Anti-malarial drug development using models of enzyme structure

Zhe Li[1], Xiaowu Chen[1], Eugene Davidson[4], Oren Zwang[4], Chandana Mendis[4], Christine S Ring[1], William R Roush[5], Glenn Fegley[5], Rongshi Li[1], Philip J Rosenthal[3], Garson K Lee[3], George L Kenyon[1], Irwin D Kuntz[1,2], and Fred E Cohen[1,2,3]*

[1]Departments of Pharmaceutical Chemistry, [2]Biochemistry and Biophysics and [3]Medicine, University of California, San Francisco, CA 94143–0446, USA, [4]Department of Biochemistry and Molecular Biology, Georgetown University, Washington DC 20007, USA and [5]Department of Chemistry, Indiana University, Bloomington, IN 47405–8300, USA

Background: The trophozoite stage of the malaria parasite infects red blood cells. During this phase of their lifecycle, the parasites use hemoglobin as their principal source of amino acids, using a cysteine protease to degrade it. We have previously reported a three-dimensional model of this cysteine protease, based on the structures of homologous proteases, and the use of the program DOCK to identify a ligand for the malaria protease.

Results: Here we describe the design of improved ligands starting from this lead. Ligand design was based on the predicted configuration of the lead compound docked to the model three-dimensional structure of the protease. The lead compound has an IC_{50} of 6 μM, and our design/synthesis strategy has resulted in increasingly potent derivatives that block the ability of the parasites to infect and/or mature in red blood cells. The two best derivatives to date have IC_{50}s of 450 nM and 150 nM.

Conclusions: A new class of anti-malarial chemotherapeutics has resulted from a computational search that was based on a model of the target protease. Despite the lack of a detailed experimental structure of the target enzyme or the enzyme–inhibitor complex, we have been able to identify compounds with increased potency. These compounds approach the activity of chloroquine (IC_{50} = 20 nM), but have a distinct mechanism of action. This series of compounds could thus lead to new therapies for chloroquine-resistant malaria.

Chemistry & Biology September 1994, **1**:31–37

Key words: docking strategies, homology modeling, novel anti-malarial compounds, structure–activity relationships, trophozoite cysteine protease

Introduction

The World Health Organization estimates that 280 million people are infected with malaria yearly [1]. Although various classes of anti-malarial agents exist, the most valuable are the quinoline-derived compounds, such as chloroquine and mefloquine. Chloroquine has been an especially effective drug in prophylactic and therapeutic settings, and thus the appearance of malaria strains resistant to chloroquine, and more recently, to mefloquine, poses a serious threat to travelers and to people in countries where malaria is endemic. Reports of multi-drug resistant strains of these parasites make the search for novel therapies especially urgent [2].

In this work, a novel target for anti-malarial chemotherapy is pursued. During the erythrocytic stage of their lifecycle, malaria parasites infect circulating red blood cells (RBCs) and use hemoglobin as their principal source of amino acids. Rosenthal and coworkers have identified a cysteine protease from malaria parasites that mediates host hemoglobin degradation [3]. Blocking this enzyme with cysteine protease inhibitors, such as L-*trans*-epoxysuccinylleucyl amido-(4-guanidino)butane (E-64) and peptide fluoromethyl ketones, arrests parasite growth *in vitro* [4] and cures murine malaria (*Plasmodium vinckei*) infection *in vivo* [5]. These results support the therapeutic rationale behind using the protease as a target. Inhibiting this protease may, either by itself or in conjunction with established therapies, offer an alternative treatment for malaria. Most cysteine protease inhibitors, however, are sulfhydryl reagents that covalently modify the enzyme active site, and thus may inactivate some host cysteine proteases. Noncovalent angiotensin-converting enzyme inhibitors have been successful in the treatment of hypertension [6], and the clinical success of these compounds has motivated us to develop a new class of compounds that inhibit the malarial cysteine protease through noncovalent interactions.

Results and discussion

Initial lead discovery

The discovery of a nonpeptidic lead compound, oxalic *bis*((2-hydroxy-1-naphthylmethylene)hydrazide) (ZLI48A), for the malaria cysteine protease has previously been described [7]. Briefly, a model structure for the malaria

*Corresponding author.

 ISSN 1074–5521

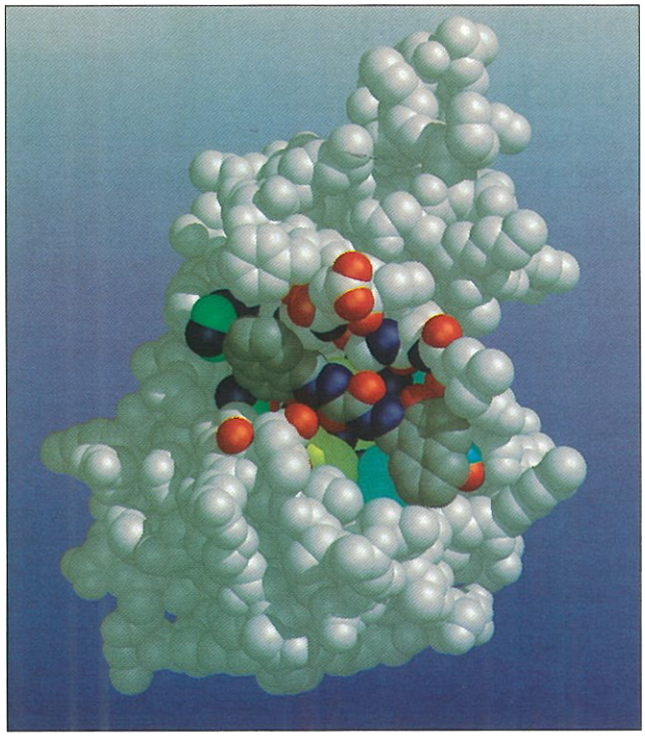

Fig. 1. The active site of the malarial protease and putative binding mode of the original lead compound, ZLI48A. The protease is shown in white and the ligand in gray. Oxygen and nitrogen atoms near the active site are colored in red and dark blue, respectively. The S_2, S_1, and S_1' binding subsites are shown in green, yellow, and cyan, respectively.

cysteine protease was proposed using the X-ray structures of papain and actinidin as a basis for modeling. The model structure was then used to search the Fine Chemicals Directory of commercially available small molecules (the Fine Chemicals Directory distributed by Molecular Design is currently known as the Available Chemical Directory) for putative ligands using the docking program, DOCK 3.0 [8]. Thirty-one compounds were finally tested and ZLI48A was identified as the best inhibitor of the protease. The IC_{50} value for enzyme inhibition against the substrate benzyloxycarbonyl-Phe-Arg-(7-amino-4-methylcoumarin) was 6 μM [7]. More importantly, this compound inhibits the growth of parasites in culture, as judged by its ability to block hypoxanthine uptake by malaria parasites, with an apparent IC_{50} value of 7 μM.

Circulating RBCs are a major target for malaria infection, and the magnitude of the infection can be quantified by counting the fraction of infected RBCs on a thick blood smear. This process is sufficiently laborious that it precludes the evaluation of dose-response curves for potential therapeutics in an acceptable time frame. To overcome this problem, we have developed a high-throughput assay that efficiently counts parasitized RBCs. Normally, circulating RBCs are enucleate and hence lack DNA. Parasitized RBCs contain malarial DNA. Thus, RBC staining by the DNA-binding dye propidium iodide can be exploited in an assay that uses a fluorescence-activated cell sorter (FACS) to count the fraction of infected RBCs in a rapid and automated fashion. In this assay, ZLI48A has an IC_{50} value of 4.5 μM.

Berger and Schechter [9] first demonstrated that papain-like cysteine proteases contain active sites that can accommodate up to seven amino acids. For notational convenience, the amino acid residues on the acyl side of the scissile bond are denoted P_1, P_2 . . . P_n and those on the leaving group side are labeled P_1', P_2' . . . P_n'. The corresponding binding sites on the enzyme are S_1, S_2 . . . S_n and S_1', S_2' . . . S_n'. The seven-residue binding pocket of papain-like cysteine proteases involves the S_4 through S_3' subsites. The S_2 and S_1 subsites are those most responsible for the peptide cleavage specificity of this class of enzymes. The most stable conformer of ZLI48A is symmetric about its midpoint, with a second axis of pseudosymmetry about the backbone. Thus, there are essentially two ways to orient the compound in the active site. In both cases, the compound lies across the active site cleft of the malaria cysteine protease with one naphthyl group fitting into the S_2 pocket and with the other stacking with the indole ring of Trp177 in the S_1' pocket. The presumed binding mode is shown in Fig. 1. This orientation was chosen as the working model for the orientation in the complex, because it maximizes the compound's interaction with the enzyme, with each hydroxyl group within hydrogen bonding distance of a suitable residue in the enzyme (Ser160 and Gln19). Alternatively, the compound could be rotated 180 degrees about the horizontal axis of pseudosymmetry. In this orientation, the hydroxyl groups seem to interact with the enzyme less effectively but might interact with solvent water molecules.

Lead optimization

The starting point for lead optimization was the protease–ZLI48A complex generated by the program DOCK 3.0 [10]. Kuntz and colleagues developed the DOCK algorithm to capture the static geometric features of a molecular recognition site. In the malarial protease work, the homology-based model of the enzyme provides the template, and the DOCK algorithm identifies a set of spheres with approximately atom-sized radii to fill the active site cleft. Frequently, many overlapping spheres are used to fill the cleft, and a clustering algorithm is used to reduce the complexity of the sphere representation. Preliminary binding modes for a compound are defined through attempts to match the centers of cleft spheres with the centers of atoms within the compound of interest. Within DOCK 3.0, the quality of a compound's fit to the binding cleft can be evaluated based on its shape complementarity (contact score) or molecular mechanics interaction energy (AMBER force field score). When searching a database of compounds, DOCK 3.0 examines only the best orientation of the small molecule within the binding cleft (DOCK database screening mode). When a single compound is studied, multiple possible binding modes can be examined (DOCK single mode). Of course, the initial orientation of the compound is dictated, in part,

by the irregular lattice of sphere centers identified originally. To overcome some of the scoring distortion that this bias could impart, a rigid body minimization algorithm has been developed to move the ligand modest distances within the binding cleft and optimize the interaction energies. The use of a rigid body minimization algorithm reflects the trade-off between the need for rapid evaluation and the reality that many ligands are flexible. ZLI48A was selected based on its score for shape complementarity. Rigid body minimization was applied to the ligand positioned in a preliminary orientation in the active site cleft so that other contributory factors (van der Waals and electrostatic interactions) required for efficient binding could be reflected in the enzyme–ligand complex (see Fig. 2).

The energetics of the enzyme–ligand assembly depends heavily upon the detailed conformation of the complex. Unfortunately, errors in the precise geometry of the complex are likely given the computational origins of the model. Potential problems could result from the fact that gross conformational changes can occur between the complexed and uncomplexed enzyme structures. In papain, the active site cleft widens upon binding a peptide-like ligand [11]. In the HIV protease, the flaps around the active site close on the substrate with backbone movements as large as 7 Å [12,13]. Ligands are also able to alter their conformations upon binding to the enzyme. For example, the structures of methotrexate alone and complexed to dihydrofolate reductase are dramatically different [14]. Consequently, choosing the relevant conformations of both enzyme and ligand can be critical for successful inhibitor design. Unfortunately, such structural changes are difficult to anticipate *a priori* and so our approach to ligand design is guided by, rather than dependent solely upon, the model of the ligand–enzyme structure.

Despite these uncertainties, a reasonable structure of the enzyme and/or enzyme–ligand complex has proven to be very useful. In this project, the malarial cysteine protease was modeled by homology to the known structures of papain and actinidin. The malarial protease has 33 % sequence identity with both enzymes. The anticipated root mean-square errors average approximately 1.5 Å for the core of the molecule [15]. Fortunately, the errors within the active site should be smaller, as it is the most conserved portion of the structure. The ligand conformation was calculated using CONCORD, an algorithm that relies on heuristic rules for translating chemical connectivities into three-dimensional coordinates [16]. This could create errors in the models of the ligand conformation. Because of the uncertainties inherent in the models of both the enzyme and the ligands, sophisticated energy calculations such as free energy perturbation are not warranted. The interpretation of these results would be difficult at best.

The model of the protease–ZLI48A complex has aryl groups filling the S_1' and S_2' pockets, but the S_1 pocket is only partially filled. In addition, although the conjugated eight-atom backbone of ZLI48A is most likely to be found in the all *anti*-conformation, *syn* isomers are possible that should not fit into the protease active site. We therefore designed analogs of ZLI48A that have a less flexible backbone, fill the S_2 and S_1 or S_2, S_1, and S_1' subsites, and effectively interact with the side chains lining the subsite specificity pockets. Because malaria is endemic principally in developing countries, useful anti-malarial agents must be inexpensive to produce. We thus favored analogs that are easily synthesized from relatively inexpensive starting materials in a small number of steps. The basic strategy for lead optimization is outlined in Fig. 3.

Starting with the rigid-body minimized complex of enzyme–ZLI48A, and using the strategy highlighted above, we explored the S_2, S_1, and S_1' subsites by varying the size of the aromatic rings, the linker lengths, and the substituents of the aromatic rings. A third aromatic ring system was also introduced (tri-aryl compounds) to study the concerted binding of the ligand molecule to the S_2, S_1 and S_1' subsites. Newly synthesized compounds were tested by FACS assay for their potency in inhibiting

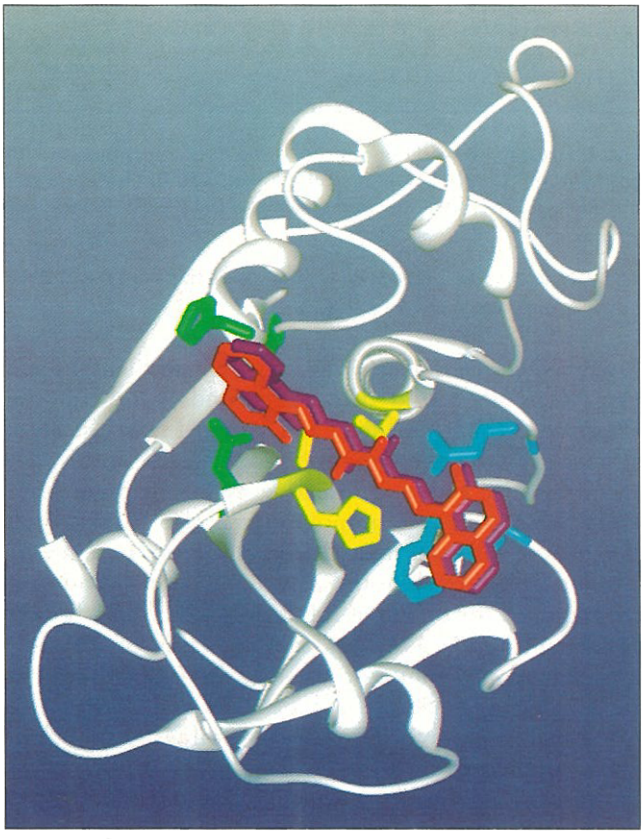

Fig. 2. Optimization of the orientation of ZL148A using rigid body minimization. The initial positioning of the ligand in the active site cleft (red) was estimated to have an interaction energy of 27.7 kcal mol^{-1} using the AMBER force field. After 12 steps of minimization, the ligand (purple) moved 0.74 Å and the energy was lowered to −35.3 kcal mol^{-1}. Although the change in the position of the ligand was small, the difference between the energies of the starting and final conformations is significant. Green, yellow, and cyan colors highlight key residues in the S_2, S_1, and S_1' binding subsites, respectively.

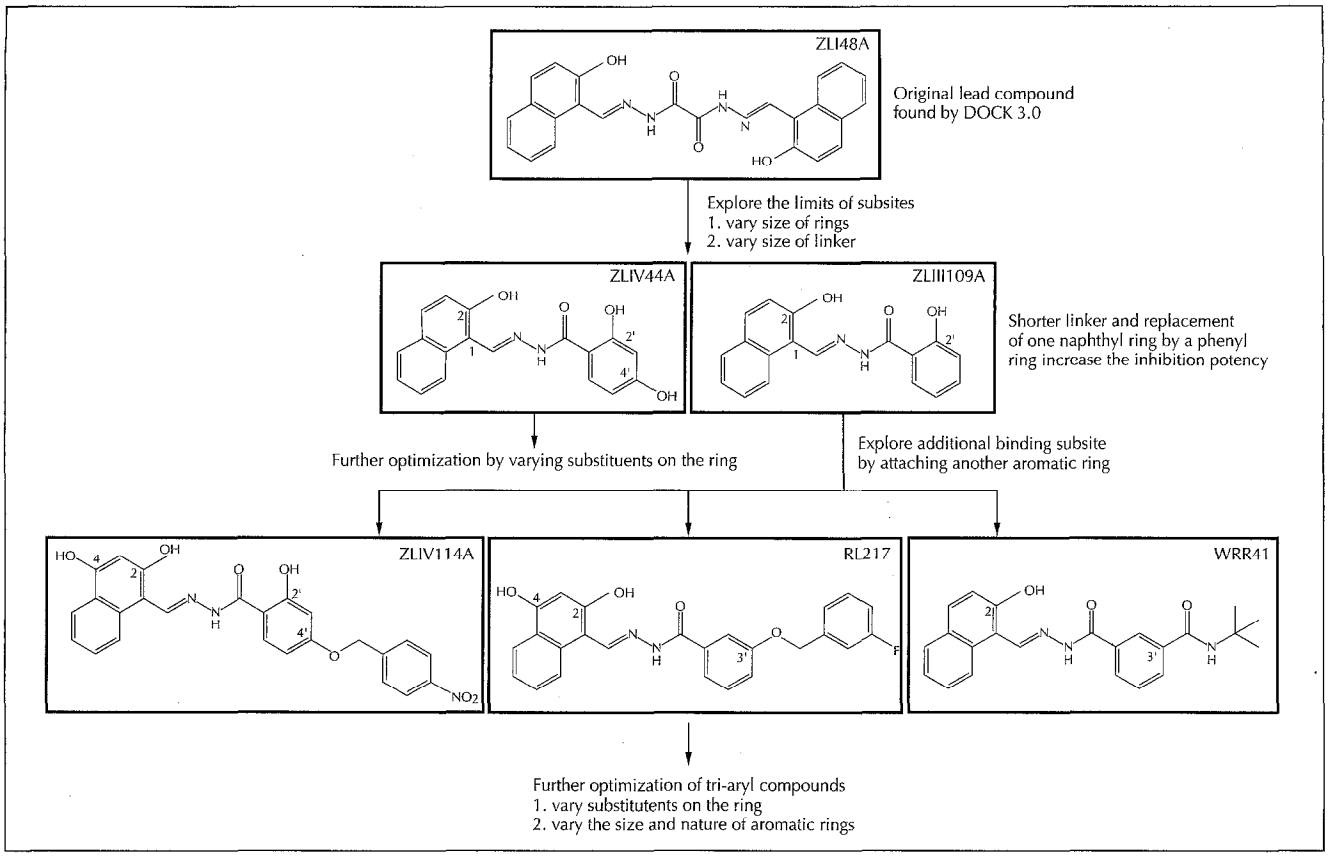

Fig. 3. Strategy for lead optimization. The fit of the compound within the S_2, S_1 and S_1' subsites was explored by varying the size of the aromatic rings and the length of the linker, and further optimized by varying ring substituents. A third aromatic system was also introduced to explore the concerted binding of the ligand to these subsites. Six compounds are shown in this chart to illustrate the strategy.

parasite growth. The assay results of some representative compounds are shown in Fig. 4.

From a study of analogs with varying linker lengths and ring sizes, we found that compounds with half the length of the original linker and a phenyl ring instead of a naphthyl ring (compound ZLIII109A in Fig. 4) seemed to be twice as effective as the original lead, ZLI48A. A series of compounds was also made to test the role of ring substituents on inhibitor potency. It seems that hydroxyl groups at the 2, 2', and 4' positions are required for effective inhibition. The most potent compound (ZLIV44A) has an IC_{50} value of 150 nM. The best tri-aryl inhibitor tested so far, ZLIV114A, has an IC_{50} value of 450 nM. Overall, the four compounds listed in Fig. 4 have absolute IC_{50} values at or below 1 μM. It is clear from Fig. 4 that there is a progressive increase in inhibition potency following the design, synthesis, and testing paradigm.

Modeling of ligand binding

DOCK calculations (in DOCK single mode) of the binding of a series of *bis* aromatic compounds with one naphthyl ring, a short linker, and one smaller aromatic ring reveal some interesting insights. The results indicate that there are essentially two ways to orient an asymmetric ligand. Both binding modes are similar to that for the symmetrical ZLI48A, shown in Fig. 1. In one binding mode (Fig. 5a), the larger naphthyl ring interacts with the deeper, well-defined S_2 specificity subsite, whereas the phenyl ring binds to the more accessible, but less well-defined, S_1' subsite. Alternatively, the ligand molecule can be rotated 180 degrees around the S_1 subsite (Fig. 5b). In this binding mode, the naphthyl ring now interacts with the S_1' subsite, whereas the phenyl ring binds to the S_1 and S_2 subsites. Compounds with three aromatic components seem more likely to adopt the binding mode shown in Fig. 5a. In this orientation, the larger naphthyl ring binds to the well-formed S_2 subsite and the third ring now interacts fully with the more accessible, less well-defined S_1' subsite.

A recent report on the structures of inhibitors bound to the serine protease elastase, with K_is as small as 10 nM, shows that chemically similar inhibitors can adopt different binding modes and interact with different subsites [17]. Several inhibitors were found to bind to the enzyme in an orientation opposite to that of natural substrate and other chemically similar inhibitors. The compounds with two-fold degeneracy that we have designed and tested may also bind to the malarial protease with more than one binding mode. Obviously, this could complicate the development of a structure–activity relationship. Further work is under way to explore the availability of alternative binding modes.

In either binding mode, it seems that polar substituents on either aromatic ring can potentially form hydrogen bonds with a number of side chains lining the binding subsites as well as with backbone oxygen and nitrogen atoms from some loop residues. This observation is consistent with our finding that the 2 and 2' hydroxyl groups are required for effective inhibition. Substitutions at other positions, such as 4 and 4', are also associated with increased potency. For example, calculation of the model of the protease–ZLIII109A complex suggests possible hydrogen bonding interactions between residues Gln19, His67, Cys25, Asn133, His159, Ser160, Trp177 and the ligand oxygen and nitrogen atoms. Presumably, this explains part of the correlation between these ligand hydroxyl groups and compound potency. The modeling studies also provide a basis for the design of other tri-aryl compounds, a group of ligands that could lead to even more potent inhibitors via their extensive interactions with the protease subsites.

Conclusions

We report here the results of a structure-based approach for inhibitor design targeted toward a critical malarial protease. Starting with a lead compound identified by computer screening of a three-dimensional small molecule database, we have designed molecules with significantly increased inhibition potency against malaria parasites. This method provides a practical strategy for future work on structure-based drug design even in the absence of a crystallographic structure of the target enzyme. Our approach not only emphasizes the structural rationale for inhibitor design, but also uses simple chemistry and commercially available starting materials for inhibitor synthesis. This approach permits us to screen a relatively large number of potential inhibitors rapidly and inexpensively.

The potency of this new class of compounds in preventing parasite growth *in vitro* begins to approach that of available traditional quinoline-based drugs. As their target is distinct from that of chloroquine, these compounds should be effective against chloroquine-resistant strains of the malaria parasite. The FACS assay we used in this report provides a practical and relevant method for evaluating inhibitor potency. The combination of structure-based screening and design coupled with simple chemical synthesis and a relevant biological assay has rapidly led to identification of a series of increasingly potent anti-malarial agents. Although therapeutic candidates will need to be active at low nanomolar concentrations, we anticipate that modifications of existing analogs will result in molecules that are suitable for preclinical efficacy and toxicology studies.

Significance

Malaria parasites, usually *Plasmodium falciparum* and *Plasmodium vivax*, infect 280 million people annually [1], and multi-drug resistant *P. falciparum* has become a significant pathogen in areas where the disease is endemic. We have chosen a cysteine protease that is central to the parasite's life cycle but distinct from the target of chloroquine action as a target for a structure-based drug development program. The choice of this target means that the chemotherapeutics that we develop should be active against chloroquine-resistant organisms.

Structure-based drug design usually depends upon the experimental determination of the target structure by X-ray crystallography or NMR spectroscopy. We have circumvented this step and relied exclusively on the sequence homology between the malaria enzyme and other cysteine proteases of known structure. A homology-based

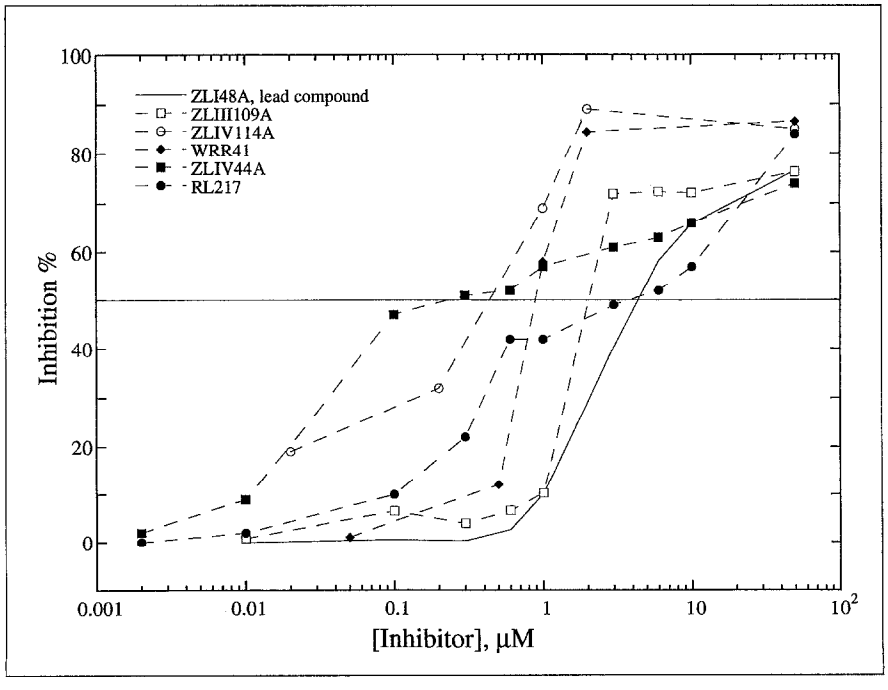

Fig. 4. Inhibition of the growth of malaria parasites in red blood cells by ZLI48A, the initial lead compound, and selected derivatives.

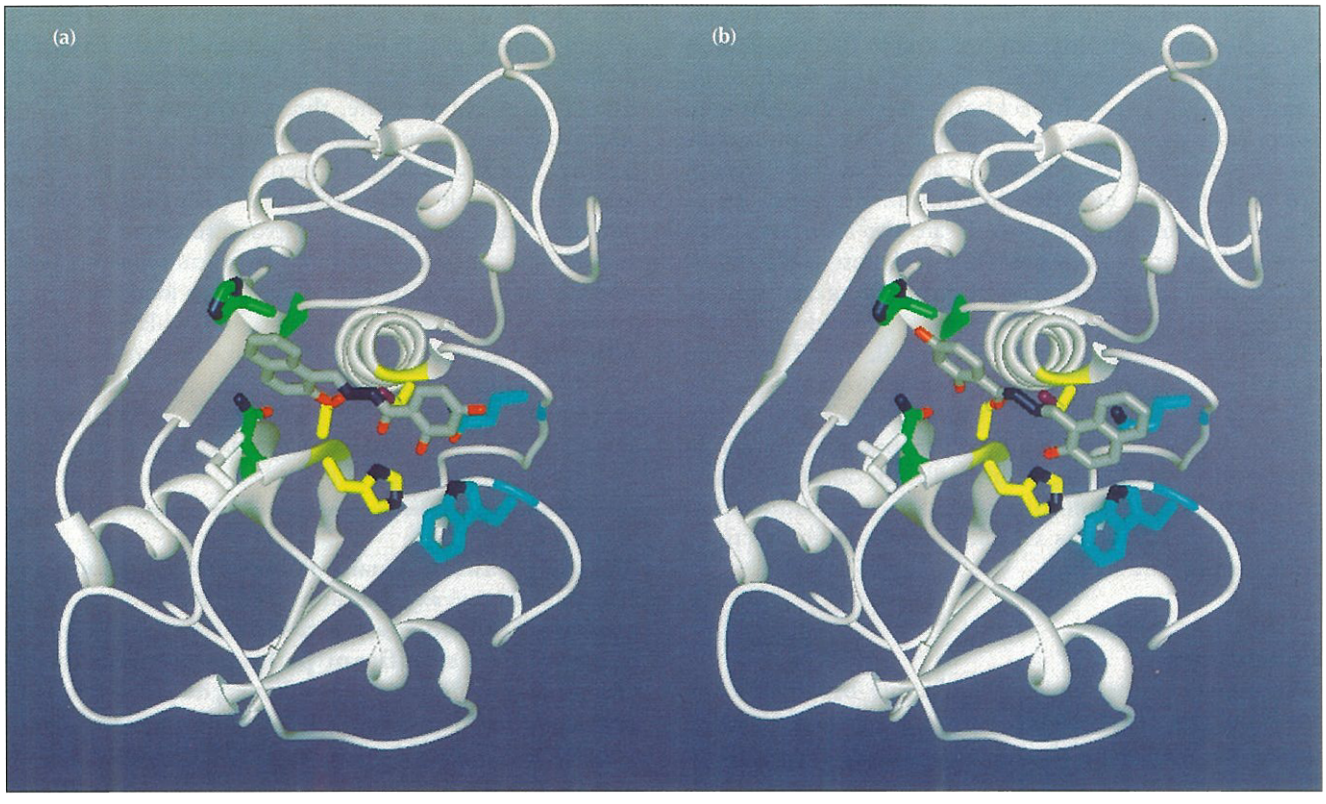

Fig. 5. Two putative binding modes of a representative *bis*-aryl compound ZLIV44A. In **(a)** the larger naphthyl ring interacts with S_2, whereas in **(b)** it interacts with the S_1 subsite. The protease is shown in white and the ligand in gray. Oxygen and nitrogen atoms near the active site are colored in red and dark blue, respectively. The S_2, S_1, and S_1' binding subsites are shown in green, yellow, and cyan, respectively.

model of the malaria enzyme was used as the template for a computer-based ligand docking calculation that identified a useful lead compound. Lead optimization was achieved by a combined approach of computational and synthetic analysis. Derivatives of the lead were first optimized for fit using the computer docking program, then synthesized and experimentally tested.

The entire process of lead optimization was completed in eighteen months and resulted in the identification of compounds that are about ten times more effective than the initial lead and are simple and inexpensive to produce. This study shows that it is possible to start with a protein sequence and end with non-peptidic small molecule inhibitors of a medically relevant target enzyme. Thus, there may be many more systems amenable to structure-based drug development efforts than was previously believed.

Materials and Methods

Computer modeling

The three-dimensional structures of potential inhibitors were constructed interactively using the molecular modeling program SYBYL and the CONCORD conversion algorithm (Tripos Associates, St Louis, MO). The Cambridge Crystallographic Database was used to determine probable low-energy conformations of certain ligand groups. Partial charges of ligands were calculated using the Gasteiger–Marsili method. DOCK 3.0 was used in single ligand mode to aid inhibitor design. All plausible binding orientations of a single ligand that meet certain user-defined criteria (contact and force field scores) were obtained by DOCK. Rigid body minimization of DOCK-derived ligand–protease complexes were performed to optimize the interaction energy between the protease and the ligand. This method provides a detailed profile of potential binding orientations of a ligand molecule in the active site of the malarial protease. All computer-assisted modeling and docking studies were carried out using a Silicon Graphics workstation IRIS4D/35 or Indigo2. All color figures were produced using the MidasPlus program from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH RR-01081) [18,19].

Biological assay

An assay based on FACS analysis was used to evaluate the potency of the compounds reviewed in this paper against parasite growth. Synchronized trophozoite-stage parasites were cultured in human blood at various inhibitor concentrations. The parasites were allowed to mature, lyse the host cell and attempt invasion of fresh red blood cells. Using propidium iodide to stain DNA, the FACS can discriminate between infected and uninfected cells and between stages of intra-erythrocytic parasite development, as only infected red blood cells contain DNA [20].

Because all of the clinical manifestations of malaria are caused by the erythrocytic cycle of lysis and reinfection, this assay is especially relevant for evaluating the efficacy of potential anti-malarial agents. The full dose inhibition curves for oxalic *bis*((2-hydroxy-1-naphthylmethylene)hydrazide) and several of the key derivatives are shown in Fig. 4.

Chemical syntheses

2-Hydroxy-1-naphthaldehyde, oxalic dihydrazide, salicylic hydrazide, methyl 2,4-dihydroxybenzoate and 4-nitrobenzyl bromide were all purchased from Aldrich. 2,4-Dihydroxybenzoic hydrazide was obtained from Trans World Chemicals. 2,4-Dihydroxynaphthaldehyde was prepared from 1,3-dihydroxynaphthalenes (Aldrich) according to published procedures [21].

General procedure for condensation of aldehyde with acylhydrazine: to a solution of the aldehyde (1 mmol) in methanol (20 ml) was added the corresponding acylhydrazine (1 mmol) in one portion. The resulting mixture was heated at reflux for 3 h. In most cases, a precipitate was observed after 10 min. The precipitate was filtered, washed with hot methanol (50 ml) and dried in vacuum (2 mm Hg). If needed, additional purification was performed by recrystallization using appropriate solvents.

For the tri-aryl compounds, the acylhydrazine was made as follows: a mixture of benzyl bromide (10.0 mmol), methyl hydroxybenzoate (10.0 mmol) and cesium carbonate (10.0 mmol) in acetone (80 ml) was heated at 56 °C for 18 h. This mixture was filtered, and the filtrate was concentrated to give the corresponding methyl phenylmethyleneoxybenzoate. This crude methyl ester was then dissolved in EtOH (80 ml) and treated with hydrazine monohydrate (5.01 g, 100.0 mmol). The resulting mixture was stirred overnight at 20 °C and then concentrated to give the corresponding acylhydrazide, which was further purified by recrystallization from EtOH/H_2O (7:3, v/v). The structures of these compounds were confirmed by spectroscopic methods.

Acknowledgements: We thank Dr Elaine Meng and Daniel Gschwend for useful comments on DOCK simulations. This work was supported by grants from the Advanced Research Projects Agency (MDA-972-91-J1013; N00014-90-2032), the National Institutes of Health (GM 39900), and the UNDP/World Bank/WHO Special Program for Research and Training in Tropical Diseases (TDR890499). The MidasPlus program from the Computer Graphics Laboratory, University of California, San Francisco was supported by the National Institutes of Health (RR-01081). Molecular Design Ltd Information Systems and Tripos Associates kindly provided software materials.

References

1. Gibbons, A. (1992). Researchers fret over neglect of 600 million patients. *Science* **256**, 1135.
2. Bruce-Chwatt, L.J. (1985). *Essential Malariology*. (2nd edn), Wiley, New York.
3. Rosenthal, P.J., McKerrow, J.H., Aikawa, M., Nagasawa, H. & Leech, J.H. (1988). A malarial cysteine proteinase is necessary for hemoglobin degradation by *Plasmodium falciparum*. *J. Clin. Invest.* **82**, 1560–1566.
4. Rosenthal, P.J., Wollish, W.S., Palmer, J.T. & Rasnick, D. (1991). Antimalarial effects of peptide inhibitors of a *Plasmodium falciparum* cysteine proteinase. *J. Clin. Invest.* **88**, 1467–1472.
5. Rosenthal, P.J., Lee, G.K. & Smith, R.E. (1993). Inhibition of a *Plasmodium vinckei* cysteine proteinase cures murine malaria. *J. Clin. Invest.* **91**, 1052–1056.
6. Williams, G.H. (1988). Converting-enzyme inhibitors in the treatment of hypertension. *New England J. Med.* **319**, 1517–1525.
7. Ring, C.S., *et al.*, & Cohen, F.E. (1993). Structure-based inhibitor design using model built structures. *Proc. Natl. Acad. Sci. USA* **90**, 3583–3587.
8. Kuntz, I.D. (1992). Structure-based strategies for drug design and discovery. *Science* **257**, 1078–1082.
9. Berger, A. & Schechter, I. (1970). Mapping the active site of papain with the acid of peptide substrates and inhibitors. *Philos. Trans. R. Soc. Lond. [Biol.]* **257**, 149–264.
10. Meng, E.C., Shoichet, B.K. & Kuntz, I.D. (1992). Automated docking with grid-based energy evaluation. *J. Comp. Chem.* **13**, 505–524.
11. Drenth, J., Kalk, K.H. & Swen, H.M. (1976). Binding of chloromethyl ketone substrate analogues to crystalline papain. *Biochemistry* **15**, 3731–3738.
12. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J. & Kent, S.B. (1988). Conserved folding in retroviral proteases: crystal structure of synthetic HIV-1 protease. *Science* **245**, 616–621.
13. Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R. & Clawson, L. (1989). Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* **246**, 1149–1152.
14. Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. & Kraut, J. (1982). Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. I. General features and binding of methotrexate. *J. Biol. Chem.* **257**, 13650–13662.
15. Chothia, C. & Lesk, A. (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5**, 823–826.
16. Pearlman, R.S. (1991). 3D-searching: an overview of a new technique for computer-assisted molecular design. *Nida Research Monograph* **112**, 62–77.
17. Mattos, C., Rasmussen, B., Ding, X., Petsko, G.A. & Ringe, D. (1994). Analogous inhibitors of elastase do not always bind analogously. *Nature Struct. Biol.* **1**, 55–58.
18. Ferrin, T.E., Huang, C.C., Jarvis, L.E. & Langridge, R. (1988). The Midas display system. *J. Mol. Graph.* **6**, 13–37.
19. Huang, C.C., Pettersen, E.F., Klein, T.E., Ferrin, T.E. & Langridge, R. (1991). Conic: A fast renderer for space-filling molecules with shadows. *J. Mol. Graph.* **9**, 230–236.
20. Clark, D.L., Chrisey, L.A., Campbell, F.R. & Davidson, E.A. (1994). Non-sequence specific antimalarial activity of oligodeoxynucleotides. *Molec. Biochem. Parasitol.* **63**, 129–134.
21. Morgan, G.T. & Vining, D.C. (1921). Dihydroxynaphthaldehydes. *J. Chem. Soc.* **119**, 177–187.

Received: **13 Apr 1994**; revisions requested: **29 Apr 1994**; revisions received: **17 Jun 1994**. Accepted: **22 Jun 1994**.