Antimalarial Drug Discovery: *In Silico* Structural Biology and Rational Drug Design

TAP de Beer[1,#], GA Wells[1,#], PB Burger[1,#], F Joubert[1], E Marechal[2], L Birkholtz[1] and AI Louw[1,*]

[1]*Bioinformatics and Computational Biology Unit, Department of Biochemistry, School of Biological Sciences, University of Pretoria, Pretoria 0002, South Africa,* [2]*UMR 5168 CNRS-CEA-INRA-Université Joseph Fourier, Laboratoire de Physiologie Cellulaire Végétale; Institut de Recherches en Technologies et Sciences pour le Vivant; CEA Grenoble, 17 rue des Martyrs, F-38054, Grenoble cedex 09, France*

Abstract: Malaria remains one of the most burdensome human infectious diseases, with a high rate of resistance outbreaks and a constant need for the discovery of novel antimalarials and drug targets. For several reasons, Plasmodial proteins are difficult to characterise structurally using traditional physical approaches. However, these problems can be partially overcome using a number of *in silico* approaches. This review describes the peculiarities of malaria proteins and then details various *in silico* strategies to select and allow descriptions of the molecular structures of drug target candidates as well as subsequent rational approaches for drug design. Chiefly, homology modelling with specific focus on unique aspects of malaria proteins including low homology, large protein size and the presence of parasite-specific inserts is addressed and alternative strategies including multiple sequence and structure-based prediction methods, sampling-based approaches that aim to reveal likely global or shared features of a Plasmodial structure and the value of molecular dynamics understanding of unique features of Plasmodial proteins are discussed. Once a detailed description of the drug target is available, *in silico* approaches to the specific design of an inhibitory drug thereof becomes invaluable as an economic and rational alternative to chemical library screening.

Keywords: Malaria, antimalarials, pharmacophore, structure-based drug discovery, *in silico.*

1. BACKGROUND

More than 2 billion people are at risk of malaria and current clinical episodes are estimated to be as high as 500 million cases and nearly 3 million deaths per year, mainly children and pregnant women in resource-poor environments [1]. The scale of the malaria problem emphasises the fragile nature of prevailing control programmes and the importance of developing more effective methods for the prevention, treatment and ultimately, the eradication of malaria [2]. Combating malaria (caused by *Plasmodium* species) requires significant financial and organizational resources, yet malaria itself restrains economic development, creating a vicious cycle in developing countries [3-5].The devastating socio-economic and public health impact of malaria is mostly experienced in sub-Saharan Africa and is galvanized by the emergence and rapid spread of drug-resistant parasites and the lack of a licensed vaccine. Even if and when effective vaccines do become available, chemotherapy will still be required. The discovery of new and robust anti-malarial drugs, preferably acting on new targets that have not mutated yet into resistant forms, is therefore urgently needed. However, because malaria is considered a disease of poverty, there is very little incentive for pharmaceutical companies to partake in a global antimalarial effort except for provision of funds for selected projects and research institutes [6]. The onus therefore falls predominantly on publicly funded research groups, academic institutions and public-private partnerships established since 2000, to identify and develop novel antimalarial strategies [7].

In general, there has been a steady decline in the number of new molecular entities entering clinical development and reaching the market over the past 10-15 years due to high levels of drug attrition mainly attributed to unanticipated efficacy and toxicity problems [8]. Part of the blame seems to reside in the extensive use of High-Throughput Screening (HTS) against ambiguous or single targets which in effect reduces the biological context by separating the target from other cellular proteins and processes that might impact its function [9] and lack of diversity in existing chemical libraries [10]. The phenotypic robustness of biological systems often reduces the effectiveness of a single-target compound [11]. One compound-one target strategies therefore need to be adapted and it is suggested that the focus should be on promiscuous compounds that modulate multiple target proteins to achieve the desired results [12]. Cell-based high content screening (HCS) circumvents this problem, since it allows the detection of small molecules acting in the cellular context [13], but it leaves the question of the actual target unresolved. HTS technology is often limited to big pharmaceutical companies due to the high cost involved in screening of targets but is also limited by high attrition, with a hit-rate of between 0.01-1% of compounds screened [14]. The process is sequential, with ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicology) properties determined later on in the discovery

*Address correspondence to this author at the Bioinformatics and Computational Biology Unit, Department of Biochemistry, School of Biological Sciences, University of Pretoria, South Africa; Tel: +27 12 420 2480; Fax: +27 12 362 5302; E-mail: braam.louw@up.ac.za

#These authors contributed equally to this manuscript as part of their PhD studies.

1871-5265/09 $55.00+.00

process. Many molecules that look promising in early stages of the process fail later on because of weak ADMET properties [15]. A common view is that disease or biology relevant screens should be introduced much earlier into the drug discovery process [15-17]. Strategies to improve the drug discovery process include models to globally map pharmacological space, the introduction of functional and structural genomics for target validation and MOA prediction [18], and *in silico* approaches at the very early stages of the discovery pipeline [19-21].

Current medium to longer-term strategies in antimalarial drug discovery programs have been extensively reviewed and include "piggy-backing" strategies (screening of existing drugs for novel antiparasite activity or screening of lead series already applied to orthologous targets in other disease settings), *de novo* drug discovery involving target-based HTS or medium throughput screening (MTS) against parasite cultures, use of reformulated combinations of existing antimalarial drugs and information on established networks for compound screening and medicinal chemistry and pharmacokinetics/metabolism [5-8, 22]. Additional recent approaches include chemical scaffold alterations, the use of bioprecursors, double drugs, multiple targets or natural products [23]. The same concerns referred to above with respect to the quality of the chemical libraries and chosen molecular targets, are equally applicable to antimalarial drug discovery. However, carefully defined chemical compound libraries are being established (see e.g. http://www.drugdiscovery.dundee.ac.uk [24]) and state-of-the-art criteria for target assessment for antiparasitic drug discovery has been reviewed [25].

Drug discovery strategies, defined as the investigative work leading up to the selection of drug candidates, have been extensively described. As in other organisms, the same iterative process focussing around target identification, lead identification and lead optimisation applies to antiparasitic drug discovery (Fig. (**1**)). The proposed drug discovery pipeline as applied to parasitic diseases including malaria has been reviewed [5, 26-31]. The inherent challenges of organizing and mining of malaria genomic and post-genomic data, *in silico* applications of the embedded information and the necessity of integration of *in vivo* and *in vitro* data has been comprehensively described [22]. This review will focus on the use of *in silico* approaches to antimalarial drug discovery and discuss each component of this highly iterative process. The potential and caveats associated with the *in silico* discovery of antimalarial drug targets have been highlighted [21] and only new developments related to this will be described. Detailed descriptions of the application of *in silico* structure-based drug discovery (SBDD) in antimalarial lead identification and lead optimisation steps have not been comprehensively reviewed elsewhere and are therefore the major focus of this review (Fig. (**1**)).

2. *IN SILICO* DRUG AND DRUG TARGET DISCOVERY IN MALARIA

Anti-parasitic drug and drug target discovery has mostly benefited from prior knowledge of beneficial or effective herbal extracts or compounds (e.g. quinine and artemisinin) with the targets of these compounds initially unknown. To our knowledge, none of the currently-used pharmaceuticals has originated from a *de novo* approach based on rationally-selected targets and most recent antimalarial drugs are principally based on compounds derived from plant extracts. Breaking with this modern exploitation of traditional pharmacopoeia, the next generation of drugs is expected to be derived from diversity-oriented synthetic chemistry and screening programs on characterized new targets.

The completed sequencing of the *P. falciparum* genome sparked great hopes regarding the use of genomic information in the identification of new drug target proteins and the development of new drugs. This has been complicated by the relatively poor annotation of the malaria parasite genome, and thus it should be emphasized that absence of a protein from the genome annotation does not necessarily imply its absence from the organism. One of the first papers

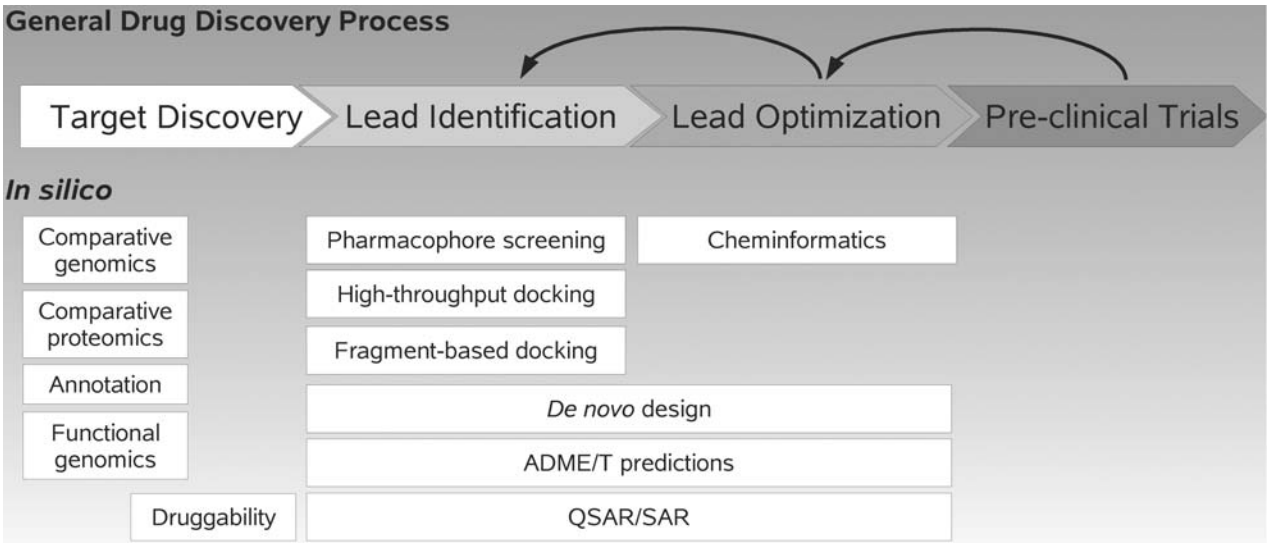

Fig. (1). The drug discovery pipeline with a focus on *in silico* strategies. These strategies are used in an integrated and iterative manner.

on the specific impact and value of the sequencing of the malaria genome on drug discovery was by Joachimiak *et al*. in 2001, where the analysis of different falcipains and the impact thereof was discussed [32]. In another well-known example, a series of plant-like genes were identified as possible new drug targets by the group of McFadden *et al*. [33]. A case in point in which the genome information contributed to a drug currently in clinical trials is that of fosmidomycin, which inhibits 1-deoxy-D-xylulose 5-phosphate (DXP) reductoisomerase, a key enzyme in the non-mevalonate pathway. In this case, a legacy drug was used to target a protein which was targeted from malaria genome data, based on information available in bacteria [34]. Apicoplast-related targets include proteins of the transcriptional/translational machinery (targets for lincosamides, e.g. clindamycin; macrolides, e.g. azithromycin, thiopeptides, e.g. thiostrepon, micrococcin), proteins involved in DNA replication (quinolones and fluoroquinolones), the fatty acid synthase of type II (target for aryloxyphenoxypropionate - fop- herbicides, e.g. haloxyfop, clodinafop, quizlofop, diclofop; fenoxaprop; tralkoxymid; thiolactomycin and analogues), the peptide deformylase (target for actinomycin) [35]. All these targets and compounds are currently investigated for the development of optimized drug candidates by various groups worldwide.

Some of the main criteria used in the *in silico* identification of putative parasite drug targets may include, but is not limited to, selecting the aspect of the parasite's biology to be interfered with; finding proteins or protein orthologues with sequence, functional and structural properties of interest; determining the level of conservation with host orthologues which may affect cross-reactivity; defining the classes of compounds that the proteins interact with; analyzing the druggability of the protein active site and validating the protein as a suitable target or choosing targets that have been clinically validated in other species [25]. Various types of antimalarial drug targets including proteins may be investigated but the role of rational criteria (including well-organised Boolean criteria) in the discovery of targets should not be underestimated [21, 36]. The sparse annotation of the malaria genomes, and the relatively poor performance of existing approaches and analysis methods on malaria data is still currently limiting and necessitates the need for more attention on the "druggability" of proteins [37]. This information provides the basis for drug target databases including the TDR Targets Database (http://tdrtargets.org) [38]. Systems level understanding of the malaria parasite is highly challenging but remains of critical importance in drug discovery endeavours [18, 39, 40]. This could be particularly evident in the case of the, as yet, unannotated or 'hypothetical' proteins of *P. falciparum*.

Experimental and *in silico* interactome data has revealed unique protein-protein networks within the malaria parasite with controlling nodes indicating a 'rich-club phenomenon' of interconnectivity [41]. Computational prioritization of drug targets is utilised in PlasmoCyc containing an integrated pathway/genome database and resulted in the identification of 216 chokepoint enzymes [42]. More recently, the *P. falciparum* metabolic pathways [43] have been used to identify an additional 22 potential new targets using *in silico* knock-out approaches [44]. Particular problems in malaria includes the limited efficiency of homology-based methods (like BLAST and HMMER) to assign functionality to the more than 60% of the malaria proteome. Innovative new approaches resulted in the creation of GO-databases, *i.e*. the *Plasmodium* OPI Databases [21, 45] or PlasmoDRAFT, containing annotated predictions based on guilt-by-association methods using post-genomic data including that from the transcriptome, proteome and interactome [46]. Alternative resources aim at utilising structural, functional and interaction features to allow druggability descriptions of malaria proteins. The Structural Annotation of Malaria Proteins (SAMP) project attempts to additionally provide information regarding possible ligands as established by the Small Molecule Interaction Database (SMID) software suite (Unleashed Informatics) [47].

However, for effective target identification, the integration of protein annotation information with existing chemoinformatics resources is of critical importance. There is little ligand information available in PlasmoDB (http://www.plasmodb.org), and while the TDR database contains some ligand information, it is not focused at enabling searches based on chemical compounds. The mapping of existing drugs to targets together with the related beneficial drug and target characteristics is of paramount importance, as the elucidation of as many key characteristics as possible will be advantageous in the selection of new targets as well as leads [48]. The development of a system integrating detailed structural and functional information for malaria proteins with information for chemical compounds, providing at least a basic chemoinformatics environment, would prove to be a huge asset to antimalarial drug development [49]. A selection of some key *in silico* resources is provided in Table **1**. A recent review provides comprehensive lists of free web resources [50].

Table 1. Selected *In Silico* Resources for Malaria Drug Discovery

Resource	URL	Reference
PlasmoDB	http://www.plasmodb.org	[161]
TDR Targets Database	http://tdrtargets.org	[38]
Malaria Parasite Metabolic Pathways	http://sites.huji.ac.il/malaria	[43]
PlasmoCyc	http://plasmocyc.stanford.edu	[42]
PlasmoDRAFT	http://atgc.lirmm.fr/plasmo_draft	[46]
Structural Genomics Consortium	http://sgc.utoronto.ca	[162]
Structural Genomics of Pathogenic Protozoa	http://www.sgpp.org	[118]
SAMP	http://malport.bi.up.ac.za	[47]
PubChem	http://pubchem.ncbi.nlm.nih.gov	[163]
Super Drug DataBase	http://bioinf.charite.de/superdrug	[97]
DrugBank	http://www.drugbank.ca	[98]
ZINC	http://zinc.docking.org	[96]
WISDOM	http://wisdom.eu-egee.fr	[104]

3. *IN SILICO* STRUCTURE-BASED DRUG DISCOVERY IN MALARIA

In silico structure-based drug design (SBDD) aims at rational or knowledge-driven descriptions of new inhibitory compounds and can be classified into receptor-based design and ligand-based design (Fig. (**2**)). In particular, receptor-based design exploits the three-dimensional structural description of a drug target to predict the *in silico* binding of hypothetical ligands to the target. These hypothetical ligands can be obtained from the modification of a ligand known to bind to the target and include fragment-based inhibitor design (scaffold structures or *de novo* design), receptor-based pharmacophore design and virtual screening of *in silico* compound libraries against the target (Fig. (**2**)).

Ligand-based design aims to predict the effect of new compounds based on the properties of compounds known to affect the target. This may be pursued in the absence of a target structure. Various virtual screening methodologies exist for lead identification and include receptor-based pharmacophores, HTS and fragment-based design. Structure-activity relationships (SAR and quantitative SAR), chemo-informatics and toxicity predictions (ADMET) should be evaluated, in addition to screening procedures and fragment assemblies based on medicinal chemistry principles. Three questions which are pivotal in deciding on a particular methodology to be followed include: 1) are molecules available which can be modified to become inhibitors, 2) is there a means for synthesizing novel molecules and 3) what is the degree of accuracy required at a particular stage of the design process versus the time needed for the calculations [51]? The latter include factors such as either protein or ligand flexibility or both, the inclusion of solvent effects, etc. [51]. These questions are not stage-specific but should be asked continuously during the drug discovery process.

Both of the abovementioned design avenues are highly integrated, iterative and knowledge-based and all sub-strategies should be investigated. The knowledge available on both the structure and inhibitors of a specific target largely determines the approach to be followed. Identified compounds are scored and ranked based on their physio-chemical interactions with the target structure and the best scoring compounds are biochemically tested for inhibitory activity. Promising lead compounds with low micromolar activities are then optimized by solving the structure of the target-lead complexes to confirm predicted data. This is followed by *in silico* optimization of the lead compound and iterative testing.

3.1. Receptor-Based Drug Design

Receptor-based drug design entails the use of 3D target structures to reduce the chemical search space and provides the molecular framework representative of the essential physiochemical features required for biological activity of the inhibitory compound. The advantage this provides has made it the preferred approach for *in silico* drug design and has further gained popularity due to the increase in computing power and wide variety of software suites available. This strategy starts by describing the 3D structure of the target of interest. However, Plasmodial protein structures have been difficult to solve experimentally [18] and therefore make recent advances in computational techniques for the description of protein structures that much more enticing to the malaria community.

3.1.1. Receptor-Based Drug Design: In Silico Approaches to Obtaining Plasmodial Protein Structures

The majority of malaria drug targets are proteins, as are all the potential novel targets currently under investigation, but these are notoriously difficult to express in heterologous systems [41]. Compounding characteristics of proteins from *P. falciparum* include large protein sizes, greater protein disorder, more basic pI than host systems, low-complexity containing parasite-specific inserted regions and a marked A+T bias of the *P. falciparum* genome. These factors additionally may contribute towards low crystallisation efficiencies of Plasmodial proteins. At the time of writing, querying the PDB (http://www.pdb.org) for structures of Plasmodial proteins and excluding sequences with greater than 90% identity, yields 118 entries. A closer inspection of all released Plasmodial protein structures reveals 100 orthologues from multiple *Plasmodium* species. In contrast, querying the PDB for human protein entries (excluding > 90% sequence identity) reveals more than 4500 structures. Even though the number of Plasmodial protein structures is still alarmingly sparse, there has been an almost doubling in Plasmodial protein structures since 2005, largely due to the advent of structural genomics programs including the Structural Genomics Consortium, (http://sgc.utoronto.ca) and the Structural Genomics of Pathogenic Protozoa (http://www.sgpp.org). The Structural Genomics Consortium (SGC) reported 25 distinct Plasmodial protein crystal structures from five species. The success rate of this study is similar to other structural genomics programs, and demonstrates the viability of structural genomics for protozoa. This was partly due to treating orthologues from multiple species as alternative expression constructs [52]. The SGPP Consortium has solved 40 structures from the parasitic organisms *Leishmania*, *Trypanosoma brucei*, *T. cruzi* and *Plasmodium* of which 16 are Plasmodial proteins. The success is attributed to pioneering a number of developments such as domain prediction, the use of co-crystallents, capillary crystallization and "fragment cocktail crystallography".

In lieu of the paucity of crystal structures for Plasmodial proteins, many groups have resorted to homology modelling. The number of studies employing this technique are too numerous to cite, however, reviewing them reveals certain trends. The modelling component usually makes a small contribution to the overall study and models are typically constructed to aid visualisation and rationalisation of experimental results. Identification of novel inhibitors using these models are often mentioned, however, there seems to be little experimental follow-up to pre-clinical phases. In few cases, popular targets for homology modelling have been superseded by recent experimental structures, notably proteases specific to *Plasmodium*.

Successful homology modelling depends critically on the alignment of the target sequence with template structures. A review of the present literature suggests that targets chosen for modelling tend to align unambiguously or easily with templates. However, Plasmodial proteins are infamous for

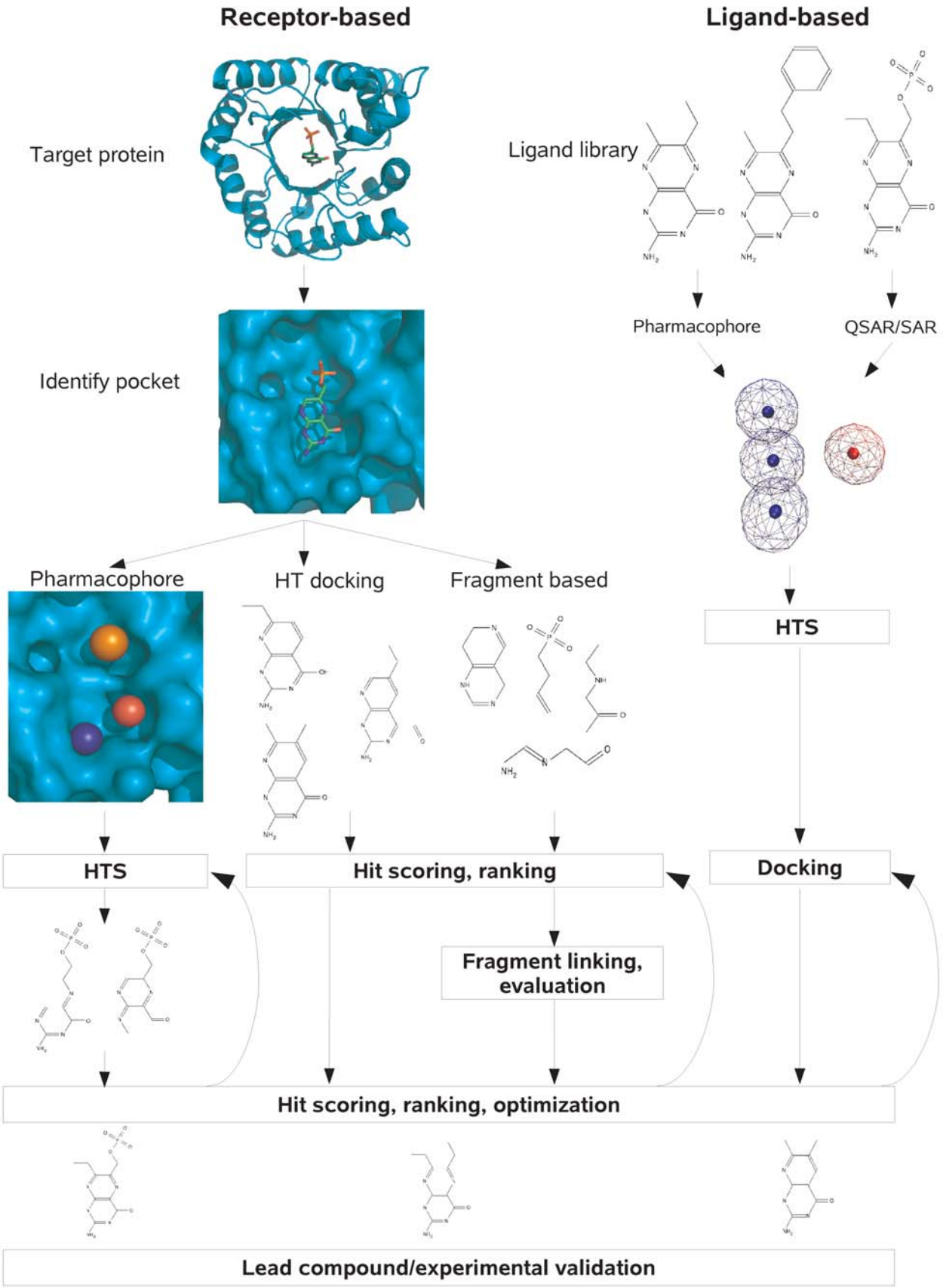

Fig. (2). Structure-based drug design as applied to the discovery of antimalarials. Parallel and integrative strategies include receptor-based design and ligand-based design.

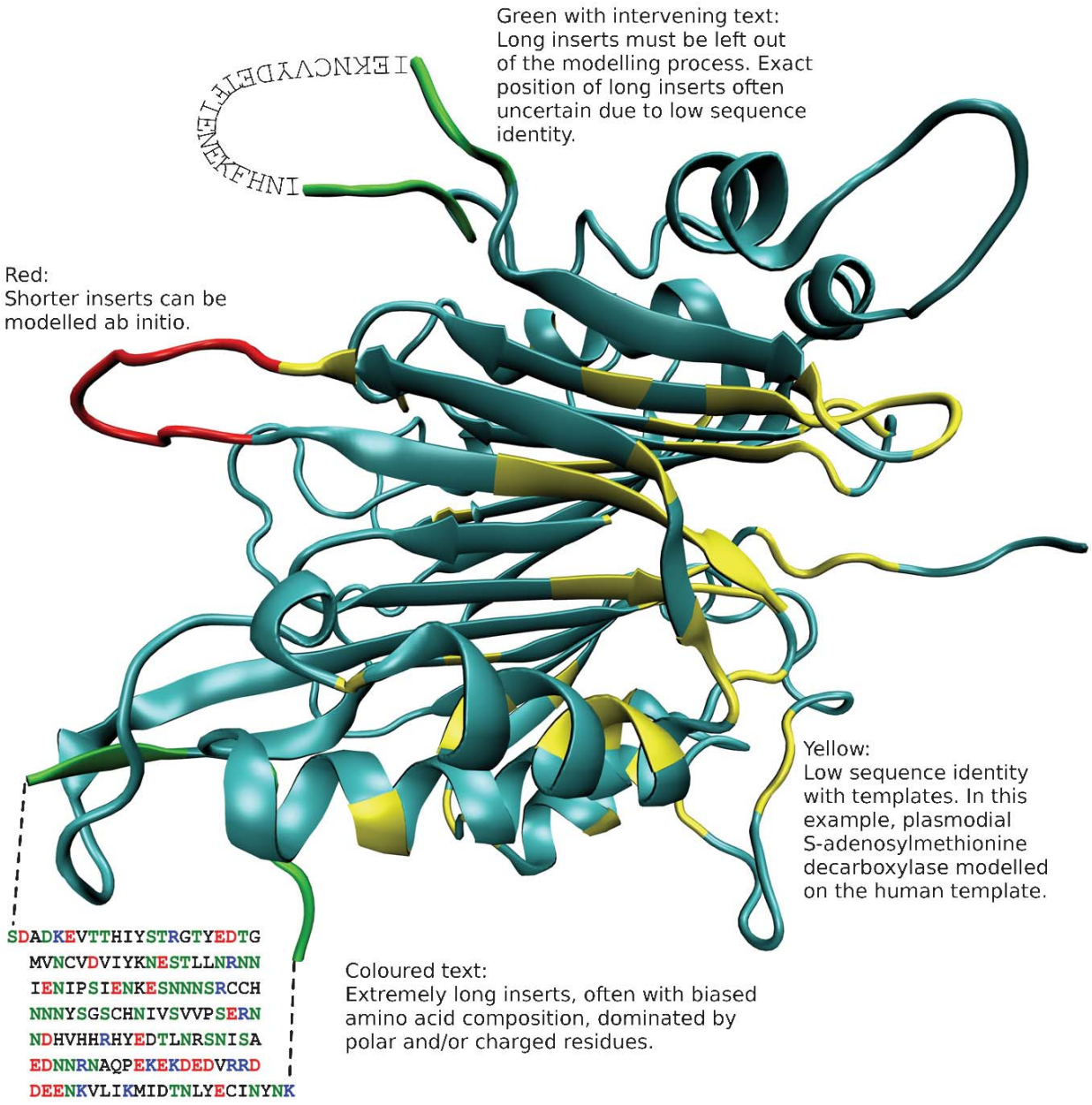

Fig. (3). Problems frequently encountered with modelling of Plasmodial proteins.

having long inserts that, along with low sequence similarity, make alignment problematic. Not surprisingly, proteins with long inserts appear to be avoided for modelling, and the problem of obtaining reliable alignments in their case is seldom discussed. The biased nucleotide and amino acid composition [53] and *Plasmodium*-specific inserts make it difficult to correctly identify core-conserved regions. The presence of inserts often confuses multiple and structural-alignment programs (Fig. (**3**)). A number of techniques can be used to circumvent this problem. From a first pass alignment approximate insert positions can be determined. Sequences can then be split according to long inserts and re-aligned. Inserts can vary considerably across different *Plasmodium* species [54, 55]. Therefore, while adjusting an alignment for modelling, it is useful to refer to phylogenetically diverse multiple alignments including as many *Plasmodium* protein sequences as possible [56]. Multiple alignments can be further improved by employing software that incorporates environment specific structural information, like FUGUE which combines profile and Hidden Markov Model (HMM) methods using both sequence and structural information. Such approaches have been successfully used for modelling of PfEMP-1 as well as for constructing reference alignments (Wells *et al.*, unpublished; [57]). Refined alignments might benefit from species-specific matrices that take into account the differences of amino acid distribution between the aligned proteins [58, 59]. As an adjunct to alignment, independent motif identification (e.g. the MEME system, http://meme.sdsc.edu/meme; [60]) can be used to fix mistakes that alignment programs frequently make when aligning long Plasmodial proteins with homologues [56, 61]. Further improvements

can be made by using hydrophobic cluster analysis [62] and secondary structure predictions to align homologous regions within inserts. Once an alignment has been decided on often based on necessary visual assessment, a series of models can be built. Because of the high degree of uncertainty that often accompanies alignments used for modelling Plasmodial proteins, it is usually not feasible to rectify all structural anomalies. However, by performing standard quality checks on a large sample of models and summarizing the results, it is possible to identify parts of the alignment causing most problems.

Despite the difficulties with homology modelling of Plasmodial proteins, there have been some notable successes and a diversity of applications. Plasmodial DHFR (dihydrofolate reductase) forms part of a bifunctional protein that also carries thymidylate synthase. A number of existing drugs such as cycloguanil and pyrimethamine targets the DHFR domain, and have been used effectively in the past. However, drug resistance has evolved that reduces the usefulness of this important class of drugs. Hence, Plasmodial DHFR has been a popular target for homology modelling efforts (e.g. [63-68]), which allowed the identification of new inhibitors in the nano- and micromolar range [63, 64], the rationalization of the antifolate resistance mechanisms [64-66, 68], as well as the ability for the drug WR99210, to inhibit both pyrimethamine and cycloguanil resistant mutants [66]. The impact of molecular modelling studies based on docking, pharmacophore mapping, QSAR, homology modelling, and quantum chemical studies in the design of Plasmodial DHFR and other antimalarial inhibitors has been recently reviewed [69]. The high accuracy of the alignment used for modelling and dockings of Plasmodial DHFR were subsequently confirmed with the crystal structure of the complete bifunctional enzyme [70]. Considerable work has also gone into modelling Plasmodial proteases essential to the parasite's intra-erythrocytic life stage. A number of these models has been used to identify new inhibitors [71-74], although the increasing number of crystal structures for these proteases is likely to gradually replace the need for homology models.

Guitérrez-de-Terán *et al.* [75] demonstrated the advantages of using multiple structures with plasmepsin IV from *P. falciparum*. A homology model and a low resolution crystal structure were both used for inhibitor identification. The homology model performed better on structural quality indicators and was more robust when calculating binding energy for an inhibitor series. The enhanced structural quality of the homology model was put down to the intermediate resolution of the X-ray structure (2.8 Å). Further improvements in predicting binding were gained by using a combined model employing both structures, as well as using molecular dynamics to increase sampling. The improved docking performance argues for making use of multiple experimental and predicted models instead of relying on a single structure (see also [76]).

Singh *et al.* [77] used homology modelling to derive a chimeric berghepain-2 that more closely resembled falcipain-2 in it's sensitivity to inhibitors. The motivation behind this approach was to create an *in vivo* rodent model of the *P. berghei* protein that mimics this important human drug target in *P. falciparum*. Homology modelling with molecular dynamics was used to predict the structure, substrate binding and MOA of histo-aspartic protease from *P. falciparum* [78]. Other noteworthy examples include homology models of dihydropteroate synthase (DHPS) from *P. vivax* and *P. falciparum* to explain the refractoriness of the *P. vivax* enzyme to sulfadoxine [79]. A homology model of histone deacetylase 1 from *P. falciparum* was successfully used to identify inhibitors in the nanomolar range with significant selectivity compared to mammalian cells [80]. Homology models combined with molecular dynamics were used to explain sulfadoxine resistance in mutants of *P. falciparum* DHPS [61].

A remarkable achievement is exemplified by the homology model obtained for *P. falciparum* farnesyltransferase (Ras FTase) based on a rat homologue [81]. The sequence identity between the target and template was quite low (23%) including a parasite-specific insert of approximately 100 residues in the Plasmodial protein. Using this model in the docking program GOLD, a range of ethylenediamine based inhibitors with IC_{50} < 50nM was identified of which two had an IC_{50} of less than 1 nM. This range of inhibitors was subsequently used together with the model for further rounds of optimization to derive new structures with better selectivity (up to 145 fold) towards the *P. falciparum* enzyme compared to its mammalian counterpart. Preliminary pharmacokinetics promisingly indicated that some of the compounds were metabolically stable [81-83]. The results of this work are encouraging and demonstrate that low sequence identity and the presence of inserts need not be a barrier to inhibitor discovery.

After a reliable structure for a Plasmodial drug target has been obtained, whether through modelling, X-ray crystallography or NMR, it has to be extensively analysed as part of the start of the lead discovery process (Fig. (**1**)). Protein quality assessment should be done to identify the limitations of the target structure to be used. The most reliable structures to be used are believed to be those from X-ray and NMR although one should be mindful of inaccuracies inherent in some crystal structures. Most deposited structures assume isotropic variation of atomic positions and do not fully capture the dynamic and anisotropic nature of protein crystals [84, 85]. It is essential that the dynamic nature of the target should be taken into account for successful SBDD and this may necessitate the use of multiple structures from crystallography or NMR. This can be further supplemented with various *in silico* methods such as molecular dynamics or Monte Carlo sampling. While homology models are a valuable tool to help fill the gap of undetermined structures, they are expected to be less accurate than X-ray or NMR structures due to errors in alignment, as well as errors "inherited" from experimental structures used as templates [76]. It is generally believed that homology models with a 50% sequence similarity can be accurately and independently used in SBDD [86, 87].

3.1.2. Receptor-Based Drug Design: Virtual High-Throughput Screening

Virtual screening is the process whereby a library of compounds is screened against a target using computational methods such as docking thereby quickly and efficiently

eliminating the majority of compounds that will not bind to a defined active site and serve as a filter of the chemical search space (Fig. (**2**)).

Docking is a computational tool that assesses the fit of a ligand into a protein cavity while evaluating protein-ligand interactions. These include approaches using genetic algorithms (GOLD, [88]), energetic evaluations and flexible protein/ligand docking (DOCK, [89]; AutoDock, [90]; FlexX, [91]; LigandFit, http://www.accelrys.com; Glide, http://www.schrodinger.com; Cdocker, http://www.accelrys.com; ICM, http://www.deltahpc.com). Comparisons between these programs are difficult due to the use of individual algorithms and various ways of evaluating ligand poses resulting in the scoring of ligand fits being one of the major problems in docking. Consensus scoring resulting from the use of different programs to perform the same docking may provide alternatives [92]. Whichever program is used, there is always a trade-off between docking accuracy and speed resulting in the use of a coarse docking in virtual HTS to eliminate obvious unsuitable compounds. This is usually followed by a more accurate screen with the remaining compounds to identify better binders. Alternatively, different programs may be used to re-dock a crystallized ligand into its receptor to allow the user to identify the program that produces the best fit and which could be used in further docking studies.

Patel *et al.* [93] used docking partly to validate dihydroorotate dehydrogenase in *P. falciparum* as a viable drug target and to provide structural information about the ability of a selection of compounds to bind to the active site. Docking of known inhibitors has also been used against wild-type and quadruple resistant mutant forms of *P. falciparum* DHFR [94] to define a common interaction pattern between inhibitors and the different forms of the protein that describes selection criteria for further screening strategies.

Virtual HTS of compounds to a target structure entails selection of either commercially or publicly available chemical libraries to be screened. The libraries of ligands to be docked are usually constructed using two main methods [95] firstly, by designing a library that includes a diversity of compounds which samples most of the conformational space and secondly, designing a library based on a rational structure approach where information from known interacting ligands is used to construct a diverse library based on certain functional and structural constraints.

Some of the major efforts to generate chemical databases include the Zinc Is Not Commercial (ZINC) database [96], National Cancer Institute (NCI, http://cactus.nci.nih.gov), PubChem (http://pubchem.ncbi.nlm.nih.gov), the Super Drug DataBase [97], the Drug Bank [98] and the SuperNatural database [99]. These databases are not all freely available for download and screening but are available on-line for similarity searches. Irwin and Shoichet [96] suggested that the "gold standard" for docking databases in academia are the commercially available, Available Chemical Database (ACD; http://www.mdli.com), ACD-SC (screening compound set (http://www.ccdc.cam.ac.uk)), Cambridge Structural Database (CSD, http://www.ccdc.cam.ac.uk) and the ChemNavigator database (http://www.chemnavigator.com). These chemical databases are but a few of the most popular ones used in virtual screening and contain compounds in the range of between a few hundred thousand up to 10 million. However, these databases have the drawback that they leave the user with the challenges of deciding on the protonation states, charges, tautomeric forms and removal of salts [96]. The ZINC database, containing over 8 million purchasable compounds, is the first database where all of these aspects have been addressed by the curators [96] and provides subsets such as lead-like, drug-like, fragment-like, Verneralis-filtered, etc., which are pre-filtered using specific criteria such as Lipinski's rule-of-five (http://zinc.docking.org).

One of the most remarkable examples of the success that can be attained with virtual HTS comes from the Cancer Project where a library of 3.5 billion molecules was screened against 12 anticancer protein targets [100]. A search of this magnitude would have taken about 100, 000 years on a desktop computer. The Cancer Project ended 27 April 2007 and a variety of hits were identified and are in the process of being synthesized and tested (http://www.chem.ox.ac.uk/curecancer.html). In one reported case in the Cancer Project, over 10% of the predicted molecules were experimentally active. This seems to imply that virtual HTS may provide rapid identification of compounds against malaria as well, potentially even in a 'piggy-backing' strategy.

Although virtual HTS is mainly achieved through clusters of computers physically connected to one another that can screen compound sets against the target, recent advances in network linking of computers are allowing powerful grid-computing strategies to be applied to HTS. Grid sites are typically distributed over a large geographical area linked via a high speed network [101] with a large number of grids recently established, attesting to the power of grid computing. These grids include Auvergrid (http://www.auvergrid.fr), E-science grid for Europe and Latin America (EELA, http://www.eu-eela.org), Enabling Grids for E-sciencE (EGEE) [102], EUChinaGrid (http://www.euchinagrid.org), EUMedGrid (http://www.eumedgrid.org), North Carolina BioGrid (http://www.ncbiogrid.org), the Canadian BioGrid (http://www.cbr.nrc.ca), the Asia Pacific BioGrid (http://www.apbionet.org/grid) and the Cancer Biomedical Informatics Grid [103]. These grids focus on different problems ranging from genetic linkage analysis [101] to molecular docking [104, 105] and metabolic pathway modelling [106].

Malaria presents various problems which can benefit from a grid-based approach and includes searching the *Plasmodium* genome and proteome for new drug targets, identification of single nucleotide polymorphisms (SNPs) on human as well as *Plasmodium* genomes relating to drug sensitivity, drug resistance mechanism elucidations as well as epidemiological monitoring of outbreaks [107]. Of these, drug discovery against malaria was identified as a key area. Various projects were initiated to use grids for large-scale docking of ligands in target proteins to assist in the discovery of new drugs against malaria. WISDOM-1 (World-wide *In Silico* Docking On Malaria) used EGEE to screen a filtered ZINC library against two *P. falciparum* plasmepsin proteins, (plasmepsin II and IV) with FlexX [108] and Autodock [90]. Around 1 million compounds were docked into each of the

targets and ultimately 41 million dockings were achieved in 6 weeks (the equivalent of 80 years of CPU power). WISDOM-I correctly identified known inhibitors as well as a new group of guanidino-based compounds, which are being investigated further [104]. Subsequently, in WISDOM-II, four different Plasmodial proteins (glutathione-S transferase, tubulin and DHFR from both *P. vivax* and *P. falciparum*) were targeted [104]. EGEE, Auvergrid, EELA, EUChinaGrid, EUMedGrid and FlexX was used to dock the same library used in WISDOM-I into the four selected proteins but re-docking against the co-crystallized compounds was performed to evaluate the docking parameters. During the 76 days duration of the project, nearly 140 million dockings were performed at a rate of almost 80 000 dockings per hour (equivalent to 413 years on a single PC). The outcome of these applications needs to be experimentally validated but illustrates the power of virtual HTS in substantially reducing search time as well as providing a coarse filtering of large libraries. Libraries can be further reduced using more accurate docking or screened using more stringent approaches. The use of grids as an initial screening tool will contribute significantly in the fight against malaria as more grids become available that can be applied to the search for new compounds.

3.1.3. Receptor-Based Drug Design: Pharmacophore Models

Receptor-based pharmacophore approaches use resolved structures to derive pharmacophore features and subsequently, pharmacophore models, which are a set of structural features in a molecule that is recognized at a receptor site and is responsible for the molecule's biological activity [109]. Preference should be given to structures resolved in complex with ligands due to conformational changes associated with ligand binding and the direct inference of protein-ligand interactions from the complexes. From these structures, a negative image of the active site can be constructed, which complements the interactions between the receptor and ligand described by pharmacophore models. These pharmacophore models are subsequently used to screen chemical libraries to find compounds matching the desired features. Hits identified during virtual HTS should then be filtered and ranked using docking techniques and the best scoring compounds then tested *in vitro*. Advantages of pharmacophore-based methods lie in the generation of divergent sets of compounds consisting of different scaffold structures and the derivation of the correct geometric orientation of the pharmacophore thereby providing directionality during the search for ligands and the identification of novel features [110]. The inclusion of dynamic descriptors in receptor-based pharmacophore strategies were developed to incorporate the inherent flexibility of protein structures in the drug design process and to reduce the entropic penalties that occur upon ligand binding to a target structure [111]. This led to a remarkable improvement in results compared to rigid pharmacophore models in a test case on HIV-1 proteases where 85-90% of known inhibitors were distinguished from drug-like non-inhibitors [112].

Due to the difficulty in obtaining 3D structures for Plasmodial proteins, very few receptor-based pharmacophore studies have been performed. Examples include the screening of a compound library of 2.6 million compounds against a receptor-based pharmacophore of *P. falciparum* spermidine synthase [113]. Seven potential inhibitors were identified from a subset of 28 compounds which were confirmed to bind to the protein using NMR techniques. Using a dynamic receptor-based pharmacophore approach resulted in the identification of two unique inhibitors (out of 9 tested *in vitro*) with micromolar Ki values against the protein. These inhibitors show great potential for lead optimization and further studies are currently underway (Burger *et al.* unpublished).

3.1.4. Receptor-Based Drug Design: Fragment-Based Design

Fragment-based drug design is recognised as a viable alternative to high-throughput screening [114, 115] and relies upon a library of smaller but more diverse ligands (molecular weights less than 200-300 Da) that are docked into the cavities of a protein. The high scoring hits are then used in subsequent steps of the rational drug design process. The motivation behind fragment-based screens is that the chemical space can be sampled much more efficiently with smaller, less complex molecules [115] of which the binding affinity per atom that binds to a protein can be just as good as an effective drug. Moreover, the resulting molecules are likely to have a better ligand-efficiency [116]. With structural insights, these fragments can be optimized much quicker to a lead compound stage [117] although linking the smaller ligands together to form an active compound can be a challenge [114]. The contribution of the specified group to the overall binding energy or potency can be calculated using a Group Efficiency score [115, 118]. This score is based upon changes in free energy of binding to a particular group of matched pairs of compounds during structure-based fragment optimization divided by the number of heavy atoms in the added group. Group efficiencies reveal the global efficiency or inefficiency of theoretically-derived additions and allows inference of which parts of the active site are responsible for contributing most to the affinity of the lead series by comparison of compounds with known binding modes. Currently, fragment optimization or fragment growing methods contributes to *de novo* drug design following classical medicinal chemistry strategies. As such, most of the work done on this aspect of receptor-based drug design is highly exploratory with a few examples encompassing integrated medicinal chemistry-like programs. The role of *in silico* fragment-based strategies is currently supportive and integrated with virtual screening, creation of virtual libraries, docking and experimental methodologies including NMR.

UCB Celltech proposed a novel strategy for fragment optimisation by combining modelling and medicinal chemistry resulting in very high ligand efficiencies of the resultant inhibitors [116]. Other structure-based approaches including the combination of virtual screening with NMR or crystallography, have been used by both academia and the private sector with Plexxikon, Astex and Abbott Laboratories claiming various levels of success [116, 119]. However, the exact magnitude of the contribution of these strategies to new therapeutics is unclear and evidence of applications of fragment-based drug design methods to the

malaria parasite is still lacking. There has been some criticism against the reductionist approach used in fragment-based screening but this is mainly towards the type of study that relies on smaller molecules that can be detected and used in *in vitro* studies [120].

3.2. Ligand-Based Drug Design

In the absence of a 3D target structure but with access to a set of structurally divergent compounds, *in silico* ligand-based approaches may be applied (Fig. (**2**)). All methodologies in this approach aim to reduce the chemical search space and may include similarity searching, sub-structure searching, SAR/QSAR and ligand-based pharmacophores. These methods are usually tightly integrated since this approach is based on the assumption that molecules with similar physiochemical properties would exert a similar biological activity [121].

3.2.1. Ligand-Based Drug Design: Similarity Searching Substructure Searching and QSAR

Similarity and substructure searching are some of the most diverse and useful tools in the drug design tool kit. Similarity searching, substructure searching and QSAR have been widely used to explore the chemical space of known inhibitors in the absence of a 3D target structure. Although this approach is hampered by very large chemical search spaces, it provides valuable information with regards to identifying common scaffolds and predicting activity.

Similarity searching makes use of molecular fingerprints in which are encoded fragment-type descriptors depicting the presence or absence of particular chemical features. Substructure searches can be defined as a search performed on complete structures to identify a specific query sub-structure. Maximum common substructure approaches are often preferred since they are more flexible compared to traditional similarity searching, which only considers global similarities between structures [122]. Similarity searching can be used to complement substructure searching since it often returns alternative structures. The predominant use of these methods is currently in the design of specific libraries to be used in virtual HTS [123]. However, these methods can also be used to filter databases and design custom libraries to be screened *in silico*. The use of similarity and substructure searching has become readily accessible by projects such as PubChem and the DrugBank and it is foreseen that it will play an increasingly important role in the drug discovery pipeline for malaria.

If a set of structurally divergent compounds with known inhibition activities is available, a QSAR can be determined and used to predict the inhibition potential of new compounds statistically. QSAR can include various levels of information that are captured in 2D-QSAR, 3D-QSAR or 4D-QSAR models. Several QSAR studies have been performed on malaria with various levels of success [124-129].

A linear discriminant-based QSAR model approach was used to screen a set of compounds for inhibitors of *P. falciparum* Ras FTase and resulted in two new compounds being identified. These compounds were tested *in vivo* and both compounds showed inhibition. One compound, aryl-aminomethylenemalonate, was the first of its kind to show antimalarial activity. In another study by Mahmoudi *et al*. [128] 127 compounds previously identified to act against the liver stages of *P. yoelii*, were used to derive a QSAR model which was subsequently used to screen databases for new compounds active against the liver stages. Various new compounds including known antiretroviral and antifungal agents as well as two ionophores that inhibit parasite development were identified.

3.2.2. Ligand-Based Drug Design: Ligand-Based Pharmacophore Models and QSAR

If a divergent set of compounds active against a specific target is available, pharmacophore features can be extracted and used in the generation of ligand-based pharmacophore models, which in turn can be screened against chemical databases to identify new lead compounds [130]. These pharmacophore models can additionally be used to identify inhibitors with a wide diversity of backbones (scaffold-hopping) and identify new ligands with different chemotypes but which still have a similar biological activity [131]. As with the receptor-based pharmacophore approach, the advantage lies in the ability to generate a diverse set of compounds [110].

The use of ligand-based pharmacophore approaches has evolved as an important technique in the fight against malaria. Parenti and co-workers (2004) successfully used a 3D-QSAR pharmacophore model to quantitatively predict inhibition constants of compounds for *P. falciparum* DHFR [132]. They used multiple methods to validate the model, such as Fisher's randomization, and upon testing the 3D pharmacophore was able to correctly identify active DHFR inhibitors from the MDL Drug Data Report (MDDR) database (www.mdli.com). Using this approach, and various statistical measures, they showed that the 3D pharmacophore with quantitative predictability could be utilized in virtual screening of databases and libraries. In another study Schormann *et al*. (2008) used various crystal structures of the *T. cruzi* DHFR-TS complex to derive 3D pharmacophores and a 3D QSAR model [133]. The model was tested against a selected compound set and the quality of the predictions showed that it could be used in further studies. They went on to propose refinements that can be applied to the 3D QSAR model.

The relationships derived from statistical analyses can also be embedded in pharmacophore models and used in virtual screening of chemical databases which may lead to the discovery of new inhibitors and/or novel scaffold structures. In a study by Bhattacharjee *et al*. [134, 135], it was shown that from a set of divergent compounds (tryptanthrin derivates) a 3D-QSAR pharmacophore model could be derived and were found to be both statistically and mechanistically significant in the identification of new-antimalarial compounds. Five aminoquinazoline derivatives were identified showing potent *in vivo* activity in mouse malarial screening tests. The specific target with which tryptanthrin derivatives interacts is unknown as for many other antimalarial drugs, therefore 3D-QSAR pharmacophore models may hold a key to understanding existing antimalarials and finding new antimalarials.

A pharmacophore model derived from 17 antimalarial compounds was used to search for new compounds using multiconformer libraries [135]. Various new compounds were identified and shown to be active *in vitro* against various strains of resistant *P. falciparum*. Another ligand-based pharmacophore derived from chalcones, was used to identify important features such as an aromatic ring, which plays an important role in chalcone activity [134]. Dascombe *et al.* [136] used a metaquinine pharmacophore model to identify new compounds based on the proven solubility and oral bioavailability of metaquinine. This was followed by various *in vitro* methods that confirmed the results and demonstrated the use of pharmacophores to help identify possible new 4-aminoquinine antimalarials. Drew *et al.* [137] used *in silico* quantum mechanical methods to evaluate all the known crystal structures of artemisinin from which a pharmacophore model for artemisinin was constructed consisting of various deoxyartemisinin and deoxyarteether derivatives. Subsequent experiments showed that substitution of the O3 atom has the most effect on the generation of the carbon-centred radical, which is responsible for the antiparasitic action of artemisinins. This study also identified the most important bonds in the artemisinin and provided the groundwork for the synthesis of new derivatives.

4. UNIQUE POSSIBILITIES IN EXPLOITING PLASMODIAL PROTEIN CHARACTERISTICS

Most regulatory proteins are often of a multiprotein or multidomain structure although most currently used pharmaceuticals target monomeric active sites. The complex networks of protein-protein interactions that govern coordinated cellular responses provide attractive targets for therapeutic interventions. In the malaria parasite, several Plasmodial proteins that have been described as drug targets and are clinically targeted, have the additional distinction of also being multidomain or bifunctional proteins including DHFR-TS, dihydro-6-hydroxymethylpterin pyrophosphokinase-dihydropteroate synthase, S-Adenosylmetionine decarboxylase-ornithine decarboxylase and others. It is therefore worthwhile exploiting these unique biological characteristics in the development of antimalarials. *In silico* methodologies that have been applied in other organisms become very attractive and should be tested against multidomain proteins in Plasmodia. Programs such as DOMINANT [138] allows deconvolution of protein structures such that domains and domain boundaries can be identified. This obviously depends on the availability of multidomain Plasmodial structures, of which there are few. Other computational methods for the analysis of protein-protein interactions have also been reviewed [139]. Our group has followed the approach of first generating multiple models (from multiple *Plasmodium* species) of each domain of such proteins followed by "each-against-each" docking to predict the potential organisation of bifunctional protein complexes (Wells G.A., unpublished). Again, in lieu of the unique nature of the protein-protein interaction network of *P. falciparum*, this might identify novel characteristics suitable for targeting [41].

If the multidomain nature of a drug target is essential to the activities or regulation of the said protein, non-active site inhibitory strategies can be exploited. However, the large surface areas of protein interaction sites make inhibitory ligand binding predictions problematic. The identification of binding 'hot-spots', the so-called residues directly responsible for the interaction, is becoming computationally viable through applications like normalised interface propensities derived from rigid body docking [140]. Moreover, structures of proteins may be used to identify epitopes involved in protein-protein interactions through protein epitope mapping that leads to the design of scaffold structures that bind these areas but which have drug-like ADMET properties [141].

Several malaria proteins are also characterised by parasite-specific inserts that may be functional and quite often contain intrinsically disordered regions [54]. Many of the current small molecules that block protein-protein interactions lead to an order-to-disorder transition of one of the partner proteins and this association is proposed as a novel strategy to develop inhibitory small molecules [142, 143].

5. CONCLUDING DISCUSSION AND FUTURE PERSPECTIVES

The current portfolio of the Malaria Medicine Venture (http://www.mmv.org) lists less than 10 discovery projects and another eight antimalarial drugs, mostly reformulations of drug combinations at pre-clinical or clinical trial stages. About 12 drugs or drug combinations are currently available for prophylaxis and treatment of *P. falciparum* and *P. vivax* of which only a few are still clinically useful due to widespread drug resistance ([22] and references therein). Some of the current drugs in use are already older than 50 years and have in addition serious limitations such as costs and poor safety profiles [27, 31]. The cost of launching new drugs have risen sharply in the last three decades and consequently only 1% of the drugs launched between 1975 and 2004 were aimed at tropical diseases [144] representing about 0.2% of the budget spent on drug discovery by the pharmaceutical industry in 1999 [145]. The traditional drug discovery and development process is high-risk, costly and time consuming and the implementation of computational methods is gaining in popularity in attempts to restructure and streamline this process (see e.g. [146] and references therein). In one successful study the entire drug discovery process, which lasted less than two years before reaching clinical trial status, was based on *in silico* methods integrated with medicinal chemistry [147].

New developments such as the launching of the genome sequences of *P. falciparum* in 2002 and for *P. vivax* and *P. knowlesi* in 2008 [148-150] and increased funding by governments, public-private partnerships and philanthropic organizations have raised expectations for the prompt discovery of new antiparasitic drug entities. Although a wealth of new knowledge on the basic and distinctive biology of the malaria parasite became evident, this scientific knowledge has not yet been translated into modern therapeutics [39].

In silico antiparasite drug discovery methods are only at the initial stages of development and are clearly complementary to other *in silico* methods (bioinformatics and related disciplines), which need to be integrated with the more traditional biology and chemistry disciplines, collectively described as chemogenomics [151]. The formidable

challenges facing scientists wanting to explore the chemogenomics knowledge space of the malaria parasite for discovering new leads from the millions of chemical compounds stored in databases, have been reviewed [18]. The chemical universe is estimated to contain from 10^{12} to 10^{180} drug-like compounds [152] that are impossible to screen with standard biological assays. *In silico* approaches allow screening of up to 10^{12} molecules in compound libraries and thus early identification of potential leads. This review largely focused on the application of *in silico* methods for HTS and in solving of the three-dimensional structures of potential protein targets, which are either difficult to express or to obtain in sufficient quantities for more traditional X-ray crystallography or NMR studies [18]. Various structure-based drug design strategies were next reviewed with special emphasis on receptor-based and ligand-based drug design methods applied to malaria proteins. The application of these methods reduces the number of leads to be assayed and to be chemically modified, guided by medicinal chemistry principles. It is apparent that the quality of the derived homology models is in many cases sufficient to study and identify quality leads but it is still too early to judge the value of these methods for antiparasite drug discovery since in-depth experimental validations are currently mostly lacking. This review did not include a detailed discussion of the importance of ADMET in the drug discovery process, an evolving methodology which in a new lead optimization paradigm, draws on information technology resources such as artificial neural network programs, relational databases, and principles of systems biology as well as results of toxicoproteomics and toxicogenomics experiments (see e.g. [153, 154]) and is rightfully a stand-alone topic.

It is worth noting that *in silico*-based methodologies for drug discovery are still evolving [146] and only now emerging as an integral but promising field in antimalarial drug discovery research. The required expertise, which is knowledge-based and not simply screening, is still being developed and informed decisions are dependent on the quality of the starting materials including the quality of 3D structures of proteins whether derived by X-ray crystallography, NMR or homology modelling and the quality of chemical libraries. In this context it is worth taking note of the criteria that was used in assembling three types of chemical libraries for drug discovery for neglected diseases: one diverse *in silico* library for virtual screening, one diverse screening compound library, and a focused compound library for the discovery of kinase inhibitors [24]. Natural products have been used for millennia as medical remedies and more recently as starting compounds for modern medicines. However, interest in the development of natural products by the pharmaceutical industry has declined. Given the historic role of plant remedies in the treatment of malaria [155] it is noteworthy that there are renewed attempts to bring natural products back into use within the lead generation paradigm. These strategies include the development of drug-like natural product libraries and for example, merging of ethnopharmacology with virtual screening for lead structure discovery [156-159].

In silico methods are not stand-alone methodologies and need to be closely integrated with bench-type experimental studies. Clear evidence of the benefits of such an integrated approach is revealed by a report in which the results obtained in a malaria high-throughput screening of a library of 1.7 million compounds were analysed by *in silico* methods [160]. These authors identified a subset of about 17 000 compounds with potent antimalarial activities, which after using several clustering approaches and docking experiments revealed significant information on the MOA and/or protein targets of selected compounds as well as novel chemical scaffolds as leads for further studies. The expediency of using 3D protein structures derived from both homology modelling and crystallography in antimalarial drug discovery is exemplified by the successful identification of new compounds with promising antimalarial activity against both wild-type and mutant *P. falciparum* [69].

It is apparent from the *in silico* methods described in this review and the examples provided that the embedded information in malaria parasite genome sequences has an increasing impact on the delivery of new knowledge and more importantly, on novel applications of this knowledge in the control of the malaria disease. There is thus every reason to be optimistic about significant breakthroughs in the next few years.

ACKNOWLEDGEMENTS

LMB and AIL are members of the South African Malaria Initiative (http://www.acgt.co.za/SAMI) and acknowledge the Department of Science and Technology of South Africa for funding the SAMI programme. TAPdB and PBB hold fellowships from the South African National Bioinformatics Network and GAW is the recipient of a South African National Research Foundation (NRF) Prestigious bursary. Research is funded by South African NRF grants (2047008, 2053368, FA2006040400011 and FA2007050300003). Authors are supported by the French-South African SAFETI program. EM is supported by a grant from Agence Nationale de la Recherce (ANR-06-MDCA-014). Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author(s) and therefore, the NRF does not accept any liability in regard thereto.

ABBREVIATIONS

SAR	=	Structure Activity Relationship
QSAR	=	Quantitative SAR
HTS	=	High Throughput Screening
SBDD	=	Structure-Based Drug Design

REFERENCES

[1] Snow, R. W.; Guerra, C. A.; Noor, A. M.; Myint, H. Y.; Hay, S. I. *Nature*, **2005**, 434, 214.
[2] Guerra, C. A.; Snow, R. W.; Hay, S. I. *Trends Parasitol.*, **2006**, 22, 353.
[3] Sachs, J.; Malaney, P. *Nature*, **2002**, 415, 680.
[4] Sachs, J. D. *Science*, **2002**, 298, 122.
[5] Pink, R.; Hudson, A.; Mouriès, M.-A.; Bendig, M. *Nat. Rev. Drug Discov.*, **2005**, 4, 727.
[6] Hopkins, A. L.; Witty, M. J.; Nwaka, S. *Nature*, **2007**, 449, 166.
[7] Nwaka, S. *Trans. R. Soc. Trop. Med. Hyg.*, **2005**, 99 Suppl 1, S20-S29.
[8] Bhogal, N.; Balls, M. *Curr. Drug Discov. Technol.*, **2008**, 5, 250.

[9] Hellerstein, M. K. *Metab. Eng.*, **2008**, 10, 1.
[10] Lipkus, A. H.; Yuan, Q.; Lucas, K. A.; Funk, S. A.; Bartelt, W. F., III; Schenck, R. J.; Trippe, A. J. *J. Org. Chem.*, **2008**, 73, 4443.
[11] Hopkins, A. L. *Nat. Chem. Biol.*, **2008**, 4, 682.
[12] Jenwitheesuk, E.; Horst, J. A.; Rivas, K. L.; Voorhis, W. C. V.; Samudrala, R. *Trends Pharmacol. Sci.*, **2008**, 29, 62.
[13] Muskavitch, M. A.; Barteneva, N.; Gubbels, M. J. *Comb. Chem. High Throughput Screen.*, **2008**, 11, 624.
[14] Neamati, N.; Barchi, J. J., Jr. *Curr. Top. Med. Chem.*, **2002**, 2, 211.
[15] Kassel, D. B. *Curr. Opin. Chem. Biol.*, **2004**, 8, 339.
[16] Hall, S. E. *Drug Discov. Today*, **2006**, 11, 495.
[17] Lang, P.; Yeow, K.; Nichols, A.; Scheer, A. *Nat. Rev. Drug Discov.*, **2006**, 5, 343.
[18] Birkholtz, L.; van Brummelen, A. C.; Clark, K.; Niemand, J.; Maréchal, E.; Llinás, M.; Louw, A. I. *Acta Trop.*, **2008**, 105, 113.
[19] Paolini, G. V.; Shapland, R. H. B.; van Hoorn, W. P.; Mason, J. S.; Hopkins, A. L. *Nat. Biotechnol.*, **2006**, 24, 805.
[20] Sawyer, T. K. *Nat. Chem. Biol.*, **2006**, 2, 646.
[21] Saidani, N.; Grando, D.; Valadie, H.; Bastien, O.; Marechal, E. *Infect. Genet. Evol.*, **2008**.
[22] Caffrey, C. R.; Sterverding, D. *Expert Opin. Drug Disc.*, **2008**, 3, 173.
[23] Biot, C.; Chibale, K. *Infect. Disord. Drug Targets*, **2006**, 6, 173.
[24] Brenk, R.; Schipani, A.; James, D.; Krasowski, A.; Gilbert, I. H.; Frearson, J.; Wyatt, P. G. *ChemMedChem.*, **2008**, 3, 435.
[25] Frearson, J. A.; Wyatt, P. G.; Gilbert, I. H.; Fairlamb, A. H. *Trends Parasitol.*, **2007**, 23, 589.
[26] Fidock, D. A.; Rosenthal, P. J.; Croft, S. L.; Brun, R.; Nwaka, S. *Nat. Rev. Drug Discov.*, **2004**, 3, 509.
[27] Nwaka, S.; Hudson, A. *Nat. Rev. Drug Discov.*, **2006**, 5, 941.
[28] Wolf, K.; Dormeyer, M. *Parasitol. Res.*, **2003**, 90 Suppl 2, S91-S96.
[29] Linares, G. E. G.; Ravaschino, E. L.; Rodriguez, J. B. *Curr. Med. Chem.*, **2006**, 12, 335.
[30] Rosenthal, P. J. *J. Exp. Biol.*, **2003**, 206, 3735.
[31] Renslo, A. R.; McKerrow, J. H. *Nat. Chem. Biol.*, **2006**, 2, 701.
[32] Joachimiak, M. P.; Chang, C.; Rosenthal, P. J.; Cohen, F. E. *Mol. Med.*, **2001**, 7, 698.
[33] Foth, B. J.; Ralph, S. A.; Tonkin, C. J.; Struck, N. S.; Fraunholz, M.; Roos, D. S.; Cowman, A. F.; McFadden, G. I. *Science*, **2003**, 299, 705.
[34] Lell, B.; Ruangweerayut, R.; Wiesner, J.; Missinou, M. A.; Schindler, A.; Baranek, T.; Hintz, M.; Hutchinson, D.; Jomaa, H.; Kremsner, P. G. *Antimicrob. Agents Chemother.*, **2003**, 47, 735.
[35] Wiesner, J.; Seeber, F. *Expert Opin. Ther. Targets*, **2005**, 9, 23.
[36] Zhou, Y.; Ramachandran, V.; Kumar, K. A.; Westenberger, S.; Refour, P.; Zhou, B.; Li, F.; Young, J. A.; Chen, K.; Plouffe, D.; Henson, K.; Nussenzweig, V.; Carlton, J.; Vinetz, J. M.; Duraisingh, M. T.; Winzeler, E. A. *PLoS ONE*, **2008**, 3, e1570.
[37] Hopkins, A. L.; Groom, C. R. *Nat. Rev. Drug Discov.*, **2002**, 1, 727.
[38] Aguero, F.; Al-Lazikani, B.; Aslett, M.; Berriman, M.; Buckner, F. S.; Campbell, R. K.; Carmona, S.; Carruthers, I. M.; Chan, A. W.; Chen, F.; Crowther, G. J.; Doyle, M. A.; Hertz-Fowler, C.; Hopkins, A. L.; McAllister, G.; Nwaka, S.; Overington, J. P.; Pain, A.; Paolini, G. V.; Pieper, U.; Ralph, S. A.; Riechers, A.; Roos, D. S.; Sali, A.; Shanmugam, D.; Suzuki, T.; Van Voorhis, W. C.; Verlinde, C. L. *Nat. Rev. Drug Discov.*, **2008**, 7, 900.
[39] Winzeler, E. A. *Nature*, **2008**, 455, 751.
[40] Young, D. *Am. J. Health Syst. Pharm.*, **2005**, 62, 1226.
[41] Birkholtz, L. M.; Blatch, G.; Coetzer, T. L.; Hoppe, H. C.; Human, E.; Morris, E. J.; Ngcete, Z.; Oldfield, L.; Roth, R.; Shonhai, A.; Stephens, L.; Louw, A. *Malar. J.*, **2008**, 7, 197.
[42] Yeh, I.; Hanekamp, T.; Tsoka, S.; Karp, P. D.; Altmann, R. B. *Genome Res.*, **2004**, 14, 917.
[43] Ginsburg, H. *Trends Parasitol.*, **2006**, 22, 238.
[44] Fatumo, S.; Plaimas, K.; Mallm, J. P.; Schramm, G.; Adebiyi, E.; Oswald, M.; Eils, R.; König, R. *Infect. Genet. Evol.*, **2009**, In press.
[45] Zhou, Y.; Ramachandran, V.; Kumar, K. A.; Westenberger, S.; Refour, P.; Zhou, B.; Li, F.; Young, J. A.; Chen, K.; Plouffe, D.; Henson, K.; Nussenzweig, V.; Carlton, J.; Vinetz, J. M.; Duraisingh, M. T.; Winzeler, E. A. *PLoS. ONE.*, **2008**, 3, e1570.
[46] Brehelin, L.; Dufayard, J.-F.; Gascuel, O. *BMC Bioinformatics*, **2008**, 9, 440.
[47] Joubert, Y.; Joubert, F. *Malar. J.*, **2008**, 7, 90.
[48] Sakharkar, K. R.; Sakharkar, M. K.; Chow, V. T. *Methods Mol. Med.*, **2008**, 142, 1.
[49] Birkholtz, L. M.; Bastien, O.; Wells, G.; Grando, D.; Joubert, F.; Kasam, V.; Zimmermann, M.; Ortet, P.; Jacq, N.; Saïdani, N.; Roy, S.; Hofmann-Apitius, M.; Breton, V.; Louw, A. I.; Maréchal, E. *Malar. J.*, **2006**, 5, 110.
[50] Villoutreix, B. O.; Renault, N.; Lagorce, D.; Sperandio, O.; Montes, M.; Miteva, M. A. *Curr. Protein Pept. Sci.*, **2007**, 8, 381.
[51] Anderson, A. C. *Chem. Biol.*, **2003**, 10, 787.
[52] Vedadi, M.; Lew, J.; Artz, J.; Amani, M.; Zhao, Y.; Dong, A.; Wasney, G. A.; Gao, M.; Hills, T.; Brokx, S.; Qiu, W.; Sharma, S.; Diassiti, A.; Alam, Z.; Melone, M.; Mulichak, A.; Wernimont, A.; Bray, J.; Loppnau, P.; Plotnikova, O.; Newberry, K.; Sundararajan, E.; Houston, S.; Walker, J.; Tempel, W.; Bochkarev, A.; Kozieradzki, I.; Edwards, A.; Arrowsmith, C.; Roos, D.; Kain, K.; Hui, R. *Mol. Biochem. Parasitol.*, **2007**, 151, 100.
[53] Bastien, O.; Lespinats, S.; Roy, S.; Métayer, K.; Fertil, B.; Codani, J.-J.; Maréchal, E. *Gene*, **2004**, 336, 163.
[54] Birkholtz, L. M.; Wrenger, C.; Joubert, F.; Wells, G. A.; Walter, R. D.; Louw, A. I. *Biochem. J.*, **2004**, 377, 439.
[55] Rastelli, G.; Pacchioni, S.; Parenti, M. D. *Bioorg. Med. Chem. Lett.*, **2003**, 13, 3257.
[56] Wells, G. A.; Birkholtz, L. M.; Joubert, F.; Walter, R. D.; Louw, A. I. *J. Mol. Graph. Model.*, **2006**, 24, 307.
[57] Bertonati, C.; Tramontano, A. *Proteins*, **2007**, 69, 215.
[58] Bastien, O.; Aude, J. C.; Roy, S.; Maréchal, E. *Bioinformatics*, **2004**, 20, 534.
[59] Yu, Y.-K.; Altschul, S. *Bioinformatics*, **2005**, 21, 902.
[60] Bailey, T. L.; Elkan, C. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **1994**, 2, 28.
[61] de Beer, T. A. P.; Louw, A. I.; Joubert, F. *Bioorg. Med. Chem.*, **2006**, 14, 4433.
[62] Callebaut, I.; Prat, K.; Meurice, E.; Mornon, J.-P.; Tomavo, S. *BMC Genomics*, **2005**, 6, 100.
[63] Toyoda, T.; Brobey, R. K.; Sano, G.; Horii, T.; Tomioka, N.; Itai, A. *Biochem. Biophys. Res. Commun.*, **1997**, 235, 515.
[64] McKie, J. H.; Douglas, K. T.; Chan, C.; Roser, S. A.; Yates, R.; Read, M.; Hyde, J. E.; Dascombe, M. J.; Yuthavong, Y.; Sirawaraporn, W. *J. Med. Chem.*, **1998**, 41, 1367.
[65] Lemcke, T.; Christensen, I. T.; Jorgensen, F. S. *Bioorg. Med. Chem.*, **1999**, 7, 1003.
[66] Rastelli, G.; Sirawaraporn, W.; Sompornpisut, P.; Vilaivan, T.; Kamchonwongpaisan, S.; Quarrell, R.; Lowe, G.; Thebtaranonth, Y.; Yuthavong, Y. *Bioorg. Med. Chem.*, **2000**, 8, 1117.
[67] Santos-Filho, O. A.; de Alencastro, R. B. *Biophys. Chem.*, **2001**, 91, 305.
[68] Delfino, R. T.; Santos-Filho, O. A.; Figueroa-Villar, J. D. *Biophys. Chem.*, **2002**, 98, 287.
[69] Adane, L.; Bharatam, P. V. *Curr. Med. Chem.*, **2008**, 15, 1552.
[70] Yuvaniyama, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Vanichtanankul, J.; Sirawaraporn, W.; Taylor, P.; Walkinshaw, M. D.; Yuthavong, Y. *Nat. Struct. Biol.*, **2003**, 10, 357.
[71] Li, R.; Chen, X.; Gong, B.; Selzer, P. M.; Li, Z.; Davidson, E.; Kurzban, G.; Miller, R. E.; Nuzum, E. O.; McKerrow, J. H.; Fletterick, R. J.; Gillmor, S. A.; Craik, C. S.; Kuntz, I. D.; Cohen, F. E.; Kenyon, G. L. *Bioorg. Med. Chem.*, **1996**, 4, 1421.
[72] Desai, P. V.; Patny, A.; Sabnis, Y.; Tekwani, B.; Gut, J.; Rosenthal, P.; Srivastava, A.; Avery, M. *J. Med. Chem.*, **2004**, 47, 6609.
[73] Desai, P. V.; Patny, A.; Gut, J.; Rosenthal, P. J.; Tekwani, B.; Srivastava, A.; Avery, M. *J. Med. Chem.*, **2006**, 49, 1576.
[74] de Terán, H. G.; Nervall, M.; Ersmark, K.; Liu, P.; Janka, L. K.; Dunn, B.; Hallberg, A.; Aqvist, J. *Biochemistry*, **2006**, 45, 10529.
[75] de Terán, H. G.; Nervall, M.; Dunn, B. M.; Clemente, J. C.; Aqvist, J. *FEBS Lett.*, **2006**, 580, 5910.
[76] Luksch, T.; Chan, N. S.; Brass, S.; Sotriffer, C. A.; Klebe, G.; Diederich, W. E. *ChemMedChem.*, **2008**, 3, 1323.
[77] Singh, A.; Walker, K. J.; Sijwali, P. S.; Lau, A. L.; Rosenthal, P. J. *Protein Eng. Des. Sel.*, **2007**, 20, 171.
[78] Bjelic, S.; Aqvist, J. *Biochemistry*, **2004**, 43, 14521.
[79] Korsinczky, M.; Fischer, K.; Chen, N.; Baker, J.; Rieckmann, K.; Cheng, Q. *Antimicrob. Agents Chemother.*, **2004**, 48, 2214.
[80] Andrews, K. T.; Tran, T. N.; Lucke, A. J.; Kahnberg, P.; Le, G. T.; Boyle, G. M.; Gardiner, D. L.; Skinner-Adams, T. S.; Fairlie, D. P. *Antimicrob. Agents Chemother.*, **2008**, 52, 1454.
[81] Glenn, M. P.; Chang, S. Y.; Hucke, O.; Verlinde, C. L.; Rivas, K.; Horney, C.; Yokoyama, K.; Buckner, F. S.; Pendyala, P. R.;

Chakrabarti, D.; Gelb, M.; Van Voorhis, W. C.; Sebti, S. M.; Hamilton, A. D. *Angew. Chem. Int. Ed. Engl.*, **2005**, 44, 4903.
[82] Glenn, M. P.; Chang, S. Y.; Horney, C.; Rivas, K.; Yokoyama, K.; Pusateri, E. E.; Fletcher, S.; Cummings, C. G.; Buckner, F. S.; Pendyala, P. R.; Chakrabarti, D.; Sebti, S. M.; Gelb, M.; Van Voorhis, W. C.; Hamilton, A. D. *J. Med. Chem.*, **2006**, 49, 5710.
[83] Fletcher, S.; Cummings, C. G.; Rivas, K.; Katt, W. P.; Hornéy, C.; Buckner, F. S.; Chakrabarti, D.; Sebti, S. M.; Gelb, M. H.; Voorhis, W. C. V.; Hamilton, A. D. *J. Med. Chem.*, **2008**, 51, 5176.
[84] DePristo, M. A.; de Bakker, P. I.; Blundell, T. L. *Structure*, **2004**, 12, 831.
[85] Davis, A. M.; St-Gallay, S. A.; Kleywegt, G. J. *Drug Discov. Today*, **2008**, 13, 831.
[86] Hillisch, A.; Pineda, L. F.; Hilgenfeld, R. *Drug Discov. Today*, **2004**, 9, 659.
[87] Ferrara, P.; Jacoby, E. *J. Mol. Model.*, **2007**, 13, 897.
[88] Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.*, **1997**, 267, 727.
[89] Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.*, **1982**, 161, 269.
[90] Goodsell, D. S.; Olson, A. J. *Proteins*, **1990**, 8, 195.
[91] Rarey, M.; Wefing, S.; Lengauer, T. *J. Comput. Aided Mol. Des.*, **1996**, 10, 41.
[92] Teramoto, R.; Fukunishi, H. *J. Chem. Inf. Model.*, **2008**, 48, 747.
[93] Patel, V.; Booker, M.; Kramer, M.; Ross, L.; Kennedy, L. M.; Dvorin, J. D.; Duraisingh, M. T.; Sliz, P.; Wirth, D. F.; Clardy, J. *J. Biol. Chem.*, **2008**, *283*, 35078.
[94] Fogel, G. B.; Cheung, M.; Pittman, E.; Hecht, D. *J. Mol. Graph. Model.*, **2008**, 26, 1145.
[95] Terstappen, G. C.; Reggiani, A. *Trends Pharmacol. Sci.*, **2001**, 22, 23.
[96] Irwin, J. J.; Shoichet, B. K. *J. Chem. Inf. Model.*, **2005**, 45, 177.
[97] Goede, A.; Dunkel, M.; Mester, N.; Frommel, C.; Preissner, R. *Bioinformatics*, **2005**, 21, 1751.
[98] Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. *Nucleic Acids Res.*, **2008**, 36, D901-D906.
[99] Dunkel, M.; Fullbeck, M.; Neumann, S.; Preissner, R. *Nucleic Acids Res.*, **2006**, 34, D678-D683.
[100] Richards, W. G. *Nat. Rev. Drug Discov.*, **2002**, 1, 551.
[101] Andrade, J.; Andersen, M.; Sillén, A.; Graff, C.; Odeberg, J. *Eur. J. Hum. Genet.*, **2007**, 15, 694.
[102] Gagliardi, F.; Jones, B.; Grey, F.; Bégin, M.-E.; Heikkurinen, M. *Philos. Transact. A Math. Phys. Eng. Sci.*, **2005**, 363, 1729.
[103] Covitz, P. A.; Hartel, F.; Schaefer, C.; Coronado, S. D.; Fragoso, G.; Sahni, H.; Gustafson, S.; Buetow, K. *Bioinformatics*, **2003**, 19, 2404.
[104] Kasam, V.; Salzemann, J.; Breton, V.; Jacq, N. WISDOM-II: a large in silico docking effort for finding novel hits against malaria using computational grid infrastructure. In *Proceedings of the 5th IEEE workshop on Challenges of large applications in distributed environments*; **2007**.
[105] Kasam, V.; Salzemann, J.; Jacq, N.; Mass, A.; Breton, V. Large Scale Deployment of Molecular Docking Application on Computational Grid Infrastructures for Combating Malaria. In *CC GRID 2007*; **2007**.
[106] Kimura, S.; Kawasaki, T.; Hatakeyama, M.; Naka, T.; Konishi, F.; Konagaya, A. *Bioinformatics*, **2004**, 20, 1646.
[107] Breton, V.; Jacq, N.; Kasam, V.; Hofmann-Apitius, M. *IEEE Trans. Inf. Technol. Biomed.*, **2008**, 12, 173.
[108] Rareya, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.*, **1996**, 261, 470.
[109] Gund, P. *Prog. Mol. Subcell. Biol.*, **1977**, 5, 117.
[110] Dror, O.; Shulman-Peleg, A.; Nussinov, R.; Wolfson, H. *Curr. Med. Chem.*, **2004**, 11, 71.
[111] Carlson, H. A.; Masukawa, K. M.; Rubins, K.; Bushman, F. D.; Jorgensen, W. L.; Lins, R. D.; Briggs, J. M.; McCammon, J. A. *J. Med. Chem.*, **2000**, 43, 2100.
[112] Meagher, K. L.; Carlson, H. A. *J. Am. Chem. Soc.*, **2004**, 126, 13276.
[113] Jacobsson, M.; Gäredal, M.; Schultz, J.; Karlén, A. *J. Med. Chem.*, **2008**, 51, 2777.
[114] Villar, H. O.; Hansen, M. R. *Curr. Top. Med. Chem.*, **2007**, 7, 1509.
[115] Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. *J. Med. Chem.*, **2008**, 51, 3661.
[116] Erlanson, D. A. *Curr. Opin. Biotechnol.*, **2006**, 17, 643.
[117] Hesterkamp, T.; Whittaker, M. *Curr. Opin. Chem. Biol.*, **2008**, 12, 260.
[118] 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.; Lacount, D.; Vignali, M.; Li, J.; Mondal, S.; Massey, A.; Carroll, B.; Gulde, S.; Luft, J.; Desoto, L.; Holl, M.; Caruthers, J.; Bosch, J.; Robien, M.; Arakaki, T.; Holmes, M.; Le, T., I; Hol, W. G. *Methods Mol. Biol.*, **2008**, 426, 497.
[119] Wyatt, P. G.; Woodhead, A. J.; Berdini, V.; Boulstridge, J. A.; Carr, M. G.; Cross, D. M.; Davis, D. J.; Devine, L. A.; Early, T. R.; Feltell, R. E.; Lewis, E. J.; McMenamin, R. L.; Navarro, E. F.; O'Brien, M. A.; O'Reilly, M.; Reule, M.; Saxty, G.; Seavers, L. C. A.; Smith, D. M.; Squires, M. S.; Trewartha, G.; Walker, M. T.; Woolford, A. J.-A. *J. Med. Chem.*, **2008**, 51, 4986.
[120] Peterson, R. T. *Nat. Chem. Biol.*, **2008**, 4, 635.
[121] Sheridan, R. P.; Kearsley, S. K. *Drug Discov. Today*, **2002**, 7, 903.
[122] Cao, Y.; Jiang, T.; Girke, T. *Bioinformatics*, **2008**, 24, i366-i374.
[123] Gillet, V. J. *Curr. Opin. Chem. Biol.*, **2008**, 12, 372.
[124] Marrero-Ponce, Y.; Iyarreta-Veitía, M.; Montero-Torres, A.; Romero-Zaldivar, C.; Brandt, C. A.; Avila, P. E.; Kirchgatter, K.; Machado, Y. *J. Chem. Inf. Model.*, **2005**, 45, 1082.
[125] Dheyongera, J. P.; Geldenhuys, W. J.; Dekker, T. G.; Matsabisa, M. G.; Van der Schyf, C. J. *Bioorg. Med. Chem.*, **2005**, 13, 1653.
[126] Flipo, M.; Beghyn, T.; Leroux, V.; Florent, I.; Deprez, B. P.; Deprez-Poulain, R. F. *J. Med. Chem.*, **2007**, 50, 1322.
[127] Fattorusso, C.; Campiani, G.; Kukreja, G.; Persico, M.; Butini, S.; Romano, M. P.; Altarelli, M.; Ros, S.; Brindisi, M.; Savini, L.; Novellino, E.; Nacci, V.; Fattorusso, E.; Parapini, S.; Basilico, N.; Taramelli, D.; Yardley, V.; Croft, S.; Borriello, M.; Gemma, S. *J. Med. Chem.*, **2008**, 51, 1333.
[128] Mahmoudi, N.; Garcia-Domenech, R.; Galvez, J.; Farhati, K.; Franetich, J. F.; Sauerwein, R.; Hannoun, L.; Derouin, F.; Danis, M.; Mazier, D. *Antimicrob. Agents Chemother.*, **2008**, 52, 1215.
[129] Xie, A.; Sivaprakasam, P.; Doerksen, R. *Bioorg. Med. Chem.*, **2006**, 14, 7311.
[130] Güner, O.; Clement, O.; Kurogi, Y. *Curr. Med. Chem.*, **2004**, 11, 2991.
[131] Sun, H. *Curr. Med. Chem.*, **2008**, 15, 1018.
[132] Parenti, M. D.; Pacchioni, S.; Ferrari, A. M.; Rastelli, G. *J. Med. Chem.*, **2004**, 47, 4258.
[133] Schormann, N.; Senkovich, O.; Walker, K.; Wright, D. L.; Anderson, A. C.; Rosowsky, A.; Ananthan, S.; Shinkre, B.; Velu, S.; Chattopadhyay, D. *Proteins*, **2008**, 73, 889.
[134] Bhattacharjee, A. K.; Nichols, D. A.; Gerena, L.; Roncal, N.; Gutteridge, C. E. *Med. Chem.*, **2007**, 3, 317.
[135] Bhattacharjee, A. K.; Hartell, M. G.; Nichols, D. A.; Hicks, R. P.; Stanton, B.; van Hamont, J. E.; Milhous, W. K. *Eur. J. Med. Chem.*, **2004**, 39, 59.
[136] Dascombe, M. J.; Drew, M. G. B.; Morris, H.; Wilairat, P.; Auparakkitanon, S.; Moule, W. A.; Alizadeh-Shekalgourabi, S.; Evans, P. G.; Lloyd, M.; Dyas, A. M.; Carr, P.; Ismail, F. *J. Med. Chem.*, **2005**, 48, 5423.
[137] Drew, M. G. B.; Metcalfe, J.; Dascombe, M. J.; Ismail, F. M. D. *J. Mol. Struc. THEOCHEM*, **2007**, 823, 34.
[138] Congreve, M.; Murray, C. W.; Blundell, T. L. *Drug Discov. Today*, **2005**, 10, 895.
[139] Salwinski, L.; Eisenberg, D. *Curr. Opin. Struct. Biol.*, **2003**, 13, 377.
[140] Grosdidier, S.; Fernández-Recio, J. *BMC Bioinformatics*, **2008**, 9, 447.
[141] Robinson, J. A.; Demarco, S.; Gombert, F.; Moehle, K.; Obrecht, D. *Drug Discov. Today*, **2008**, 13, 944.
[142] Cheng, Y.; LeGall, T.; Oldfield, C. J.; Mueller, J. P.; Van, Y. Y.; Romero, P.; Cortese, M. S.; Uversky, V. N.; Dunker, A. K. *Trends Biotechnol.*, **2006**, 24, 435.
[143] Dunker, A. K.; Oldfield, C. J.; Meng, J.; Romero, P.; Yang, J. Y.; Chen, J. W.; Vacic, V.; Obradovic, Z.; Uversky, V. N. *BMC Genomics*, **2008**, 9 Suppl 2, S1.
[144] Chirac, P.; Torreele, E. *Lancet*, **2006**, 367, 1560.
[145] Trouiller, P.; Olliaro, P.; Torreele, E.; Orbinski, J.; Laing, R.; Ford, N. *Lancet*, **2002**, 359, 2188.
[146] Kapetanovic, I. M. *Chem. Biol. Interact.*, **2008**, 171, 165.

[147] Becker, O. M.; Dhanoa, D. S.; Marantz, Y.; Chen, D.; Shacham, S.; Cheruku, S.; Heifetz, A.; Mohanty, P.; Fichman, M.; Sharadendu, A.; Nudelman, R.; Kauffman, M.; Noiman, S. *J. Med. Chem.*, **2006**, 49, 3116.

[148] Gardner, M. J.; Hall, N.; Fung, E.; White, O.; Berriman, M.; Hyman, R. W.; Carlton, J. M.; Pain, A.; Nelson, K. E.; Bowman, S.; Paulsen, I. T.; James, K.; Eisen, J. A.; Rutherford, K.; Salzberg, S. L.; Craig, A.; Kyes, S.; Chan, M. S.; Nene, V.; Shallom, S. J.; Suh, B.; Peterson, J.; Angiuoli, S.; Pertea, M.; Allen, J.; Selengut, J.; Haft, D.; Mather, M. W.; Vaidya, A. B.; Martin, D. M. A.; Fairlamb, A. H.; Fraunholz, M. J.; Roos, D. S.; Ralph, S. A.; McFadden, G. I.; Cummings, L. M.; Subramanian, G. M.; Mungall, C.; Venter, J. C.; Carucci, D. J.; Hoffman, S. L.; Newbold, C.; Davis, R. W.; Fraser, C. M.; Barrett, T. *Nature*, **2002**, 419, 498.

[149] Carlton, J. M.; Adams, J. H.; Silva, J. C.; Bidwell, S. L.; Lorenzi, H.; Caler, E.; Crabtree, J.; Angiuoli, S. V.; Merino, E. F.; Amedeo, P.; Cheng, Q.; Coulson, R. M.; Crabb, B. S.; Del Portillo, H. A.; Essien, K.; Feldblyum, T. V.; Fernandez-Becerra, C.; Gilson, P. R.; Gueye, A. H.; Guo, X.; Kang'a, S.; Kooij, T. W.; Korsinczky, M.; Meyer, E. V.; Nene, V.; Paulsen, I.; White, O.; Ralph, S. A.; Ren, Q.; Sargeant, T. J.; Salzberg, S. L.; Stoeckert, C. J.; Sullivan, S. A.; Yamamoto, M. M.; Hoffman, S. L.; Wortman, J. R.; Gardner, M. J.; Galinski, M. R.; Barnwell, J. W.; Fraser-Liggett, C. M. *Nature*, **2008**, 455, 757.

[150] Pain, A.; Bohme, U.; Berry, A. E.; Mungall, K.; Finn, R. D.; Jackson, A. P.; Mourier, T.; Mistry, J.; Pasini, E. M.; Aslett, M. A.; Balasubrammaniam, S.; Borgwardt, K.; Brooks, K.; Carret, C.; Carver, T. J.; Cherevach, I.; Chillingworth, T.; Clark, T. G.; Galinski, M. R.; Hall, N.; Harper, D.; Harris, D.; Hauser, H.; Ivens, A.; Janssen, C. S.; Keane, T.; Larke, N.; Lapp, S.; Marti, M.; Moule, S.; Meyer, I. M.; Ormond, D.; Peters, N.; Sanders, M.; Sanders, S.; Sargeant, T. J.; Simmonds, M.; Smith, F.; Squares, R.; Thurston, S.; Tivey, A. R.; Walker, D.; White, B.; Zuiderwijk, E.; Churcher, C.; Quail, M. A.; Cowman, A. F.; Turner, C. M.; Rajandream, M. A.; Kocken, C. H.; Thomas, A. W.; Newbold, C. I.; Barrell, B. G.; Berriman, M. *Nature*, **2008**, 455, 799.

[151] Marechal, E. *Comb. Chem. High Throughput Screen.*, **2008**, 11, 583.

[152] Gorse, A. D. *Curr. Top. Med. Chem.*, **2006**, 6, 3.

[153] Ryan, T. P.; Stevens, J. L.; Thomas, C. E. *Curr. Opin. Pharmacol.*, **2008**, 8, 654.

[154] Schwaha, R.; Ecker, G. F. *Sci. Pharm.*, **2008**, 76, 5.

[155] Ioset, J.-R. *Curr. Org. Chem.*, **2008**, 12, 643.

[156] Butler, M. S. *Nat. Prod. Rep.*, **2008**, 25, 475.

[157] Quinn, R. J.; Carroll, A. R.; Pham, N. B.; Baron, P.; Palframan, M. E.; Suraweera, L.; Pierens, G. K.; Muresan, S. *J. Nat. Prod.*, **2008**, 71, 464.

[158] Rollinger, J. M.; Haupt, S.; Stuppner, H.; Langer, T. *J. Chem. Inf. Comput. Sci.*, **2004**, 44, 480.

[159] Rollinger, J. M.; Langer, T.; Stuppner, H. *Curr. Med. Chem.*, **2006**, 13, 1491.

[160] Plouffe, D.; Brinker, A.; McNamara, C.; Henson, K.; Kato, N.; Kuhen, K.; Nagle, A.; Adrián, F.; Matzen, J. T.; Anderson, P.; Nam, T. G.; Gray, N. S.; Chatterjee, A.; Janes, J.; Yan, S. F.; Trager, R.; Caldwell, J. S.; Schultz, P. G.; Zhou, Y.; Winzeler, E. *Proc. Natl. Acad. Sci. USA*, **2008**, 105, 9059.

[161] Stoeckert, C. J.; Fischer, S.; Kissinger, J. C.; Heiges, M.; Aurrecoechea, C.; Gajria, B.; Roos, D. *Trends Parasitol.*, **2006**, 22, 543.

[162] Cottingham, K. *J. Proteome. Res.*, **2008**, 7, 5073.

[163] Sayers, E. W.; Barrett, T.; Benson, D. A.; Bryant, S. H.; Canese, K.; Chetvernin, V.; Church, D. M.; DiCuccio, M.; Edgar, R.; Federhen, S.; Feolo, M.; Geer, L. Y.; Helmberg, W.; Kapustin, Y.; Landsman, D.; Lipman, D. J.; Madden, T. L.; Maglott, D. R.; Miller, V.; Mizrachi, I.; Ostell, J.; Pruitt, K. D.; Schuler, G. D.; Sequeira, E.; Sherry, S. T.; Shumway, M.; Sirotkin, K.; Souvorov, A.; Starchenko, G.; Tatusova, T. A.; Wagner, L.; Yaschenko, E.; Ye, J. *Nucleic Acids Res.*, **2009**, 37, D5.

Received: November 21, 2008 Revised: February 4, 2009 Accepted: February 5, 2009