Fragment-Based Screening of a Natural Product Library against 62 Potential Malaria Drug Targets Employing Native Mass Spectrometry

Hoan Vu,† Liliana Pedro,† Tin Mak,‡ Brendan McCormick,‡ Jessica Rowley,† Miaomiao Liu,† Angela Di Capua,† Billy Williams-Noonan,† Ngoc B. Pham,† Rebecca Pouwer,† Bao Nguyen,† Katherine T. Andrews,† Tina Skinner-Adams,† Jessica Kim,‡ Wim G. J. Ho,‡ Raymond Hui,§ Gregory J. Crowther,⊥ Wesley C. Van Voorhis,⊥ and Ronald J. Quinn,*†

†Griffith Institute for Drug Discovery, Griffith University, 170 Kessels Road, Nathan, Queensland 4111, Australia
‡Department of Biochemistry and †Department of Medicine, University of Washington, 1705 NE Pacific Street, Seattle, Washington 98195-7350, United States
§Structural Genomics Consortium, University of Toronto, MaRS South Tower, seventh floor 101 College Street, Toronto, Ontario MSG 1L7, Canada

Supporting Information

ABSTRACT: Natural products are well known for their biological relevance, high degree of three-dimensionality, and access to areas of largely unexplored chemical space. To shape our understanding of the interaction between natural products and protein targets in the postgenomic era, we have used native mass spectrometry to investigate 62 potential protein targets for malaria using a natural-product-based fragment library. We reveal here 96 low-molecular-weight natural products identified as binding partners of 32 of the putative malarial targets. Seventy-nine (79) fragments have direct growth inhibition on Plasmodium falciparum at concentrations that are promising for the development of fragment hits against these protein targets. This adds a fragment library to the published HTS active libraries in the public domain.

KEYWORDS: natural products, fragments, native mass spectrometry, target identification, malaria

In 2015, malaria infected an estimated 212 million people globally, leading to 429,000 deaths, most of which were children aged 5 years and under in Africa.† Plasmodium falciparum is the most common species causing malaria in humans. Currently, four major classes of compounds are clinically used to treat malaria: (I) quinolines (e.g., chloroquine, quinine, mefloquine), (II) antifolates (e.g., sulfadoxine, pyrimethamine), (III) artemisinin derivatives (e.g., artesunate, artemether), and (IV) antimicrobials (e.g., doxycycline).‡ However, a combination of parasite drug resistance and vector resistance to insecticides remains a huge challenge in combating this disease.³ It is an urgent need that new classes of drugs be developed for the treatment of this pathogen.

Since 2008, almost 6 million compounds have been screened against the blood stages of P. falciparum, and approximately 0.5% of these compounds showed activity consistent with an EC₅₀ of less than 1 μM.⁴ Chemical structures and associated data have been made public from consortia led by GlaxoSmithKline (GSK), Novartis, and St. Jude Children’s Research Hospital, Memphis (St. Jude) to encourage further research into this disease.⁵−⁷ In addition, a subset of 400 compounds has been assembled into the Medicines for Malaria Venture Malaria Box that has catalyzed research into the discovery of new antimalarial clinical candidates.⁸−¹³ Finding a ligand for a disease-related protein remains one of the biggest hurdles of the postgenomic era. The identification of synthetic or natural ligands can help decrypt the biochemical functions of proteins with unknown functions.¹⁴ Such efforts require a method to detect the interaction between small molecular ligands and macromolecular targets.

Here, we describe a fragment-based approach using native mass spectrometry to detect protein–ligand interactions that exploit proteins produced by structural genomics efforts. The technique relies on non-denaturing electrospray ionization (ESI) to first recognize multicharged proteins in their native states. Native screening using high-resolution electrospray ionization Fourier transform cyclotron resonance mass spectrometry (ESI-FT-ICR-MS) is a label-free, fast, accurate method that permits the direct observation of ligand interactions that...
noncovalent protein—ligand complexes.\textsuperscript{15–17} In the resulting mass spectrum, the difference between the mass-to-charge ratio ($\Delta m/z$) for the protein—ligand complex and the unbound protein ions multiplied by the charge state ($z$) directly affords the molecular weight of the bound ligand ($\text{MW}_{\text{ligand}} = \Delta m/z \times z$). The molecular weight mass information on the ligand allows the identification of the active ligand in a pooled library. The simple mix-and-measure, label-free nature of these experiments makes native ESI-FT-ICR-MS a practical technique in screening fragments.

We previously reported the detection of protein—ligand interactions in complex natural product extracts.\textsuperscript{18–20} We also demonstrated that ESI-FT-ICR-MS can directly observe weak, noncovalent protein binding of fragment-sized natural products.\textsuperscript{21} Fragment-based drug discovery (FBDD) is an established alternative approach to high-throughput screening (HTS). Because of the simpler structure and smaller size of fragments, screening of fragment libraries can sample chemical space more efficiently and give higher hit rates in comparison to standard nonfragment chemical libraries.\textsuperscript{22,23} Although fragment hits generally have lower potency with initial affinities of up to 1 mM,\textsuperscript{24,25} their high ligand efficiency can offer more efficient and fruitful optimization campaigns, thus reducing attrition rates at all stages.\textsuperscript{36} FBDD has yielded two marketed drugs, vemurafenib\textsuperscript{27,28} and venetoclax,\textsuperscript{29} and there are many other compounds currently in clinical trials,\textsuperscript{30,31} including verubecestat (MK-8931) for Alzheimer’s disease, where the “intratable” Beta-APP cleaving enzyme (BACE) protein pockets that are long and shallow have proven difficult for traditional drug screenings to identify a hit.\textsuperscript{32}

Natural products, enriched with biosynthetic intermediates and endogenous metabolites resulting from being exposed to a long selection process to interact with biological targets, are excellent resources to search for protein binding partners. Native mass spectrometry provides key advantages for weak binding detection as it has high sensitivity, low sample consumption, does not require modifications or labeling of the protein target, and provides a direct visualization of all species present in solution under binding equilibrium.\textsuperscript{33,34}

Using this assay, we screened a fragment-sized library consisting of low-molecular-weight natural products and identified a number of natural products that bind to putative malaria targets. The hits were further investigated in biological activity assays to identify compounds that inhibit \textit{P. falciparum} and warrant hit-to-lead development. We have identified 96 fragments that bind to 32 potential targets and have demonstrated that 79 of these fragments can inhibit the growth of malaria parasites in vitro.

\section*{RESULTS AND DISCUSSION}

\textbf{Protein Selection.} The proteins were selected according to a variety of criteria from the analysis of \textasciitilde 5400 predicted open reading frames in the \textit{P. falciparum} genome. Initial analysis was restricted to proteins of less than 50 kDa as the 4.7 T FT-ICR-MS detects up to 50 kDa. This produced 2535 genes, of which there was strong evidence that 67 were essential in either the exoerythrocytic stage or intraerythrocytic stages or had demonstrated high expression in gametocytes or liver stages. The evidence for essentiality was based mainly on genetic knockouts\textsuperscript{35} but also involved pathway-based reasoning. For example, dihydroorototate dehydrogenase (DHODH) has been validated as a drug target,\textsuperscript{36} so other pyrimidine biosynthesis enzymes were also included. Of these 67 genes, 56 met the criterion of having been validated in an exoerythrocytic stage and/or being highly expressed in gametocyte\textsuperscript{37} and/or liver stages.\textsuperscript{38} Proteins that were expected to be difficult to express and/or purify were removed from the analysis, leaving a total of 49 prioritized candidates. From these prioritized genes, 13 proteins were successfully expressed and purified for this study. We attempted to clone and express all 49 proteins but were successful with only 13, broadly consistent with the low success rates previously reported for \textit{Plasmodium} proteins.\textsuperscript{39,40} The full list of these proteins is in Table S1. Another 66 proteins were identified from previous selection criteria\textsuperscript{41} as likely being essential for parasite survival and were amenable to recombinant expression. All proteins were soluble and therefore have a high likelihood of chemotherapeutic accessibility.

Seventy-nine \textit{Plasmodium} proteins were subjected to mass spectrometry under optimized conditions, yielding protein signals at low charged states between +8 and +18, depending on the molecular weight of the protein. Out of the 79 proteins, 62 proteins had a good signal-to-noise ratio (>100:1), indicating that native mass spectrometry is a powerful technology for recognizing protein in its native-like state (Table S2 and Figure S3).

\textbf{Fragment Library Assembly and Comparative Analysis of Physicochemical Properties, Shape, and Structural Diversity.} The physicochemical properties of most fragment libraries today are compliant with the rule of three (Ro3), initially proposed by Congreve et al. in 2003.\textsuperscript{42} However, both the structural and shape diversity of a fragment library are important additional factors to consider, and fragment libraries obtained synthetically or purchased from commercial suppliers are predominantly composed of small, flat heterocyclic molecules.\textsuperscript{31,43} Natural product fragments have different chemotypes and are more populated with three-dimensional (3D) structures, thus offering an opportunity to improve the structural and shape diversity of fragment libraries and consequently to search different regions of biologically relevant chemical space and find structurally different hits.\textsuperscript{32} We compared natural product fragments present in the Dictionary of Natural Products (DNP, sdf version 211.9) with commercial and synthetic fragments present in six different screening libraries (Asinex, ChemBridge, InterBioscreen (IBS), Molecular Libraries Small Molecule Repository (MLSMR), Life Chemicals, and Otava) using the following filtering criteria: molecular weight ($\text{MW}$) $\leq 250$ Da, octanol–water partition coefficient (A$\text{LogP}$) $< 4$, hydrogen bond donors (HBD) $\leq 4$, hydrogen bond acceptors (HBA) $\leq 5$, rotatable bonds (RB) $\leq 6$, percent polar surface area ($\%$PSA) $< 45$. After a cleaning process which included salt stripping, structure normalization and standardization, ionization at pH 7.4 and the selection of organic chemicals, and Otava) using the following filtering criteria: molecular weight ($\text{MW}$) $\leq 250$ Da, octanol–water partition coefficient (A$\text{LogP}$) $< 4$, hydrogen bond donors (HBD) $\leq 4$, hydrogen bond acceptors (HBA) $\leq 5$, rotatable bonds (RB) $\leq 6$, percent polar surface area ($\%$PSA) $< 45$. After a cleaning process which included salt stripping, structure normalization and standardization, ionization at pH 7.4 and the selection of organic molecules with $\text{MW}$ $\geq 100$ Da or heavy atom count $\geq 7$ and sulfur atom counts $\leq 3$,\textsuperscript{44} we found that only 6.7% of the fragments from DNP were considered identical (Tanimoto score of 1) to all of the unique fragments present in the six commercially sourced fragment libraries, as determined by pairwise similarity comparisons using two-dimensional (2D) extended-connectivity fingerprints ECFP_6. In agreement with Hert et al.,\textsuperscript{46} who noted that 83% of natural product ring scaffolds were absent in commercially available molecules, our substructure search analysis revealed that on average 91% of the scaffolds from the DNP fragments were absent in each of the commercially sourced fragment libraries.

Our natural product fragment library contains 643 molecules with fragment-like physicochemical properties. Of these
molecules, 602 (94%) comply strictly with the physicochemical property criteria initially set to guide the selection of these fragments (MW ≤ 250 Da, ALogP < 4, HBD ≤ 4, HBA ≤ 5, RB ≤ 6, %PSA < 45), while the remaining 41 constitute derivatives or structures that despite slightly violating one criterion or another were deemed of interest to screening. For the sake of accurate comparison to other generated fragment libraries, these 41 fragments were left out of the analyses discussed in this section, unless otherwise stated.

The mean distributions among the different fragment libraries for selected physicochemical properties (Figure S1) show that natural-product-based fragments (our 602 natural product fragments and fragments from DNP) have on average the same number of rings but a markedly smaller number of aromatic rings than commercially available fragments (fragments present in Asinex, ChemBridge, IBS, MLSMR, Life Chemicals, and Otava screening libraries). Our natural product fragment library presents itself with the highest average number of HBDs and the lowest average of ALogP and number of RBs (Figure S1).

To characterize the 3D shape of the fragments contained within each library, the fraction of sp3-hybridized carbon atoms (Fsp3) and principal moments of inertia (PMI) were determined for each molecule. The plot of the average of the calculated sum of the normalized PMI ratios (PMIsum) against the average Fsp3 clearly differentiates fragment libraries based on commercial and synthetic compounds from those based on natural products, which occupy a space where the average Fsp3 and average PMIsum are higher (Figure 1a). However, the benefits of populating a fragment library with more 3D-shaped fragments is still controversial.47−52 On one hand, 3D-shaped fragments have been suggested to be capable of sampling a larger chemical space40 and of exploring the pocket space of targets more extensively than 2D-shaped fragments,48 being more likely to deliver biologically active and clinically relevant molecules.51,52 On the other hand, 3D-shaped fragments have raised questions about the implications that their implicit higher molecular complexity can have for hit rates53 and for ease of synthesis and derivatization.44 It has also been shown that three-dimensionality can be introduced, to different degrees depending on the connectivity of the fragment scaffold, during fragment elaboration.47−49 While it is still unclear how important 3D character is for the core scaffold of fragments, we believe that shape diversity should exist in a fragment library. The 643-member natural product fragment library also contains a good representation of flat scaffolds (Bemis-Murcko molecular frameworks including rings, linkers, exocyclic double bonds, and terminal double bonds attached to the linkers), despite containing a higher proportion of 3D-shaped structures than other commercially available fragment libraries.

To assess the structural diversity of each fragment library, the percentage of unique fragment scaffolds (Bemis-Murcko molecular frameworks including rings, linkers, exocyclic double bonds, and terminal double bonds attached to the linkers) present in each library and the average number of unique fingerprint features per fragment (using 2D extended-connectivity fingerprints ECFP_6) were calculated. Figure 1b shows the plot of the former against the latter, demonstrating the high structural diversity achieved with our natural product fragment library.

Prevalence of Fragment Hit Scaffolds among Existing Antimalarial HTS Hits. In order to accelerate research efforts for antimalarial drug discovery, the structures of the hit molecules that resulted from different high-throughput screens (HTS) have been deposited in the public domain (http://www.ebi.ac.uk/chemblntd).

Herein, the structures provided by St. Jude’s Children’s Research Hospital,5 Novartis,55,56 GlaxoSmithKline,5 Harvard Medical School,52 and MMV58 were used to evaluate the representation of the scaffolds derived from our fragment hits within this antimalarial chemical space. Our 96 fragment hits contain 72 unique fragment scaffolds. Through a substructure search, each HTS hit was searched for in the presence of one or more of these fragment scaffolds. Twenty-six (36%) of the scaffolds were found among these libraries. While 8 of the fragment scaffolds are common to all antimalarial HTS hit libraries, some are unique to specific active antimalarial libraries (Figure S2). Table 1 gives examples of some fragment scaffolds represented among the antimalarial HTS hits as well as the originating fragment. The structures of the active drug-sized molecules clearly show that these could have been developed through FBDD by fragment linking, growing, or merging (Table 1). While this provides a validation of our fragment screening approach, it shows that, with 46 (64%) of the fragment hit scaffolds remaining to be mapped within an antimalarial HTS hit molecule, many opportunities exist for fragment elaboration into structurally new active drug-sized molecules. A PMI plot of the unique scaffolds present in our natural product fragment library (325 fragment scaffolds), where those present within fragment hits are discriminated...
from those absent, is shown in Figure 2. Within our fragment screening campaign, fragment hits were identified across all types of scaffold shapes.

Table 1. Examples of Fragment Hit Scaffolds Present among Antimalarial HTS Hits

| Fragment scaffold | Original fragment | Antimalarial HTS hit molecule | Ref. |
|-------------------|-------------------|------------------------------|------|
| ![Fragment scaffold 1](image1) | ![Original fragment 1](image2) | ![Antimalarial HTS hit molecule 1](image3) | 55 |
| ![Fragment scaffold 2](image4) | ![Original fragment 2](image5) | ![Antimalarial HTS hit molecule 2](image6) | 55 |
| ![Fragment scaffold 3](image7) | ![Original fragment 3](image8) | ![Antimalarial HTS hit molecule 3](image9) | 6 |
| ![Fragment scaffold 4](image10) | ![Original fragment 4](image11) | ![Antimalarial HTS hit molecule 4](image12) | 5 |

Figure 2. PMI of the fragment scaffolds (Bemis-Murcko molecular frameworks) contained in our fragment hit library (96 fragment hits, 72 unique fragment scaffolds) and comparison with those not contained in any fragment hit (253 unique fragment scaffolds). NPR1 constitutes the X axis, while NPR2 constitutes the Y axis.

Fragment Screening by Native MS. We performed a native MS screening using the natural product library against each of the 62 proteins. The complexes formed were ranked as...
strong, medium, or weak according to the ratio between the protein−ligand complex signal intensity and the unbound protein signal intensity (strong, ratio >0.5; medium, ratio 0.15 to 0.5; weak, ratio 0.02 to <0.15) (Figure 3).

This ratio does not provide an absolute measure of the binding affinity but can be used to obtain a relative affinity ranking and provide rough $K_D$ estimates. Assuming that the ionization, transmission, and detection efficiencies of the protein−ligand complex and free protein are the same and that relative equilibrium concentrations in solution are preserved upon transfer to the gas phase, the higher the intensity of the protein−ligand complex in the mass spectrum relative to the free protein, the higher the protein occupancy in solution and therefore the higher the ligand affinity for the protein. In our experience $^{7,20}$ and considering the range of protein and ligand concentrations used throughout this work (1−20 μM protein, 5−400 μM fragment), it is estimated that strong and medium binders have $K_D < 100 \mu M$. The categorization into strong, medium, and weak binding is particularly useful, together with hit rates, in assessing target ligandability, the ability of a protein target to bind small molecules with high affinity. $^{24}$ Taking into account the ligandability scoring system proposed by Edfeldt and colleagues, $^{24}$ we have defined high ligandable targets (with a high ligandability score) as those with $\geq 2$ strong/medium binders and $\geq 7$ hits, medium (at least 1 strong/medium binder and $\geq 2$ hits), and low (at least 1 hit). (Figure 4). The results clearly identify some targets that are more worthwhile to pursue on the basis of ligandability.

Ninety-six compounds (Figure 5) formed complexes with 32 proteins (Figure 6). Of the 96 fragments found to have protein
Figure 5. continued
partners, 48 fragments bound 1 protein, 18 bound 2 proteins, 14 bound 3 proteins, 4 bound 4 proteins, 5 bound 5 proteins, and 7 bound 6 proteins. Compounds that formed complexes with greater than 6 proteins were considered promiscuous and were excluded from further analysis. The retained hits are highly selective because they were screened against a total of 62 proteins: 48 fragments did not bind to 61 other proteins, 18 compounds did not bind to 60 other proteins, 14 compounds did not bind to 59 proteins, 4 compounds did not bind to 58 proteins, 5 compounds did not bind to 57 proteins, and 7 compounds did not bind to 56 proteins.

Although it is generally thought that molecules containing an electrophilic warhead (such as Michael acceptors) react indiscriminately and form covalent bonds with a variety of proteins, several studies have shown that such is not the case. A molecule containing a weak to moderate electrophilic warhead will bind covalently to a nucleophilic protein residue only if it has first bound noncovalently to the target protein and if the reactive electrophile has been properly oriented within the target protein so that a bond can be formed with the nucleophile on the protein. Many of our fragment hits contain electrophilic groups and are selective for a limited number of proteins (for example, see 15–20). Moreover, our experimental data shows that not all PAINS-containing molecules (such as polyphenols, β-lactams, epoxides, and rhodamine-related compounds) are promiscuous (for example, see 27, 50, 68, and 71).

To better display these data obtained from our MS screening, we built a map that shows the relationship between fragments and proteins (Figure 7). Considering hits and proteins as
nodes, we connected them to each other based on the binding interaction from Figure 6: the result is a network with 96 fragments. As previously mentioned, there were 48 selective fragments that bound a single protein so that 48 fragments bound more than 1 protein.

Figure 8 shows a limited part of the complete map with 54 fragments and 6 proteins connected to each other through 9 fragments. Twenty-seven hits are binding protein PF3D7_1127100 dUTPase, and six of these hits are joining fragments. Twenty-seven hits are binding protein fragments and 6 proteins connected to each other through 9 fragments. Twenty-seven hits are binding protein fragments and 6 proteins connected to each other through 9 fragments. As previously mentioned, there were 48 selective compounds against 31 of the 32 proteins (Figure 6, Table S3).

IC_{50} values were able to be calculated for 24 compounds, with 13 having IC_{50} < 45 \mu M. Compounds 5, 6, 7, 9, 27, 28, 85 have previously been reported to have anti-Plasmodial activity. Compound 9 had weak activity against P. falciparum 3D7 (61% inhibition at 100 \mu M) but had a previously reported IC_{50} of 9 \mu M against chloroquine-sensitive (D6, Sierra Leone) and chloroquine-resistant (W2) lines of P. falciparum. The IC_{50} values of the other compounds correlated with previously reported values. We have previously reported that 80, 81, and 87–89 bind allosterically to recombinant P. falciparum 2′-deoxyuridine 5′-triphosphate nucleotidohydrolase (PdUTPase, PF3D7_1127100), enhancing enzyme activity and inhibiting the in vitro activity viability of both P. falciparum gametocyte (sexual) and blood (asexual) stage parasites.

We identified 31 malaria proteins having a fragment–protein complex and in vitro activity against P. falciparum. These proteins warrant further investigation as possible drug targets in Plasmodium. Only one protein had a binder that did not show activity in the literature against asexual Plasmodium parasites at submicromolar concentration. The starting point for vemurafenib (27, 28) was an IC_{50} in the mM range with low affinity at ~200 \mu M. Similarly, the starting point for venetoclax (ABT-199) was two fragment hits with K_D = 400 and 2000 \mu M that were linked to initially give a compound with an IC_{50} of 6.9 \mu M on the way to the development of the approved drug.

Fragment-based drug discovery is based on the identification of low-affinity ligands. The starting point for vemurafenib (27, 28) was an IC_{50} in the mM range with low affinity at ~200 \mu M. Similarly, the starting point for venetoclax (ABT-199) was two fragment hits with K_D = 400 and 2000 \mu M that were linked to initially give a compound with an IC_{50} of 6.9 \mu M on the way to the development of the approved drug.
CONCLUSIONS

We demonstrated that enriching fragment screening libraries with natural-product-based fragments provides the means not only to increase the structural diversity of the library but also to improve the content of 3D-shaped structures. Increasing the structural diversity of fragment libraries with chemotypes usually absent in commercially available libraries has allowed searching different regions of biologically relevant chemical space and finding structurally different hits.

We have presented a native mass spectrometry method that can recognize native proteins and find a binder concurrently. We have demonstrated the value of native MS screening in fragment-based drug discovery on cloned and purified proteins. We identified 31 proteins worthy of further investigation as anti-plasmodial targets having 79 low-molecular-weight natural product hits with in vitro antimalarial activity out of a total of 96 natural product hits using native MS. This adds a fragment library to the already-published HTS active libraries in the public domain and may facilitate fragment-based drug discovery for malaria.

MATERIAL AND METHODS

Cheminformatics Analysis. Fragment Generation. Fragments from DNP (sdf version 211.9 166,379 molecules), Asinex (gold and platinum collection; 296,537 molecules), ChemBridge (express-pick collection stock; 459,091 molecules), InterBioscreen (IBS) (synthetic compounds library; 467,251 molecules), the Molecular Libraries Small Molecule Repository (MLSMR) (325,164 molecules), Life Chemicals (general HTS collection; 389,305 molecules), and Otava (collection for prompt delivery; 271,710 molecules) libraries were obtained using a series of filters available in Scitegic Pipeline Pilot (version 9.1.0.1, Accelrys Software, Inc.), which first cleaned up the structures (including salt stripping and selection of the largest fragment, structure normalization and standardization, ionization at pH 7.4, and selection of organic molecules with MW ≥ 100 Da or heavy atom count ≥ 7 and sulfur atom counts ≤ 3) and then selected those with fragmentlike physicochemical properties (MW ≤ 250 Da, atomic-based prediction of the octanol–water partition coefficient (A log P) < 4, HBD ≤ 4, HBA ≤ 5, rotatable.

Figure 7. Network visualization of 96 fragment hits against 32 proteins. The network was visualized using the Cytoscape software. Circle nodes represent fragment hits, and rectangular nodes represent proteins. The edge represents interaction between fragment hit and protein. The green edge indicates strong binding, blue indicates medium binding, and orange indicates weak binding, according to Figure 3.
bonds (RB) ≤ 6, percent polar surface area (%PSA) < 45). After the removal of duplicate molecules using canonical SMILES, 20,510 DNP fragments, 14,484 Asinex fragments, 24,291 ChemBridge fragments, 17,271 IBS fragments, 20,783 MLSMR fragments, 3,682 Life Chemicals fragments, and 17,769 Otava fragments were obtained.

**Determination of the Fraction of sp³-Hybridized Carbon Atoms (Fsp³) and Principal Moments of Inertia (PMI).** The fraction of sp³-hybridized carbon atoms (Fsp³) of each fragment was determined with Scitegic Pipeline Pilot (version 9.1.0.1, Accelrys Software, Inc.). Principal moments of inertia (PMI) were determined in Maestro (version 10.4.017, Schrodinger Inc., NY) after 3D structure minimization in Scitegic Pipeline Pilot. The normalized PMI ratios (NPR 1 and 2) were calculated by dividing the two lowest PMIs by the highest, as described by Sauer et al.⁷⁰ NPR1 was added to NPR2 to give the PMIsum.

**Generation of Fragment Scaffolds.** Within Scitegic Pipeline Pilot, the “generate scaffold tree” component was first used to generate hierarchical ring scaffolds according to the procedure described in ref ⁷¹. Then the first-level scaffolds, which correspond to the Bemis-Murcko molecular frameworks (rings, linkers, exocyclic double bonds, and terminal double bonds attached to the linkers), were selected.⁷² After the removal of duplicate molecules using canonical SMILES, the following number of unique fragment scaffolds were obtained: 3,958 DNP, 5,163 ChemBridge, 3,779 Asinex, 5,319 IBS, 13,74 Life Chemicals, 6,824 MLSMR, 3,704 Otava, and 325/72 of our global/hit fragment library. The percentage of unique fragment scaffolds was calculated by dividing the number of unique scaffolds by the number of fragments in the library.

**Determination of the Average Number of Unique Fingerprint Features per Fragment.** Using Scitegic Pipeline Pilot, each molecule was represented by its extended-connectivity fingerprint, ECFP_6, a 2D circular topological fingerprint where integer identifiers which account for different atom properties (atom type, mass, connection count, etc.) are first assigned to each heavy atom of the molecule and then are combined with identifiers of neighboring atoms in an iterative process until a diameter of six bonds is reached. This iteration process is based on the Morgan Algorithm⁷³ and converts the substructural features around each atom into integer values which constitute the extended-connectivity fingerprint. The average number of unique fingerprint features per fragment was determined by counting the total number of unique ECFP_6 fingerprint features and dividing it by the total number of molecules.

**Cleanup of Active Antimalarial HTS Hits.** St. Jude’s Children’s Research Hospital (1,523 molecules),⁶ Novartis (5,615 molecules),⁵⁶ GlaxoSmithKline (GSK) (13,467 mole-
Protein Cloning, Purification, and Purity Analysis. Recombinant histidine-tagged enzymes from *P. falciparum* and *P. vivax* were expressed in *Escherichia coli*. In general, genes were cloned into expression vectors that enabled tagging of the corresponding proteins with an N-terminal 6-histidine tag, as previously described. In some cases, truncated genes of the corresponding proteins with an N-terminal 6-histidine tag were cloned into expression vectors that enabled tagging.

Substructure Search. The fragment scaffolds were used as queries for the substructure search within the molecular structures of interest. Using Scitegic Pipeline Pilot, each query was mapped against all of the molecules to be searched. For each molecule successfully mapped, the number and name of queries mapped were retrieved.

Pairwise Similarity Comparisons. The similarity between two molecules was determined by calculating the similarity between their respective ECFP₆ fingerprints using the Tanimoto function: Tanimoto coefficient = c/(a + b − c), where a is the number of bits present in molecule A but not B, b is the number of bits present in molecule B but not A, and c is the number of bits present in both A and B.

Native MS Screening. Protein Preparation. The proteins were buffer-exchanged into a suitable volatile buffer (ammonium acetate or ammonium bicarbonate) under nearly physiological conditions (pH 7.0) using size exclusion chromatography prior to ESI–MS analysis. Depending on the protein, the buffer and its concentration were chosen to obtain the highest sensitivity in the mass spectrometer. Starting conditions were 10 μM protein in 10 mM volatile buffer.

Instrument Control and Acquisition. Experiments were performed on a Bruker Apex III 4.7T or a Bruker SolarIX XR 12T FT-ICR mass spectrometer (Bruker Daltonics Inc., Billerica, MA). An Agilent electrospay tuning mix (G2431A) was used for instrument calibration. Instrument parameters were tuned to avoid the dissociation of noncovalent protein–ligand complexes and achieve the highest signal intensities. As a control, a bovine carbonic anhydrase II (bCAII) sulfanilamide complex was used. Mass spectra were recorded in positive ion and profile modes with a mass range from 50 to 6000 m/z. Each spectrum was a sum of 16 transients (scans) composed of 256 K (Bruker Apex III 4.7T) or 1 M (Bruker SolarIX XR 12T) data points. All aspects of pulse sequence control and data acquisition were controlled by Xmass (Bruker Apex III 4.7T) or Solarix control (Bruker SolarIX XR 12T) software in a Windows operating system. For fragment library screening, pools of eight fragments (1 μL of each fragment at 5 mM in MeOH) were dried, resuspended in 1 μL MeOH, and incubated with proteins (1–20 μM) for 1 h at room temperature and analyzed by ESI-FT-ICR-MS. Final fragment/protein molar ratios varied between 5:1 and 20:1. When a noncovalent complex was found, the molecular weight of the binding fragment was calculated from the spectrum using the following equation: MW<sub>ligand</sub> = Δm/z × z. The binding of the individual active fragment was confirmed in a separate experiment (Figure 4).

Fragment Library. Approximately half of the compounds were previously isolated in our laboratory and are >90% pure. The other half were purchased. The cited papers (Table S3) are listed to provide easy access to the original publications that reported the compounds.

Asexual Blood Stage Assay. A drug-sensitive (3D7) and -resistant (Dd2) Pf line was to be tested using an established radiometric growth inhibition assay. Briefly, *P. falciparum*-infected erythrocytes (2.5% hematocrit, 0.5% parasitaemia) were incubated in 96-well plates containing test compounds or DMSO as a control. After 48 h, <sup>3</sup>H-hypoxanthine (0.5 microCi/well) was added, and the plates were incubated for another 24 h, followed by an assessment of <sup>3</sup>H incorporation.

**ASSOCIATED CONTENT**

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsinfecdis.7b00197.

Mean physicochemical properties of different fragment libraries, distribution of fragment hit scaffolds present within different libraries, and proteins studied for cloning and expression. Proteins studied by native mass spectrometry, detected hits and their antimalarial activity, and mass spectra of 62 *Plasmodium* proteins. (PDF)

**AUTHOR INFORMATION**

**Corresponding Author**

*E-mail: r.quinn@griffith.edu.au. Phone: +61 7 3735 6006.

**ORCID**

Ronald J. Quinn: 0000-0002-4022-2623

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

Funding was provided by the Bill and Melinda Gates Foundation (OPP1035218), the Australia Research Council (ARC) (LP120100485, LE20100170, and LE0237908), the National Health and Medical Research Council (NHMRC) (APP1046715), and National Institutes of Health (NIH) Grant 1P50 GM64655-01 “Structural Genomics of Pathogenic Protozoa (SGPP)”. We acknowledge the Australian Red Cross Blood Service for the provision of human blood and sera. We gratefully acknowledge access to the NatureBank biota library housed at the Griffith Institute for Drug Discovery (www.griffith.edu.au/gridd) for isolation of some of the fragments.

**REFERENCES**

1. WHO. World Malaria Report 2016, pp 1–186.
2. Cai, S. X., Risinger, A. L., Nair, S., Peng, J. N., Anderson, T. J. C., Du, L., Powell, D. R., Mooberry, S. L., and Cichewicz, R. H. (2016).
Identification of compounds with efficacy against malaria parasites from common North American plants. J. Nat. Prod. 79, 490–498.

(3) Roy, K. K. (2017) Targeting the active sites of malarial proteases for antimalarial drug discovery: approaches, progress and challenges. Int. J. Antimicrob. Agents 50, 287–302.

(4) Spangenberg, T., Burrows, J. N., Kowalczyk, P., McDonald, S., Wells, T. N. C., and Willis, P. (2013) The open access malaria box: a drug discovery catalyst for neglected diseases. PLoS One 8, e62006.

(5) Gamo, F. J., Sanz, L. M., Vidal, J., de Cozar, C., Alvarez, E., Lavandera, J. L., Vanderwall, D. E., Green, D. V. S., Kumar, V., Hasan, S., Brown, J. R., Peishoff, C. E., Cardon, L. R., and Garcia-Bustos, J. F. (2010) Thousands of chemical starting points for antimalarial lead identification. Nature 465, 305–310.

(6) Guiguemde, W. A., Shelat, A. A., Bouck, D. F., Duffey, S., Crowther, G. J., Davis, P. H., Smithson, D. C., Connelly, M., Clark, J. Z., Zhu, F. Y., Jimenez-Diaza, M. B., Martinez, M. S., Wilson, E. B., Tripathi, A. K., Gut, J., Sharlov, E. R., Bathurst, I., El Mazzouni, F., Fowble, J. W., Forquer, L., McGinley, P. L., Castro, S., Angulo-Barturen, I., Ferrer, S., Rosenthal, P. J., DeRisi, J. L., Sullivan, D. J., Lazo, J. S., Roos, D. S., Riscoe, M. K., Phillips, M. A., Rathod, P. K., Van Voorhis, W. C., Averv, V. M., and Guy, R. K. (2010) Chemical genetics of Plasmodium falciparum. Nature 465, 311–315.

(7) Meister, S., Plouffe, D. M., Kuhlen, K. L., Bonamy, G. M. C., Wu, T., Barnes, S. W., Bopp, S. E., Borboa, R., Bright, A. T., Che, J. W., Cohen, S., Dharia, N. V., Gagaring, M., Gettayacamin, M., Gordon, P., Groessl, T., Kato, N., Lee, C. M. S., McNamara, C. W., Fidock, D. A., Nagle, A., Nam, T. G., Tuntland, T., Walker, J. R., Zhou, Y. Y., Chatterjee, A., Diagana, T., Vigil, E., Vinetz, J. M., Vinh, P. V., Yu, H., Watanabe, N. A., Weatherby, K., White, P. M., Willks, A. F., Winzeler, E. A., Wojcik, E., Wree, M., Wu, W., Yokoyama, N., Zollo, P. H. A., Abla, N., Blasco, B., Burrows, J., Laleu, B., Leroy, D., Spangenberg, T., Wells, T., and Willis, P. A. (2016) Open source drug discovery with the malaria box compound collection for neglected diseases and beyond. PLoS Pathog. 12, e1005763.

(14) Carver, T. E., Berde, B., Cummings, M. D., Peprella, E. C., Pucci, M. J., Zawadzke, L. E., Dougherty, B. A., Tredup, J. A., Bryson, W. Y., Yanchunas, J., Jr., Doyle, M. L., Witmer, M. R., Nelen, M. L., Desjarlais, R. L., Jaeger, E. P., Devine, H., Asel, E. D., Springer, B. A., Bone, R., Salamme, F. R., and Todd, M. J. (2005) Decrypting the biochemical function of an essential gene from Streptococcus pneumoniae using thermofluor technology. J. Biol. Chem. 280, 11074–11072.

(15) Yu, H., Quinn, R. J., and Nguyen, N.-T. (2015) Bioaffinity mass spectrometry screening using droplet-based microfluidics. Micro Nano Lett. 7, 74–79.

(16) Vetting, M. W., Al-Obaidi, N., Zhao, S., San Francisco, B., Kim, J., Wichelecki, D. J., Bourrier, J. T., Solbiati, J. O., Yu, H., Zhang, X., Redonov, D. A., Love, J. D., Hillerich, B. S., Seidel, R. D., Quinn, R. J., Osterman, A. L., Cronan, J. E., Jacobson, M. P., Gerlt, J. A., and Almo, S. C. (2015) Experimental strategies for functional annotation and metabolism discovery: targeted screening of soluble binding proteins and unbiased panning of metabolomes. Biochemistry 54, 909–931.

(17) Pedro, L., Van Voorhis, W. C., and Quinn, R. J. (2016) Optimization of electrospray ionization by statistical design of experiments and response surface methodology: protein-ligand equilibrium dissociation constant determinations. J. Am. Soc. Mass Spectrom. 27, 1520–1530.

(18) Maresca, A., Temperini, C., Yu, H., Poulson, S. A., Scorzaffava, A., Quinn, R. J., and Supuran, C. T. (2009) Non-zinc mediated inhibition of carbonic anhydrases: coumarins are a new class of suicide inhibitors. J. Am. Chem. Soc. 131, 3057–3062.

(19) Yu, H., and Quinn, R. J. (2008) Direct screening of natural product extracts using mass spectrometry. J. Biomol. Screening 13, 265–275.

(20) Yang, B., Feng, Y. J., Yu, H., McCormick, B., Rowley, J., Pedro, L., Crowther, G. J., Van Voorhis, W. C., Forster, P. I., and Quinn, R. J. (2016) Bioaffinity mass spectrometry screening. J. Biomol. Screening 21, 194–200.

(21) Yu, H., Roullier, C., Campitelli, M., Trenholme, K. R., Gardiner, D. L., Andrews, K. T., Skinner-Adams, T., Crowther, G. J., Van Voorhis, W. C., and Quinn, R. J. (2013) Plasmodium gametocyte genotype Identification from a natural-product-based fragment library. ACS Chem. Biol. 8, 2654–2659.

(22) Hopkins, A. L., Groom, C. R., and Alex, A. (2004) Ligand efficiency: a useful metric for lead selection. Drug Discovery Today 9, 430–431.

(23) Kuntz, I. D., Chen, K., Sharp, K. A., and Kollman, P. A. (1999) The maximal affinity of ligands. Proc. Natl. Acad. Sci. U. S. A. 96, 9997–10002.
A combined transcriptome and proteome survey of malaria parasite liver stages. *Proc. Natl. Acad. Sci. U. S. A.* 105, 305–310.

39. Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A. P., Wasney, G. A., Gao, M., Hills, T., Broxk, S., Qin, W., Sharma, S., Diassi, A., Alam, Z., Melone, M., Mulichak, A., Wernimont, A., Bray, J., Loppnau, P., Plotnikova, O., Newberry, K., Sundararajan, E., Houston, S., Walker, J., Temple, W., Bochkarev, A., Kozieradzki, L., Edwards, A., Arrowsmith, C., Roos, D., Kain, K., and Hui, R. (2007) Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms. *Mol. Biochem. Parasitol.* 151, 100–110.

40. Mehlin, C., Boni, E., Buckner, F. S., Engel, L., Feist, T., Gelb, M. H., Haji, L., Kim, D., Liu, C., Mueller, N., Myler, P. J., Reddy, J. T., Sampson, J. N., Subramanian, E., Van Voorhis, W. C., Worthey, E., Zucker, F., and Hol, W. G. J. (2006) Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. *Mol. Biochem. Parasitol.* 148, 144–160.

41. Fan, E., Baker, D., Fields, S., Gelb, M. H., Buckner, F. S., Van Voorhis, W. C., Phizicky, E., Dumont, M., Mehlin, C., Grayhack, E., Sullivan, M., Verlinde, C., Detitta, G., Meldrum, D. R., Merritt, E. A., Earnest, T., Soltis, M., Zucker, F., Myler, P. J., Schoenfeld, L., Kim, D., Worthey, L., Lacoant, D., Vignali, M., Li, J., Mondal, S., Massey, A., Carrell, R., Gulde, S., Luft, J., Desoto, L., Holl, M., Caruthers, J., Bosch, J., Robien, M., Arakaki, T., Holmes, M., Le Trong, L., and Hol, W. G. (2008) Structural genomics of pathogenic protozoa: an overview. *Methods Mol. Biol.* 426, 497–513.

42. Congreve, M., Carr, R., Murray, C., and Jhoti, H. (2003) A rule of three for fragment-based lead discovery? *Drug Discovery Today* 8, 876–877.

43. Scott, D. E., Coyne, A. G., Hudson, S. A., and Abell, C. (2012) Fragment-based approaches in drug discovery and chemical biology. *Biochemistry* 51, 4900–5003.

44. Over, B., Wetzol, S., Grutter, C., Nakai, Y., Renner, S., Raub, D., and Waldmann, H. (2013) Natural-product-derived fragments for fragment-based ligand discovery. *Nat. Chem.* 5, 21–28.

45. Pascuolotti, M., Campitelli, M., Nguyen, B., Pham, N., Gorse, A.-D., and Quinn, R. J. (2015) Capturing nature's diversity. *PLoS One* 10, e0120942.

46. Hert, J., Irwin, J. J., Laggner, C., Keiser, M. J., and Shoichet, B. K. (2009) Quantifying biogenic bias in screening libraries. *Nat. Chem. Biol.* 5, 479–483.

47. Meyers, J., Carter, M., Mok, N. Y., and Brown, N. (2016) On the origins of three-dimensionality in drug-like molecules. *Future Med. Chem.* 8, 1753–1767.

48. Fuller, N., Spadola, L., Cowen, S., Patel, J., Schonherr, H., Cao, Q., McKenzie, A., Edfeldt, F., Rabow, A., and Goodnow, R. (2016) An improved model for fragment-based lead generation at AstraZeneca. *Drug Discovery Today* 21, 1272–1283.

49. Hall, R. J., Mortenson, P. N., and Murray, C. W. (2014) Efficient exploration of chemical space by fragment-based screening. *Prog. Biophys. Mol. Biol.* 116, 82–91.

50. Morley, A. D., Pugliese, A., Birchall, K., Bower, J., Brennan, P., Brown, N., Chapman, T., Drysdale, M., Gilbert, I. H., Hoelder, S., Jordan, A., Ley, S. V., Merritt, A., Miller, D., Swarbrick, M. E., and Wyatt, P. G. (2013) Fragment-based hit identification: thinking in 3D. *Drug Discovery Today* 18, 1221–1227.

51. Mok, N. Y., Brench, R., and Brown, N. (2014) Increasing the coverage of medicinal chemistry-relevant space in commercial fragments screening *J. Chem. Inf. Model.* 54, 79–85.

52. Aldeghi, M., Malhotra, S., Selwood, D. L., and Chan, A. W. E. (2014) Two-and three-dimensional rings in drugs. *Chem. Biol. Drug Des.* 83, 450–461.

53. Hann, M. M., Leach, A. R., and Harper, G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* 41, 856–864.

54. Keseru, G. M., Erlanson, D. A., Ferencey, G. G., Hann, M. M., Murray, C. W., and Pickett, S. D. (2016) Design principles for fragment libraries: maximizing the value of learnings from pharma
fragment-based drug discovery (FBDD) programs for use in Academia. J. Med. Chem. 59, 8189–8206.

(55) Plouffe, D., Brinker, A., McNamara, C., Henson, K., Kato, N., Kuhen, K., Nagle, A., Adrian, F., Matzen, J. T., Anderson, P., Nam, T. G., Gray, N. S., Chatterjee, A., Janes, J., Yan, S. F., Trager, R., Caldwell, J. S., Schultz, P. G., Zhou, Y., and Winzeler, E. A. (2008) In silico activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. Proc. Natl. Acad. Sci. U. S. A. 105, 9059–9064.

(56) Gagaring, K., Borboa, R., Francek, C., Chen, Z., Buenviaje, J., Plouffe, D., Winzeler, E., Brinker, A., Diagana, T., Taylor, J., Glynne, R., Chatterjee, A., and Kuhen, K. Novartis-GNF Malaria Box; Genomics Institute of the Novartis Research Foundation (GNF), 2010; www.ebi.ac.uk/chemblntd.

(57) Derbyshire, E. R., Prudencio, M., Mota, M. M., and Clardy, J. (2012) Liver-stage malaria parasites vulnerable to diverse chemical scaffolds. Proc. Natl. Acad. Sci. U. S. A. 109, 8511–8516.

(58) Avery, V. M., Bashyam, S., Burrows, J. N., Duffy, S., Papadatos, G., Puthukkuti, S., Sambandan, Y., Singh, S., Spangenberg, T., Waterston, D., and Willis, P. (2014) Screening and hit evaluation of a chemical library against blood-stage Plasmodium falciparum. Malar. J. 13, 190.

(59) Singh, J., Petter, R. C., Baillie, T. A., and Whitty, A. (2011) The resurgence of covalent drugs. Nat. Rev. Drug Discovery 10, 307–317.

(60) Mah, R., Thomas, J. R., and Shafer, C. M. (2014) Drug discovery considerations in the development of covalent inhibitors. Bioorg. Med. Chem. Lett. 24, 33–39.

(61) Baell, J. B., and Holloway, G. A. (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. J. Med. Chem. 53, 2719–2740.

(62) Baell, J. B. (2016) Feeling nature’s PAINS: natural products, natural product drugs, and pan assay interference compounds (PAINS). J. Nat. Prod. 79, 616–628.

(63) Shannon, P., Markiel, A., Ozier, O., Baliga, N., Wang, J., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

(64) Kaur, K., Jain, M., Kaur, T., and Jain, R. (2009) Antimalarials from nature. Bioorg. Med. Chem. 17, 3229–3256.

(65) Bajsa, J., Singh, K., Nanayakkara, D., Duke, S. O., Rimando, A. M., Evidente, A., and Tekwani, B. L. (2007) A survey of synthetic and natural pyrrolopyrazine and pyrroloazepine analogs as potential antimalarial compounds. J. Nat. Prod. 70, 1740–1744.

(66) Gengenbacher, M., Fitzpatrick, T. B., Raschle, T., Flicker, K., Sinning, I., Muller, S., Machereux, P., Tews, I., and Kappes, B. (2006) Vitamin B6 biosynthesis by the malaria parasite Plasmodium falciparum - Biochemical and structural insights. J. Biol. Chem. 281, 3633–3641.

(67) Bajsa, J., McCluskey, A., Gordon, C. P., Stewart, S. G., Hill, T. A., Sahu, R., Duke, S. O., and Tekwani, B. L. (2010) The antiplasmodial activity of norcantharidin analogs. Bioorg. Med. Chem. Lett. 20, 6688–6695.

(68) Suberu, J. O., Gorka, A. P., Jacobs, L., Roepe, P. D., Sullivan, N., Barker, G. C., and Lapkin, A. A. (2013) Anti-Plasmodial polyvalent interactions in Artemisia annua L. aqueous extract - possible synergistic and resistance mechanisms. PLoS One 8, e80790.

(69) Calderon, A. I., Romero, L. L., Ortega-Barria, E., Bruen, R., Correa, A. M. D., and Gupta, M. P. (2006) Evaluation of larvicidal and in vitro antiparasitic activities of plants in a biodiversity plot in the Altos de Campana National Park, Panama. Pharm. Biol. 44, 487–498.

(70) Sauer, W. H. B., and Schwarz, M. K. (2003) Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. J. Chem. Inf. Comput. Sci. 43, 987–1003.

(71) Schuffenhauer, A., Ertl, P., Roggo, S., Wetzel, S., Koch, M. A., and Waldmann, H. (2007) The scaffold tree - visualization of the scaffold universe by hierarchical scaffold classification. J. Chem. Inf. Model. 47, 47–58.

(72) Bemis, G. W., and Murcko, M. A. (1996) The properties of known drugs. 1. molecular frameworks. J. Med. Chem. 39, 2887–2893.

(73) Morgan, H. L. (1965) Generation of a unique machine description for chemical structures-a technique developed at chemical abstracts service. J. Chem. Doc. 5, 107–113.