–850.
- (84) Baell, J. B. Observations on screening-based research and some concerning trends in the literature. *Future Med. Chem.* **2010**, *2*, 1529–1546.
- (85) Mok, N. Y.; Maxe, S.; Brenk, R. Locating sweet spots for screening hits and evaluating pan-assay interference filters from the performance analysis of two lead-like libraries. *J. Chem. Inf. Model.* **2013**, *53*, 534–544.
- (86) Bruns, R. F.; Watson, I. A. Rules for identifying potentially reactive or promiscuous compounds. *J. Med. Chem.* **2012**, *55*, 9763–9772.

- (87) Mendgen, T.; Steuer, C.; Klein, C. D. Privileged scaffolds or promiscuous binders: a comparative study on rhodanines and related heterocycles in medicinal chemistry. *J. Med. Chem.* **2012**, *55*, 743–753.
- (88) Thorne, N.; Auld, D. S.; Inglese, J. Apparent activity in high-throughput screening: origins of compound-dependent assay interference. *Curr. Opin. Chem. Biol.* **2010**, *14*, 315–324.
- (89) Lowe, D. Polluting the literature with PAINs. *In the Pipeline*, 2010, [http://blogs.sciencemag.org/pipeline/archives/2010/02/08/polluting\\_the\\_literature\\_with\\_pains](http://blogs.sciencemag.org/pipeline/archives/2010/02/08/polluting_the_literature_with_pains), accessed Dec 8, 2015.
- (90) Lowe, D. PAINs go mainstream. *In the Pipeline*, 2014, [http://blogs.sciencemag.org/pipeline/archives/2014/09/26/pains\\_go\\_mainstream](http://blogs.sciencemag.org/pipeline/archives/2014/09/26/pains_go_mainstream), accessed Dec 8, 2015.
- (91) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discovery* **2011**, *10*, 307–317.
- (92) SmartsFilter, <http://pasilla.health.unm.edu/tomcat/biocomp/smartsfilter>, accessed Dec 8, 2015.
- (93) Huth, J. R.; Mendoza, R.; Olejniczak, E. T.; Johnson, R. W.; Cothron, D. A.; Liu, Y.; et al. ALARM NMR: a rapid and robust experimental method to detect reactive false positives in biochemical screens. *J. Am. Chem. Soc.* **2005**, *127*, 217–224.
- (94) Ekins, S.; Williams, A. J. When pharmaceutical companies publish large datasets: an abundance of riches or fool's gold? *Drug Discovery Today* **2010**, *15*, 812–815.
- (95) OSM Consortium, OSDD Malaria public consultation 1, 2012, <https://www.youtube.com/watch?v=ooM8kuo14Bg>, accessed Dec 8, 2015, available for download from The University of Sydney eScholarship Repository at <http://hdl.handle.net/2123/14110>.
- (96) Todd, M. H. Desired compounds consultation phase 2. *Synaptic Leap* 2012, Node 412. <http://www.thesynapticleap.org/node/412>.
- (97) Ylioja, P. M. Consultation outcome. *Synaptic Leap* 2012, Node 416. <http://www.thesynapticleap.org/node/416>.
- (98) Boström, J.; Hogner, A.; Llinàs, A.; Wellner, E.; Plowright, A. T. Oxadiazoles in medicinal chemistry. *J. Med. Chem.* **2012**, *55*, 1817–1830.
- (99) OSM electronic laboratory notebook, pharmacophore modeling of OSM compounds: [http://malaria.oureperiment.org/pharmacophore\\_modeling](http://malaria.oureperiment.org/pharmacophore_modeling), accessed Dec 8, 2015.
- (100) Olesen, P. H. The use of bioisosteric groups in lead optimization. *Curr. Opin. Drug Discovery Dev.* **2001**, *4*, 471–478.
- (101) Open Source Malaria Wiki, GSK Arylpyrrole Series: [http://openwetware.org/wiki/OpenSourceMalaria:GSK\\_Arylpyrrole\\_Series](http://openwetware.org/wiki/OpenSourceMalaria:GSK_Arylpyrrole_Series), accessed Dec 8, 2015. (Snapshot available for download at The University of Sydney eScholarship Repository at <http://hdl.handle.net/2123/15389>).
- (102) Berry, L. M.; Wollenberg, L.; Zhao, Z. Esterase activities in the blood, liver and intestine of several preclinical species and humans. *Drug Metab. Lett.* **2009**, *3*, 70–77.
- (103) Danker, T.; Möller, C. Early identification of hERG liability in drug discovery programs by automated patch clamp. *Front. Pharmacol.* **2014**, DOI: 10.3389/fphar.2014.00203.
- (104) Ruecker, A.; Mathias, D. K.; Straschil, U.; Churcher, T. S.; Dinglasan, R. R.; Leroy, D.; et al. A male and female gametocyte functional viability assay to identify biologically relevant malaria transmission-blocking drugs. *Antimicrob. Agents Chemother.* **2014**, *58*, 7292–7302.
- (105) Meister, S.; Plouffe, D. M.; Kuhlen, K. L.; Bonamy, G. M. C.; Wu, T.; Barnes, S. W.; et al. Imaging of *Plasmodium* liver stages to drive next-generation antimalarial drug discovery. *Science* **2011**, *334*, 1372–1377.
- (106) Lee, A. Y.; St-Onge, R. P.; Proctor, M. J.; Wallace, I. M.; Nile, A. H.; Spagnuolo, P. A.; et al. Mapping the cellular response to small molecules using chemogenomic fitness signatures. *Science* **2014**, *344*, 208–211.
- (107) Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. A.; et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 15545–15550.
- (108) Gupta, A. P.; Chin, W. H.; Zhu, L.; Mok, S.; Luah, Y.-H.; Lim, E.-H.; et al. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite *Plasmodium falciparum*. *PLoS Pathog.* **2013**, *9*, e1003170.
- (109) Jiang, L.; Mu, J.; Zhang, Q.; Ni, T.; Srinivasan, P.; Rayavara, K.; et al. PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum*. *Nature* **2013**, *499*, 223–227.
- (110) Yu, L.; Lopez, A.; Anafloos, A.; El Bali, B.; Hamal, A.; Ericson, E.; et al. Chemical-genetic profiling of imidazo[1,2-*a*]pyridines and –pyrimidines reveals target pathways conserved between yeast and human cells. *PLoS Genet.* **2008**, *4*, e1000284.
- (111) Sanz, L. M.; Crespo, B.; De-Cózar, C.; Ding, X. C.; Llergo, J. L.; Burrows, J. N.; et al. *P. falciparum* *in vitro* killing rates allow to discriminate between different antimalarial mode-of-action. *PLoS One* **2012**, *7*, e30949.
- (112) Prathipati, P.; Ma, N. L.; Manjunatha, U. H.; Bender, A. Fishing the target of antitubercular compounds: *in silico* target deconvolution model development and validation. *J. Proteome Res.* **2009**, *8*, 2788–2798.
- (113) Koutsoukas, A.; Simms, B.; Kirchmair, J.; Bond, P. J.; Whitmore, A. V.; Zimmer, S.; et al. From *in silico* target prediction to multi-target drug design: current databases, methods and applications. *J. Proteomics* **2011**, *74*, 2554–2574.
- (114) Martínez-Jiménez, F.; Papadatos, G.; Yang, L.; Wallace, I. M.; Kumar, V.; Pieper, U.; et al. Target prediction for an open access set of compounds active against *Mycobacterium tuberculosis*. *PLoS Comput. Biol.* **2013**, *9*, e1003253.
- (115) Baldwin, J.; Farajallah, A. M.; Malmquist, N. A.; Rathod, P. K.; Phillips, M. A. Malarial dihydroorotate dehydrogenase: substrate and inhibitor specificity. *J. Biol. Chem.* **2002**, *277*, 41827–41834.
- (116) Spillman, N. J.; Allen, R. J. W.; McNamara, C. W.; Yeung, B. K. S.; Winzeler, E. A.; Diagana, T. T.; et al. Na<sup>+</sup> regulation in the malaria parasite *Plasmodium falciparum* involves the cation ATPase PfATP4 and is a target of the spiroindolone antimalarials. *Cell Host Microbe* **2013**, *13*, 227–237.
- (117) White, N. J.; Pukrittayakamee, S.; Phyto, A. P.; Rueangweerayut, R.; Nosten, F.; Jittamala, P.; et al. Spiroindolone KAE609 for falciparum and vivax malaria. *N. Engl. J. Med.* **2014**, *371*, 403–410.
- (118) Vaidya, A. B.; Morrissey, J. M.; Zhang, Z.; Das, S.; Daly, T. M.; Otto, T. D.; et al. Pyrazoleamide compounds are potent antimalarials that target Na<sup>+</sup> homeostasis in intraerythrocytic *Plasmodium falciparum*. *Nat. Commun.* **2014**, *5*, 5521.
- (119) Jiménez-Díaz, M. B.; Ebert, D.; Salinas, Y.; Pradhan, A.; Lehane, A. M.; Myrand-Lapierre, M.-E.; et al. +)-SJ733, a clinical candidate for malaria that acts through ATP4 to induce rapid host-mediated clearance of *Plasmodium*. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E5455–E5462.
- (120) Flannery, E. L.; McNamara, C. W.; Kim, S. W.; Kato, T. S.; Li, F.-W.; Teng, C. H.; et al. Mutations in the P-type cation-transporter ATPase 4, PfATP4, mediate resistance to both aminopyrazole and spiroindolone antimalarials. *ACS Chem. Biol.* **2015**, *10*, 413–420.
- (121) Lehane, A. M.; Ridgway, M. C.; Baker, E.; Kirk, K. Diverse chemotypes disrupt ion homeostasis in the malaria parasite. *Mol. Microbiol.* **2014**, *94*, 327–339.
- (122) Open Source Malaria Wiki, Triazolopyrazine (TP) Series: [http://openwetware.org/wiki/OpenSourceMalaria:Triazolopyrazine\\_%28TP%29\\_Series](http://openwetware.org/wiki/OpenSourceMalaria:Triazolopyrazine_%28TP%29_Series), accessed Dec 8, 2015. (Snapshot available for download at The University of Sydney eScholarship Repository at <http://hdl.handle.net/2123/15389>).
- (123) van Voorhis, W. C.; Adams, J. H.; Adelfio, R.; Ah Yong, V.; Akabas, M. H.; Alano, P.; et al. Open source drug discovery with the malaria box compound collection for neglected diseases and beyond. *PLoS Pathog.* **2016**, *12*, e1005763.
- (124) Lowe, D. Company time for your own ideas, or not? *In the Pipeline*, 2016 <http://blogs.sciencemag.org/pipeline/archives/2016/07/21/company-time-for-your-own-ideas-or-not>, accessed July 25, 2016.