This is the Author’s version of the paper published as:

**Author:** T. Spithill, S. Yanow, L. Purcell and M. Lee  
**Author Address:** tspithill@csu.edu.au  
**Title:** Genomics-based drug design targets the AT-rich malaria parasite: implications for anti-parasite chemotherapy  
**Year:** 2007  
**Journal:** Pharmacogenomics  
**Volume:** 8  
**Pages:** 1267-1272  
**ISSN:** 1462-2416  
**URL:** Keywords: malaria, Plasmodium, drug design, genomics, chemotherapy, DNA, review  
**Abstract:** Evaluation of: Woynarowski JM, Krugliak M, Ginsburg H: Pharmacogenomic analyses of targeting the AT-rich malaria parasite genome with AT-specific alkylating drugs. Mol. Biochem. Parasitol. 154(1), 70-81 (2007) [1]. The sequencing of the malaria genome sought to expose the parasite's ability to cause disease and identify new targets for antimalarial drugs and vaccines. In this study, the authors discovered how malaria genomic DNA, which is unusually rich in adenine and thymine nucleotides, is intrinsically a target for a selective class of compounds. AT-specific DNA-binding agents have previously been shown to have potent antimalarial activity in vitro. The authors used high-resolution bioinformatic tools to explore the genomic basis for this drug susceptibility, first at the level of individual DNA-binding sites, then expanding to the entire genomic context of each malaria chromosome. Their findings revealed a nonrandom distribution and organization of drug-binding sites that can be further exploited to target these AT sequences. Based on these findings, comparative bioinformatics analyses with other parasite genomes may lead to the identification of new target organisms for these AT-specific drugs and have wide implications for the treatment of human and animal parasitic diseases.
Genomics-based drug design targets the AT-rich malaria parasite: implications for anti-parasite chemotherapy

Stephanie K. Yanow¹, Lisa A. Purcell², Moses Lee³, and Terry W. Spithill², ⁴*

¹Provincial Laboratory for Public Health, WMC Rm 2B4.59, 8440 112th Street, Edmonton, AB, Canada, Tel: +1 780 407 7558, Fax: +1 780 407 3864, s.yanow@provlab.ab.ca
²McGill University, Institute of Parasitology and Centre for Host-Parasite Interactions, 21,111 Lakeshore Road, Ste. Anne-de-Bellevue, Qc, Canada, H9X 3V9, Tel: +1 514 398 8668, Fax: +1 514 398 7857, lisa.purcell@mail.mcgill.ca; terry.spithill@mcgill.ca
³Hope College, Division of Natural and Applied Sciences & Department of Chemistry, 35E 12th Street, Holland, MI, 49423, Tel: +1 616 395 7190, Fax: +1 616 395 7923, lee@hope.edu
⁴School of Agricultural and Veterinary Sciences, Charles Sturt University, Wagga Wagga, NSW, Australia, 2678

* Author for correspondence

Abstract

The sequencing of the malaria genome sought to expose the parasite’s ability to cause disease and identify new targets for anti-malarial drugs and vaccines. In this study, the authors discovered how the malaria genomic DNA, which is unusually rich in adenine and thymine nucleotides, is intrinsically a target for a selective class of compounds. AT-specific DNA binding agents have previously been shown to have potent anti-malarial activity in vitro. The authors used high-resolution bioinformatic tools to explore the genomic basis for this drug susceptibility, first at the level of individual DNA binding sites then expanding to the entire genomic context of each malaria chromosome. Their findings revealed a non-random distribution and organization of drug binding sites that can be further exploited to target these AT sequences. Based on these findings, comparative bioinformatics analyses with other parasite genomes may lead to the
identification of new target organisms for these AT-specific drugs and have wide implications for the treatment of human and animal parasitic diseases.

**Keywords:** malaria, *Plasmodium falciparum*, genomics, *in silico*, AT-rich, adozelesin, bizelesin, matrix attachment regions (MARs), super-AT islands, parasite

**Introduction**

The results from a recent pharmacogenomic analysis of drugs that target the AT-rich genomic DNA of the malaria parasite have thrown the spotlight on a new chemotherapeutic approach to control this significant human pathogen [1]. Malaria, the most deadly parasitic infectious disease worldwide [2], is caused by the parasite *Plasmodium* that infects humans via the bite of an infectious mosquito. Treatment for malaria exists but the arsenal of anti-malarial drugs is limited. Widespread drug resistance has rendered many anti-malarial drugs ineffective and chemotherapy now requires a cocktail of drugs [3]. In order to re-gain a stronghold on the disease, new strategies for anti-malarial drug design are needed. One approach is to exploit the extraordinary AT-richness (~80%; [4]) of the *Plasmodium* genome with AT-specific DNA binding agents. A number of AT-specific compounds have been tested against *Plasmodium* over the years, many exhibiting significant anti-plasmodial activity (Table 1). While the rationale for this approach was intuitively based on preferential sequence specificity in the parasite compared to the host, Woynarowski and colleagues have now validated this strategy using advanced pharmacogenomic methods. Their findings provide
in silico support for this approach and strengthen the rationale to further develop this family of anti-malarials.

**Results**

The authors examined the frequency and distribution of binding sites for two AT-specific compounds, adozelesin and bizelesin, which have low *in vitro* IC$_{50}$ values against *P. falciparum*, the most lethal species of malaria [1, 5]. Both adozelesin and bizelesin recognize selective sequences with the following AT-motifs: (A/T)$_3$A and T(A/T)$_4$A, respectively (Table 1). Using computational methods previously developed by the group for scanning the human genome, the authors searched for these binding sites within the entire genome of *P. falciparum* (~23 Mbp). They calculated the frequency of these binding sites within the parasite genome relative to the human (host) genome and, for both of these compounds, the ratio strongly favoured the parasite. For adozelesin, the frequency of binding sites in the parasite is 3.9-fold greater than in humans and for bizelesin it is 7.0-fold higher. While the frequency of these sites is similar across all 14 chromosomes of *P. falciparum*, the distribution within each chromosome is clearly not random. The majority of these binding sites localize within clusters of excessively AT-rich DNA sequences which they termed ‘super-AT islands’. These regions are approximately 1-4 kbp in length, nearly 100% AT and highly repetitive in sequence. Most striking was the discovery that each of the 14 chromosomes contains a single, unique super-AT island [1].

What is the function of these super-AT islands? Previous work by the authors led them to hypothesize that these AT clusters may be associated with matrix attachment
regions (MARs) [6]. These are sequences in the DNA that facilitate the attachment of chromatin loops to the nuclear matrix and play an important role in nuclear structure. Using an algorithm to identify MARs within the *P. falciparum* genome, they discovered that each chromosome contains regions with putative high MAR potential along with a single ‘super-MAR’ domain that is highly reminiscent of the super-AT islands. In fact, there is precise overlap between the super-AT islands and the predicted super-MAR domains. With further bioinformatic analyses, these regions also co-localize with putative centromeres. These super-AT islands and super-MAR domains contain the highest concentration of adozelesin binding sites, suggesting a preferential localization for adozelesin in regions containing highly repetitive DNA, high AT content and a sparseness of genes and nucleosomes [1].

**Significance and conclusions**

The development of exquisitely refined bioinformatic tools by Woynarowksi and colleagues to mine the genome of the malaria parasite has revealed a highly unusual chromosomal organization in *Plasmodium* that could not have been anticipated based on sequence analysis alone. The concordance of adozelesin binding sites, the super-AT islands and the predicted super-MAR domains highlights the importance of genomic context in the evaluation of specific DNA sequences. This triple co-localization is impressive given the heterogeneity and generally poor quality of sequencing data from AT-rich DNA, especially within centromeric and intergenic regions. This is particularly challenging with *P. falciparum*, given its high AT content [4]. It is remarkable that four different algorithms (Msearch for identifying adozelesin binding sites, Oligo for
calculating %AT within 50 bp segments, Tandem Repeats Finder to determine consensus motifs within repetitive regions, and MAR Finder for identifying putative MARs) precisely identified these overlapping sequences within one region on each malaria chromosome.

These findings provide a genomic basis to explain the anti-plasmodial activity of this AT-binding class of drugs. Given the concentration of drug binding sites within these regions, multiple drug adducts could form on the DNA in the super-AT islands/super-MARs domains. Accessibility of compounds to their target sequence would be facilitated by the lack of nucleosomes in these centromere-like regions. The covalent binding of these compounds would ensure irreversibility and enhance their lethality against the parasite. This lethality could also be compounded by an inability of the parasite to excise drug adducts from the DNA. In human cells, adozelesin adducts are effectively repaired by the nucleotide excision pathway [7]. In Plasmodium, only a base excision repair pathway has been described thus far [8]. If the parasite lacks the enzymatic machinery required to remove these AT adducts, it may be more vulnerable to these agents. In addition, inefficient DNA repair has been proposed as one mechanism contributing to the emergence of drug-resistance in Plasmodium [9]. If the adozelesin adducts are never recognized by the Plasmodium repair machinery, and the effect of drug binding is highly lethal, the parasite will die without undergoing mutagenesis and thus will not contribute to the parasite population. Furthermore, the fact that there are multiple binding sites for adozelesin and bizelesin and the majority of these are localized within highly repetitive, non-coding DNA also reduces the probability of generating adaptive mutations in P. falciparum.
One of the concerns associated with this family of DNA binding agents is the potential for mutagenic effects on the host DNA. However, the greater abundance of adozelesin and bizelesin binding motifs in the parasite genome compared to the human genome, and the high lethality of these compounds for host cells, counters this argument. While the results imply that these compounds would preferentially bind to parasite DNA over human DNA, the potential for mutagenic effects on the host DNA must be further examined using animal models. The pharmacological properties of each compound and the dosing regimen used for treatment will also be important factors for consideration. If a strategy embracing drugs that target AT-rich sequences is adopted for malaria chemotherapy, the specific binding site of each individual compound must be considered within the context of the *P. falciparum* genome. For example, Woynarowski and colleagues found that the binding sites for another AT-specific agent, tallimustine, had a very low frequency within the malaria genome. This suggests that a simple preference for AT regions may not be a sufficient determinant for a successful anti-malarial compound. In contrast, the low frequency of GC-motifs in the *P. falciparum* genome clearly penalizes this class of compounds over AT-specific agents as anti-malarials.

**Future perspectives and implications for anti-parasitic drug discovery**

**Malaria:** AT-rich sequences within the minor groove of DNA are important targets for the design and synthesis of novel medicinal agents [10]. For *Plasmodium*, the activity of AT-specific agents that target these regions has been recognized for decades (Table 1). However, the precise contribution of sequence selectivity to their anti-plasmodial activity remained unclear. The binding sites for many of these compounds overlap with the target
site for adozelesin, suggesting they may also preferentially bind to super-AT islands and super-MAR domains within *Plasmodium*. Yet in terms of drug development, few of these compounds have progressed past the preclinical stage. Factors such as potency and selective uptake of the compound by parasites, shelf-life, cost, accessibility of the synthetic chemistry, as well as toxicity of the agent can all limit the therapeutic potential of an anti-malarial. Even though adozelesin and bizelesin are extremely potent anti-plasmodial agents, they are too toxic for development as anti-malarials. They exhibit severe toxicity to the bone marrow in animals and in human anticancer trials [11-13]. Medicinal chemists have instead focused on simplifying the structures related to adozelesin and bizelesin, retaining biological potency as well as AT-sequence specificity and reactivity, yet lowering systemic toxicity. The new compound centanamycin, an achiral *seco*-amino analog of (+)-duocarmycin SA, has emerged as the next incarnation of adozelesin and bizelesin-like agents [14]. This achiral compound, which is chemically stable, displayed potent activity against *P. falciparum* in culture (IC$_{50}$ 1.8 nM), and did not show appreciable toxicity to mice at a dose of 15 mg/kg (i.p.) [15]. At this dose, given in a single administration, centanamycin cured mice infected with various strains of *Plasmodium* and also blocked the transmission of parasites in mosquitoes [15].

Centanamycin displays excellent features of a potentially useful anti-malarial drug that may also be targeting super-AT islands/super-MAR domains in DNA: if so, centanamycin may impede the attachment of the parasite DNA to the nuclear matrix, disrupt chromosome organization within the nucleus and/or interfere with centromere function.
Another prominent AT-specific compound that has entered the anti-malarial pipeline is furamidine or DB75 [16-18]. It binds tightly to the minor groove of AT-rich sequences and exhibits potent anti-plasmodial activity (Table 1). However, due to its poor oral bioavailability, a pro-drug (DB289) has been developed and entered clinical trials [19]. In a recent study supported by the MMV and Immtech International, a 96 % cure rate was observed against *P. vivax* and *P. falciparum* with a treatment regimen of three doses per day for five days [20]. Unfortunately, treatment failure was observed when DB289 was administered with a regimen of one dose daily for three days [101]. Nevertheless, DB289 is currently undergoing a Phase II trial for malaria prevention and new combinations of dications are currently in the discovery stage for malaria therapy [101, 102].

**Other parasitic diseases:** AT-specific DNA binding drugs may have broad application as anti-parasitics since many parasites that cause significant disease in humans or animals (such as *Brugia malaya*, *Strongyloides stercoralis*, *Dirofilaria immitis*, *Cryptosporidium parvum*, *Theileria parva*, and the human body louse, *Pediculus humanus humanus*) exhibit AT-rich genomic DNA [21-24; S.C. Barker, personal communication]. Sequencing of the genomes from a range of clinically important parasite species is in progress and may reveal the presence of super-AT islands and the frequency of potential binding sites for AT-specific drugs in these genomes [103]. For example, the chromosomes of *T. parva*, a major cause of disease in cattle in Africa, each exhibit a >97% AT island of about 3kb in size which may represent the centromere, suggesting this parasite may be susceptible to AT binding drugs [22]. The genome of *B. malayi* (causative agent of lymphatic filariasis) encodes a Hha 1 repeat (79% AT) of unknown
function which comprises 9% of the genome [24]: such a repeat is likely to be targeted by AT-specific drugs. Other putative targets for AT binding drugs are the parasite organelle genomes such as the apicoplast and mitochondrial DNAs (mt DNA). The apicoplast genome of malaria is 86% AT [25] and the mitochondrial DNA of many parasites is also AT rich: malaria mtDNA is 69% AT [25] and the mtDNAs of trematode and cestode parasites is up to 74% AT [26]. Interestingly, the anti-malarial drug DB289 has broad anti-parasitic activity and is also effective against *Trypanosoma* and *Leishmania* species [16, 17]. The genomes of these parasites are not AT-rich, yet they possess an AT-rich mitochondrial genome (termed maxicircle kinetoplast DNA) critical for their survival [27]. Bioinformatic analyses of maxicircle DNA using the algorithms developed by Woynarowski and colleagues may reveal multiple binding sites for DB75 and identify new target organelles for these AT-specific agents within parasites whose genomes are not especially AT-rich. In conclusion, while the work by Woynarowski and colleagues is limited to *in silico* validation of drug target sites in malaria, it may have a profound impact on the design of new generations of drugs targeting AT sequences in other parasite species that cause significant disease in humans and animals.
<table>
<thead>
<tr>
<th>Compound</th>
<th>Sequence motif(^a)</th>
<th>Bond type(^b)</th>
<th>Mode of DNA binding</th>
<th>IC(_{50}) against P.f.</th>
<th>Drug discovery stage</th>
<th>Limitations of therapeutic use</th>
<th>Refs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adozelesin</td>
<td>(A/T)$_3$A</td>
<td>Covalent</td>
<td>Adduct formation</td>
<td>0.07 nM</td>
<td>Phase II trial for cancer; not pursued Preclinical testing in animal models of malaria</td>
<td>Myelotoxicity</td>
<td>[1, 5]</td>
</tr>
<tr>
<td>Bizelesin</td>
<td>T(A/T)$_4$A</td>
<td>Covalent</td>
<td>Interstrand crosslinking</td>
<td>0.01 nM</td>
<td>Phase I trial for cancer \textit{In vitro} testing for malaria</td>
<td>Myelotoxicity</td>
<td>[1, 28, 29]</td>
</tr>
<tr>
<td>Centanamycin</td>
<td>AAAAA</td>
<td>Covalent</td>
<td>Adduct formation</td>
<td>1.8 nM</td>
<td>Preclinical testing in animal models of malaria</td>
<td>Not determined</td>
<td>[14, 15]</td>
</tr>
<tr>
<td>Furamidine (DB75)</td>
<td>AAAA and ATTA</td>
<td>Non-covalent</td>
<td>Unknown(^c)</td>
<td>15.5 nM</td>
<td>\textit{In vitro} testing for malaria Pro-drug DB289 pursued</td>
<td>Lacks oral bioavailability</td>
<td>[16-19]</td>
</tr>
<tr>
<td>Pafuramidine (DB289; pro-drug of Furamidine)</td>
<td>Not applicable</td>
<td>Not applicable</td>
<td>Not applicable</td>
<td>11400 nM</td>
<td>Phase IIb trial for malaria treatment Phase II trial for malaria prevention</td>
<td>Ineffective at once daily, 3-day regimen for malaria therapy</td>
<td>[17-20, 30, 101, 102]</td>
</tr>
<tr>
<td>Pentamidine</td>
<td>AATT</td>
<td>Non-covalent</td>
<td>DNA Intercalation(^c)</td>
<td>58.4 nM</td>
<td>Phase II proof-of-concept study for malaria Clinically used for treatment of \textit{T.b.g.}, antimony-resistant Leishmaniasis and AIDS-related \textit{P. jiroveci} pneumonia</td>
<td>High toxicity, poor oral bioavailability</td>
<td>[16-18, 31]</td>
</tr>
<tr>
<td>Distamycin A</td>
<td>AAATTT</td>
<td>Non-covalent</td>
<td>Reversible</td>
<td>618 nM</td>
<td>\textit{In vitro} testing for malaria</td>
<td>High toxicity</td>
<td>[32, 33]</td>
</tr>
<tr>
<td>Compound</td>
<td>AATT</td>
<td>Mode of Binding</td>
<td>Stability</td>
<td>Concentration (nM)</td>
<td>Test Method</td>
<td>Toxicity</td>
<td>Notes</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>-----------------</td>
<td>-----------</td>
<td>--------------------</td>
<td>------------------------------</td>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Netropsin</td>
<td>AATT</td>
<td>Non-covalent</td>
<td>Reversible</td>
<td>3000 nM</td>
<td>In vitro testing for malaria</td>
<td>High toxicity</td>
<td>[32]</td>
</tr>
<tr>
<td>DAPI</td>
<td>AATT</td>
<td>Non-covalent</td>
<td>Reversible</td>
<td>9.7 nM</td>
<td>In vitro testing for malaria</td>
<td>High toxicity</td>
<td>[32]</td>
</tr>
<tr>
<td>Hoechst 33258</td>
<td>AATT</td>
<td>Non-covalent</td>
<td>Reversible</td>
<td>35.6 nM</td>
<td>In vitro testing for malaria</td>
<td>High toxicity</td>
<td>[32]</td>
</tr>
</tbody>
</table>

A “Pu” denotes a purine, “N” denotes any nucleotide.

B All compounds bind preferentially in the minor groove of DNA.

C The mode of action of these diamidine compounds still remains to be elucidated, but it has been suggested that the intracellular targets include mitochondrial respiration, hemoglobin degradation and DNA replication.
**Bibliography**

Papers of special note have been highlighted as either of interest (*) or of considerable interest (**) to readers.


**This review provides a detailed overview of the antimalarial drug discovery and development process.**


* This study showed that a single treatment with adozelesin protects mice against infection with *P. chabaudi adami*.


** First AT-specific DNA binding agent to enter human clinical trials for treatment of malaria.


*This study was the first to test AT-specific agents against *P. falciparum* in vitro. Although none of these agents progressed past the pre-clinical stage, it validated the
original hypothesis that targeting AT sequences within the *Plasmodium* genome was an effective strategy to identify new anti-malarial compounds.


**Websites**


103. [http://www.sanger.ac.uk/pathogens](http://www.sanger.ac.uk/pathogens)