

## Chapter 1

# Malaria, Drug Discovery and Polyamines

Malaria is an infectious disease that affects the daily lives of more than 2 billion people worldwide and has been estimated to result in 300-500 million clinical cases annually leading to approximately 2 million deaths (Cambell *et al.*, 2004; Guerra *et al.*, 2008). These deaths are mainly due to the most virulent malaria species, *Plasmodium falciparum* (Cambell *et al.*, 2004; Snow *et al.*, 2005). In a recent study performed by Snow *et al.* (2005) an underestimation of the malaria burden has been reported showing the global distribution of clinical episodes of only *P. falciparum* malaria to be responsible for 515 (range 300-660) million cases in 87 malaria endemic countries. It was also estimated that 1 million people die in Africa from infection by *P. falciparum* alone. The incidences of *P. falciparum* was found to be 50% higher globally than what was estimated by the World Health Organization (WHO; Snow *et al.* 2005). The reasons for these conflicting results were ascribed to the WHO's policy of relying on passive national reporting from endemic countries (Snow *et al.*, 2005).

Population groups most at risk to malaria are pregnant women, who lose their acquired immunity and children (Barnes *et al.*, 2008). Malaria is the biggest killer of children of any infectious disease in the world at present and unfortunately the mortality from malaria appears to be increasing (Olliaro, 2005). Malaria poses a great socio-economic burden on endemic countries and it is estimated to cost these already poverty stricken countries nearly US\$ 12 billion annually (Cambell *et al.*, 2004). The lack of a vaccine and the rapid emergence and spread of drug resistant strains of *P. falciparum* therefore necessitate the

development of new drugs and the identification and validation of new parasite-specific therapeutic targets (Hyde, 2005).

## 1.1. Malaria

Human malaria is caused by a protozoan parasite from the phylum *Apicomplexa*, order *Haemosporidia*, family *Plasmodiidae* and genus *Plasmodium*. Four species from the *Plasmodium* genus are able to cause human malaria and include *P. falciparum*, *P. malariae*, *P. ovale* and *P. vivax*. Malaria is a vector-borne parasitic infection which is initiated by the injection of sporozoites into a human by a female Anopheline mosquito (genus *Anopheles*). This starts the life-cycle of *Plasmodium*, which can be divided into three stages schizogony, gametogony and sporogony (Figure 1.1).

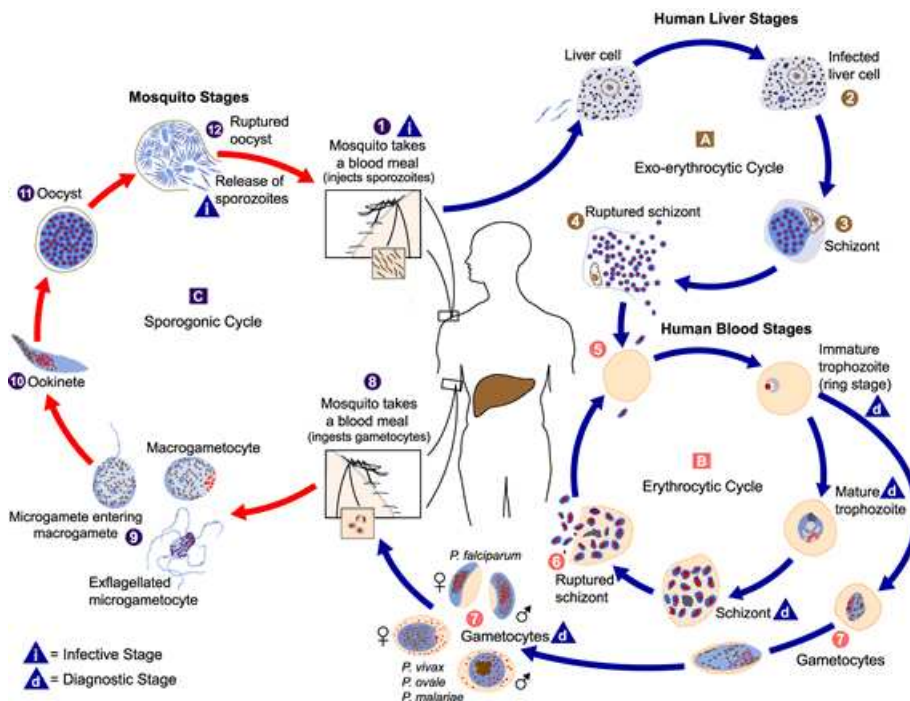


Figure 1.1: The life-cycle of the malaria parasite. The malaria parasite life-cycle involves two hosts, humans and *Anopheles* mosquitoes. **A** represents the exo-erythrocytic schizogony stage. **B** erythrocytic schizogony stage. **C** sporogonic cycle. 1) Injection of sporozoites 2) Infected liver cells 3) Schizonts 4) Ruptured schizonts 5) Ring stage trophozoites 6) Ruptured schizonts 7) Gametocytes 8) Ingested gametocytes 9) Microgametocytes 10) Ookinetes 11) Oocysts 12) Release of new sporozoites. (National Center for Infectious Disease, 2004).

Schizogony is the asexual stage in the host cells where the sporozoites form new individuals. The sporozoites leave the circulation by localizing themselves in the hepatocytes where they transform, multiply and develop into tissue schizonts. This is also known as the pre-erythrocytic stage of infection and is asymptomatic, lasting between 5 and 15 days depending on the *Plasmodium* species (Tracy and Webster, 2001). The rupturing of the tissue schizonts releases thousands of merozoites into circulation initiating the erythrocytic stage of infection. Both *P. falciparum* and *P. malariae* release all merozoites from the liver, however some of the *P. ovale* and *P. vivax* tissue parasites remain in the liver which can cause relapses of erythrocytic infection months to years after the initial infection (Tracy and Webster, 2001). After the invasion of the erythrocytes, asexual development takes place where the merozoites develop to ring forms, which develops into trophozoites and later mature to form schizonts. Rupturing of schizont-containing erythrocytes releases between 6 and 24 merozoites depending on the *Plasmodium* species. The released merozoites invade more erythrocytes to continue the cycle (Figure 1.1).

Malaria can either be “tertian” or “quartan” depending on the periodicity of the rupturing of schizont-containing erythrocytes. Tertian malaria causes febrile attacks to occur on day 1 and 3 due to synchronous ruptures of infected erythrocytes resulting in the release of merozoites into the host’s circulation. Thus, the periodicity of the parasites and febrile clinical manifestation depends on the timing of schizogony generation within the erythrocytes. Tertian malarial infections are caused by *P. falciparum*, *P. ovale* and *P. vivax* where the schizogony generation takes about 48 hours to complete, resulting in febrile attacks. In quartan malaria, caused by *P. malariae*, this process takes 72 hours resulting in malarial attacks on days 1 and 4. Infection by the different human malarial species results in both characteristic illnesses and blood morphological features. The most virulent strain of these species, *P. falciparum* results in an overwhelming parasitemia, sequestration of infected erythrocytes in the peripheral microvasculature, hypoglycemia, hemolysis and shock with multi-organ failure (Tracy and Webster, 2001). The tertian malaria caused by both *P. vivax* and *P. ovale* show milder conditions than *P. falciparum* with low mortality rates in untreated patients. The quartan malaria caused by *P. malar-*

*iae* is rare with a high survival rate.

The intra-erythrocytic ring stage can also undergo gametocytogenesis to start the sexual stage of the malarial life-cycle. During a blood meal on an infected patient, gametocytes are taken up by a female Anopheline mosquito. The male gametocytes undergo exflagellation and gametogenesis which is followed by the fertilization of the female gametocytes resulting in zygotes within the insect's gut. These zygotes develop into oocytes and give rise to sporozoites within the gut-wall of the vector. From the gut-wall the sporozoites invade the salivary glands which then get injected into the host during a blood meal (Tracy and Webster, 2001).

### 1.1.1. Antimalarials

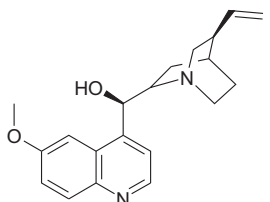
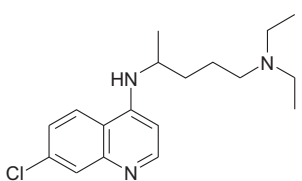
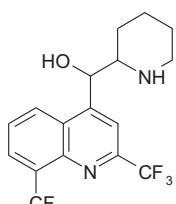
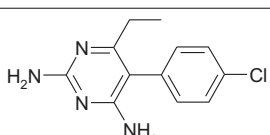
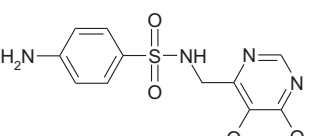
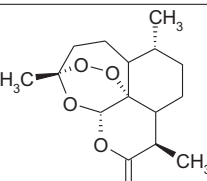
Due to the emergence of drug resistant parasites, insecticide-resistant mosquitoes and the return of malaria to regions where it has previously disappeared from, there is an urgent need for new antimalarials (Figure 1.2; Snow *et al.* 2005). Currently available anti-malarials include quinine, chloroquine and its congeners, primaquine, sulfadoxine-pyrimethamine, mefloquine, atovaquone-proguanil and artemisinin (Nwaka and Hudson, 2006). The most widely used antimalarials in the control of malaria are quinolines (the mainstay of malaria control), antifolates (traditional second-line agents) and more recently artemisinin compounds (Table 1.1; Olliaro 2005).

#### 1.1.1.1. Quinolines

Quinine is an alkaloid found in the bark of a South American chinchona tree. The first resistance to quinine was reported 100 years ago in Brazil although widespread resistance only started during the 1960's (Figure 1.2; Hyde 2005). Quinine was reserved after the emergence of resistance and is currently only used in the treatment of severe malaria of both chloroquine- and multi-drug resistant (MDR) strain infections of *P. falciparum* due to its undesirable side effects (Tracy and Webster, 2001; Tripathi *et al.*, 2005).

Chloroquine forms part of a large series of 4-aminoquinolines which were discovered simultaneously by both the United States and German Armies during World War II

Table 1.1: The three current most widely used drug classes in the control of malaria (Olliario, 2005).

Antimalarial	Structure
<b>Quinoline</b>	
Quinine	
Chloroquine	
Mefloquine	
<b>Antifolates</b>	
Pyrimethamine	
Sulfadoxine	
<b>Artemisinin</b>	
Artemisinin	

(Tracy and Webster, 2001). For 40 years after its discovery chloroquine was the most effective and widely used antimalarial. Chloroquine is highly effective against the erythrocytic forms of *P. vivax*, *P. ovale*, *P. malariae* and the chloroquine-sensitive strain of *P. falciparum* (Tracy and Webster, 2001; Randrianariveojosia *et al.*, 2008). It exerts

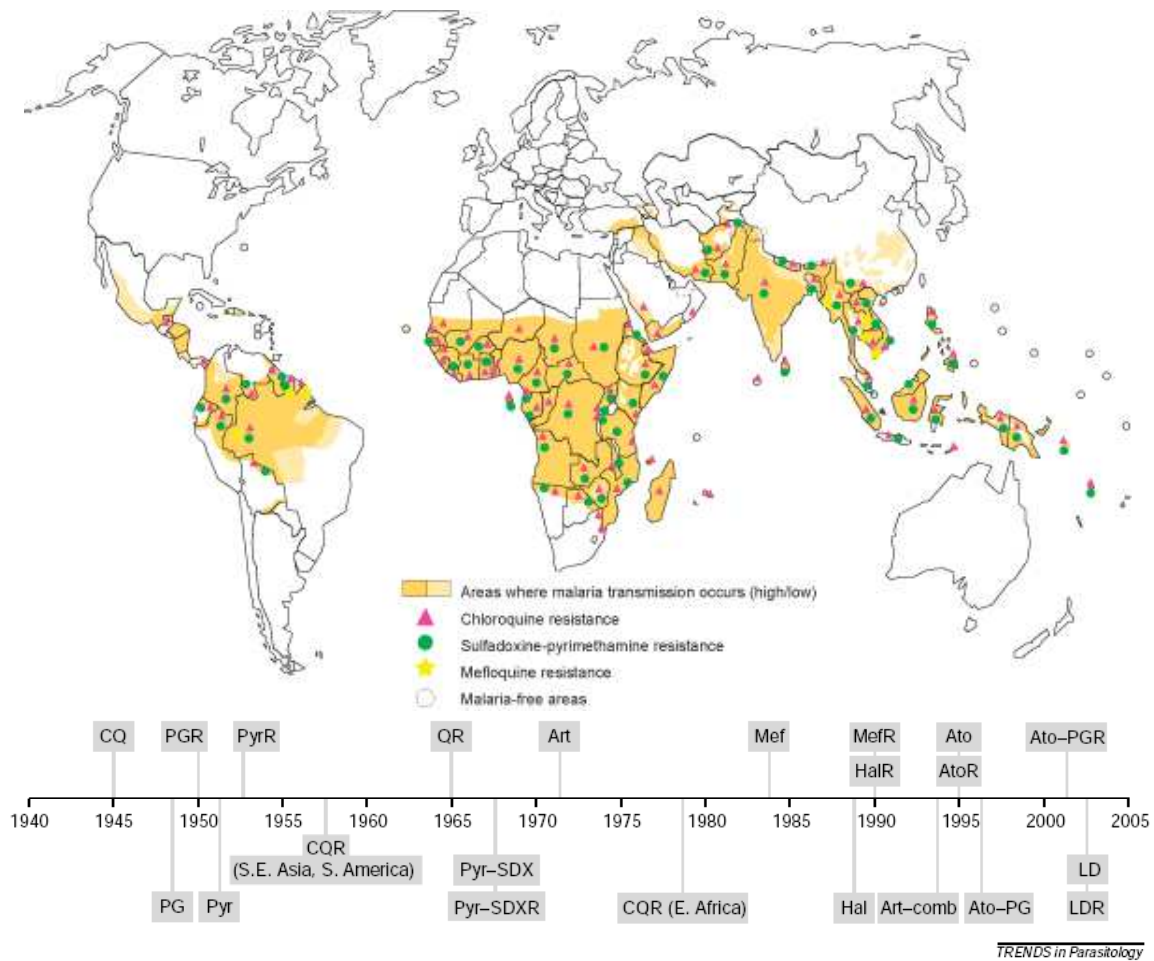


Figure 1.2: (Top) Drug resistance of antimalarials throughout the world in 2005 (<http://www.rollbackmalaria.org>). (Bottom) Approximate dates of introduction of antimalarial drugs as clinical agents, and observations of treatment failure. Q, Quinine; R, resistance; CQ, chloroquine; Pyr, pyrimethamine; SDX, sulfadoxine; Ato, Atovaquone; PG, proguanil; Art, Artemisinin; comb, combinations; Hal, halofantrine; LD, LapDap. Adapted from Hyde (2005).

activity against the gametocytes of *P. vivax*, *P. ovale* and *P. malariae* but not that of *P. falciparum*. Chloroquine is also not effective against the latent stages of recurrent malaria caused by *P. ovale* and *P. vivax* (Tracy and Webster, 2001). Malaria presented resistance to chloroquine the first time in the late 1950's in Southeast Asia and South America and in the late 1970's in Africa (Figure 1.2; Hyde 2005). The resistance to chloroquine is multigenic and epistatic and affects drug accumulation in the food vacuole in the parasite (Olliaro, 2005).

Mefloquine is similar in structure to quinine and was developed in the 1960's in the

Walter Reed Army Institute of Research (WRAIR; Tracy and Webster 2001). Resistance to mefloquine was first reported in the 1980's (Figure 1.2; Hyde 2005). It is postulated that this resistance may have been influenced by prior use of drugs chemically related to quinine. Mefloquine is most effectively used today in the treatment of malaria in combination with artemisinin (Henry *et al.*, 2008b; Agomo *et al.*, 2008; Tangpukdee *et al.*, 2008).

#### 1.1.1.2. Antifolates

Pyrimethamine, a diaminopyrimidine, has the ability to antagonize folic and folinic acids and was found to be the most effective diaminopyrimidine against *P. falciparum*. As chemotherapeutic agent, pyrimethamine has been used in combination with the antibacterial agent sulfadoxine. The sulfadoxine-pyrimethamine combination known as Fansidar has been used for prophylaxis and suppression of human malaria, especially against chloroquine-resistant strains of *P. falciparum* (Tracy and Webster, 2001; Tjitra *et al.*, 2008). Sulfadoxine inhibits dihydropteroate synthase (DHPS) and pyrimethamine, dihydrofolate reductase (DHFR; Oliaro 2005). Drug resistance against sulfadoxine-pyrimethamine developed very rapidly and was present already in the late 1960's (Figure 1.2; Hyde 2005; Fernandes *et al.* 2007). Currently sulfadoxine-pyrimethamine are mostly used as a prophylaxis or in combination with artemisinin derivatives (Rulisa *et al.*, 2007).

#### 1.1.1.3. Artemisinin compounds

Artemisinin is a sesquiterpene lactone endoperoxide derived from the weed, *qing hao* (*Artemisia annua*). Tea made from *qing hao* has been used for more than 2 000 years to treat fevers and in 1596, Li Shizhen recommended its use to relieve the symptoms of malaria for the first time (Klayman, 1985). In 1972 Chinese scientists identified artemisinin to be the active compound in *qing hao*. Three artemisinin derivatives, dihydroartemisinin, artemether and artesunate were later synthesized with greater anti-malarial potency. The potency of these drugs showed 10-100 fold increase to that of other antimalarials available (Tracy and Webster, 2001). Artemisinin is fast acting, effective and safe for use against severe malaria. It is suggested that it should not be used as a prophylactic or against mild attacks of malaria due to its short half-life in humans

(White, 1999; Greenwood *et al.*, 2008). Therefore artemisinin is associated with a high rate of recrudescence after monotherapy and is used in a combination with longer-acting antimalarials such as mefloquine (White, 1999; Greenwood *et al.*, 2008). The use of artemisinin combination treatment (ACT) has become common throughout the world today (Duffy and Mutabingwa, 2004; Yeung *et al.*, 2004). In South Africa lumefantrine and coartem are currently used as a combination therapy (artemether-lumefantrine). The increased popularity of ACT can be ascribed to the distinct modes of action between the use of atemisinins and partner drugs, which in theory should enable ACTs to kill parasites with a decreased susceptibility to one agent (Greenwood *et al.*, 2008). Recently, resistance to artemisinin derivatives along the Thai-Cambodian border have been reported (Wkly Epidemiol Rec. (2007)). The south-eastern border of Thailand with Cambodia has for decades been an epicenter of drug-resistant malaria with the resistance to chloroquine starting here in the late 1950's, followed by resistance to sulfadoxine-pyrimethamine and finally to mefloquine (Figure 1.2); Wkly Epidemiol Rec. (2007)). This highlights the urgent need to develop antimalarials targeting alternative targets to conventional anti-malarials.

#### 1.1.1.4. Multidrug Resistance

Above and beyond the development of drug resistance against the above-mentioned antimalarials used in the control of this disease, multi-drug resistance against these anti-malarials is developing throughout the world. *P. falciparum* can be defined as multi-drug resistant if it shows resistance to more than two groups of operational antimalarial compounds from different chemical classes and has 25% treatment failure (Wernsdorfer and Payne, 1991). Numerous reports of multi-drug resistant malaria have been published over the last couple of years and include countries from the following regions; South-east Asia, Africa and South America (Poespoprodjo *et al.*, 2008; Fanello *et al.*, 2007; Karyana *et al.*, 2008; Tjitra *et al.*, 2008; Tasanor *et al.*, 2006; Clark *et al.*, 2004; Krudsood *et al.*, 2005; Schwab *et al.*, 2005; Gardella *et al.*, 2008). Multi-drug resistance are mainly against the most commonly used classes of 4-aminoquinolines (chloroquine) and antifolates (sulfadoxine-pyrimethamine; Wongsrichanalai *et al.* 2001; Henry *et al.* 2008b). As a result artemisinin combination therapies as first-line treatment have been adapted



almost worldwide after recommendation that combination therapy should be standard for malaria treatment by the WHO and the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) in 2003 (Sibley *et al.*, 2008). Drug resistance in malaria is receiving much attention and a World Antimalarial Resistance Network (WARN) database has been proposed and is currently being constructed (Barnes *et al.*, 2007; Price *et al.*, 2007). This project aims to incorporate clinical data for malaria throughout the world allowing for the identification of accurate and timely recognition of trends in drug efficacy (Price *et al.*, 2007). This will allow for the appropriate intervention and dealing with established multi-drug resistant strains of malaria and also facilitate prompt action against the emergence of new *Plasmodium* resistant strains (Price *et al.*, 2007). It can therefore be concluded that it is pivotal to continue the development of new antimalarials so that a replacement will be ready if and when ACT's reach the end of their clinical life (Greenwood *et al.*, 2008).

### 1.1.2. The Fight Against Malaria

#### 1.1.2.1. Vaccines and Vector Control

Vaccines can be developed against various stages of malaria infection including the pre-erythrocytic, erythrocytic and sexual stages of the parasite (Greenwood *et al.*, 2008). Vaccines against the pre-erythrocytic stages aim to prevent infection and have shown modest results with the best vaccine being RTS,S/AS02A (Aponte *et al.*, 2007; Alonso *et al.*, 2005). The development of vaccines against the asexual or erythrocytic stage of the parasite is preventative in nature and does not target the blocking of invasion of the parasite (Greenwood *et al.*, 2008). Little success has been encountered in finding a vaccine against this stage. Vaccination against the asexual stage of the parasite is aimed to prevent parasite transmission and has been successful in animal models (de Souza *et al.*, 2002). Vaccination against this stage will not have a direct benefit for the patient but will help reduce the transmission of parasites to the community (Greenwood *et al.*, 2008).

The epidemiology of infection and thus the malaria disease is highly dependent on the seasonality, abundance and feeding habits of the vector mosquito *Anopheles* (Snow *et al.*,

2005). For this reason vector control has been an integral part of previous successful malaria eradication programs implemented in Europe and the USA during the last century (Greenwood *et al.*, 2008). Contemporary vector control strategies include the use of insecticide-treated bed nets (ITN) and indoor residual spraying (IRS; Greenwood *et al.* 2008). The use of dichloro-diphenyl-trichloroethane (DDT) as IRS in combination with artemisinin in KwaZulu-Natal, South Africa has been shown to be highly successful in reducing the malarial infections, highlighting the importance of vector control in the fight against malaria (Barnes *et al.*, 2005; Coleman *et al.*, 2008). One of the major concerns in vector control at present is the resistance to the insecticide, pyrethroid, used to treat bed nets (Casimiro *et al.*, 2006). Pyrethroid is the only licensed insecticide to be used in the treatment of ITN's and targets the same biological target as DDT (N'Guessan *et al.*, 2007; Greenwood *et al.*, 2008). Therefore, there is a great need for new insecticides targeting a different biological target to that of DDT and pyrethroid.

### 1.1.2.2. Malaria and Drug Discovery

The global call to halve the world's malaria burden by 2015 has seen a significant increase in the international financing for malarial control (Snow *et al.*, 2008). As a result, many funding agencies have been established including the WHO (<http://www.who.int/malaria/>), the Special Program for Research and Training in Tropical Diseases (TDR; <http://www.who.int/tdr/svc/diseases/malaria>), the Bill & Melinda Gates Foundation (<http://www.gatesfoundation.org>), the GFATM (<http://www.theglobalfund.org/>) and the Medicines for Malaria Venture (MMV; <http://www.mmv.org>). These organizations have developed new collaborative mindsets to scale up drug discovery for malaria as well as other tropical diseases (Hopkins *et al.*, 2007). For example, the MMV (sponsored by philanthropic, public and private sector partners) provide both funding and pharmaceutical-styled research and development management for private-public partnerships with its main aim being the development of drugs with suitable product profiles (Biagini *et al.*, 2005). The MMV currently has three drug candidates in the pre-clinical stage, one in both Phase I and Phase II and two in Phase III trials (<http://www.mmv.org>). The MMV has also succeeded in the registration of Coartem (approved in December

2007; <http://www.mmv.org>). These results highlight the success of these organizations in fighting malaria worldwide.

The development of antimalarials needs to comply to normal product profiles which apply in drug discovery (Rosenthal, 2003). Therefore, it needs to be efficient, safe and has additional properties important for specific disease indications. The MMV (<http://www.mmv.org>) has also proposed additional criteria for the development of antimalarials against uncomplicated malaria; the antimalarials needs to be efficient against drug resistant strains; it should cure within 3 days making use of a single dose per day strategy; it should have low toxicity especially in children and in pregnancy; the drug should have a low risk of developing resistance; it should be adaptable in formulation and packaging with a low manufacturing cost (Biagini *et al.* 2005; <http://www.mmv.org>). The development of antimalarials adhering to such criteria is challenging and therefore it is estimated that only between 1-2% of projects would make it to the later stage of clinical development (Biagini *et al.*, 2005). The cost involved in developing a drug candidate which can enter clinical development has been estimated by the TDR to be in the order of US\$ 20 million (assuming the historical 16% success rate; Kola and Landis 2004; Hopkins *et al.* 2007). It is thus not surprising that most of the drugs developed over the last decade were derivatives from either quinolines, antifolates or artemisinin compounds. The development of resistance against these antimalarial drugs currently used emphasizes the desperate need for new antimalarials targeting alternative drug targets. Therefore, new antimalarials interfering with novel drug targets have been proposed and are currently being investigated (Yeh and Altman, 2006). Yeh and Altman (2006) listed more than 60 drug targets which were classified by their broad biological function and included the following energy metabolism, cofactor and prosthetic group synthesis, protein modification, glycosylphosphatidylinositol biosynthesis, lipid metabolism, DNA replication and transcription, hemoglobin digestion, antioxidant defense and miscellaneous targets (Yeh and Altman, 2006).

The call to halve the malaria burden by 2015 and the subsequent increase in funding

for malarial drug discovery has seen many reviews on how to improve the drug discovery process for malaria and other parasitic protozoa diseases (Rosenthal, 2003; Fidock *et al.*, 2004; Nwaka and Hudson, 2006; Caffrey and Steverding, 2008). The development of new antimalarial chemotherapy has been divided into six main approaches by Rosenthal (2003) and include; 1) optimization of therapy with existing agents; 2) development of analogs of existing agents; 3) natural products; 4) compounds active against other diseases; 5) drug resistance reversers and 6) compounds active against new targets.

Optimization of therapy with existing agents involves the optimization of dosing regimens or formulation. This approach has attracted much attention lately due to the WHO/TDR and GFATM recommendation in 2002 that ACT's be used worldwide as first-line therapy for antimalarials (Sibley *et al.*, 2008). A consideration that needs to be taken into account during this approach is that combination therapies should increase the efficacy of the antimalarials by providing additive or synergistic anti-parasitic activity (Rosenthal, 2003). The combination therapy should also be formulated as to slow the progression of parasitic resistance and prolong the use of the antimalarial. The pharmacokinetics should also be of such a nature that neither of the combination drugs are present as a single antimalarial in the blood level, thus preventing resistance (Laufer *et al.*, 2007).

The development of analogs of existing agents has been successful with various analogs for quinolines and artemisinin currently in use (Nwaka and Hudson, 2006). Analogues of quinolines include chloroquine, mefloquine, aminoquine tefanoquine, primaquine with more analogues being developed (Rosenthal, 2003; Egan, 2001). Analogues of artemisinin include artesunate, artemether and dihydroartemisinin (Weina, 2008). Analogues of antifolates are also being developed at present (Tarnchompoo *et al.*, 2002). The use of computational methods play an increasing important role as is emphasized in a study by Marrero-Ponce *et al.* (2005).

The use of natural products (extracts) for the treatment for malaria dates back more than 2000 years for artemisinin and to the 1600's for quinine. Natural products are usually

obtained by compound extractions from plants but also recently the focus has been placed on marine natural products (Blunt *et al.*, 2008). A recent success story of this approach is the identification of tazopsine, which was identified from extracts of the *Strychnopsis thouarsii* plant growing in Madagascar rainforest's (Flemming, 2007). Tazopsine shows promising results as a prophylaxis. N-cyclopentyl-tazopsine, a tazopsine derivative has shown full protection from malarial challenge in mouse malaria models (Flemming, 2007).

The use of compounds active against other diseases is also known as piggy-back projects. This approach capitalizes on investments made in other diseases. The use of such an approach makes sense for neglected diseases such as malaria. Folate antagonists, tetracyclines and other antibiotics were developed for their antibiotic properties and were later found to have antimalarial activity (Ellis *et al.*, 2001). The development of antimalarials by interfering with polyamine biosynthesis is another such an approach (Muller *et al.*, 2008). The polyamine pathway has been studied extensively in cancer cells and its interferences in other parasitic protozoa has resulted in the successful treatment of African trypanosomiasis. The polyamine pathway will be discussed in detail in section 1.2.

The use of drug resistance reversers in combination with previously effective drugs is yet another approach which can be followed for the treatment of malaria (Rosenthal, 2003; Egan and Kaschula, 2007). The anti-hypertensive verapamil and the anti-depressant desipramine have been shown to have the effect of resistance reversal (Ryall, 1987; Bitonti *et al.*, 1988; Lehane *et al.*, 2008). Verapamil, desipramine and trifluoperazine have been shown to provide relevant clinical resistance reversal when used in combination (van Schalkwyk *et al.*, 2001). More than 40 compounds have been identified to reverse chloroquine resistance with some proposed to have the potential to be used in clinical applications in the near future (Egan and Kaschula, 2007). A hybrid molecule having both antimalarial and reversal characteristics have recently been reported by Burgess *et al.* (2006) and shows more promise for this approach.

The final approach, is the discovery of novel compounds active against new antimalar-

ial targets (Nwaka and Hudson, 2006). New antimalarial targets include targets within the cytosol, parasite membrane, food vacuole, mitochondrion apicoplast and extracellular targets (Fidock *et al.*, 2004; Yeh and Altman, 2006). The discovery of drugs against new drug targets can either be derived by following a random or rational approach. In the past the discovery of drugs has mainly relied on serendipity for drug discovery which includes strategies such as *in vitro* high-throughput screening (HTS). The random approach has been dominated by HTS over the last decade. The HTS technology is limited to big pharmaceutical companies due to the high cost involved in screening of targets. HTS is also limited by high attrition, with a hit-rate of between 0.01-1% of compounds screened (Neamati and Barchi, 2002). As with most drug discovery tools, many efforts are made to increase the hit rate. The advances in HTS include the use of robotics taking HTS to ultra-HTS and the improvement of data quality obtained after screening. Focused libraries which satisfy properties such as drug-likeness are also currently being screened to increase the hit rates and it is proposed that this will reduce the high rate of failure in later stages of the drug discovery process.

The rational approach is based on the use of 3D protein target structures in the design of inhibitors and is known as structure-based drug design (SBDD). Various methodologies exist for SBDD and includes technologies such as virtual screening (high throughput docking (Shahid *et al.*, 2008) and pharmacophore models (Jacobsson *et al.*, 2008)) and NMR-based techniques (fragment-based drug discovery; (Ciulli *et al.*, 2006; Hajduk *et al.*, 2005; Hajduk, 2006)). SBDD include the disciplines of structural biology and bioinformatics, which has resulted in more than 40 drugs that have entered clinical trails of which seven has been approved and marketed by 2003 (Blundell *et al.*, 2006). These successes highlight the importance of SBDD in the drug discovery process. The advent of the “omics” era has seen a shift to the more rational SBDD approach (Sawyer 2006). The recent advances in the biological field has resulted in vast resources of information that need to be mined as to better understand biological systems and ultimately be used in the drug design process (Sawyer, 2006). The mining of data for drug discovery includes the following disciplines: functional genomics, structural genomics, proteomics,

chemoinformatics and bioinformatics (Birkholtz *et al.*, 2006, 2008a). This is followed by making sense of the data and the incorporation thereof in the drug discovery process, in particular the SBDD process. The SBDD methodology followed in this study is the development of a dynamic receptor-based pharmacophore model, which allows for the incorporation of protein flexibility within the drug design process. Pharmacophores and their importance will be discussed in detail in Chapter 3.

Although the random and rational approaches differ significantly, the techniques used to accomplish the goal of finding new drugs or lead compounds can frequently be used interchangeably. Therefore, the array of both different computational and experimental disciplines available to the drug discovery process should be seen as a multifaceted discipline which can contribute to the early stages of this process, regardless of the approach followed (Stahl *et al.*, 2006). As a result, smart drug discovery platforms have been put in place in most pharmaceutical companies (Sawyer, 2006). Hopkins *et al.* (2007) proposed a smart drug discovery platform for the discovery of new antimalarials (Figure 1.3). This smart drug discovery platform can generally be seen as iterative but having discrete decision stages allowing projects to continue to the next stage of the drug discovery process or to be terminated (Figure 1.3). The first stage in this platform is target discovery and involves the identification of a target by exploiting information derived from biochemical properties and the data mining of the genome sequence of the organism of interest, in this case malaria. The identified protein target should be subjected to target assessment to determine if it is druggable and what the potential is for finding selective inhibitors. This is subsequently followed by target validation (Figure 1.3; decision stage: Target selection). The latter entails confirmation that intervention at the specific step of a biological pathway will have the desired biological effect (Neamati and Barchi, 2002). Drug target validation is still considered as the bottle-neck in drug discovery and there is no standard way to validate targets identified. Some of the widely used techniques currently implemented are targeted gene disruption, gene knockouts and reverse genetics by specific suppression of genes (Neamati and Barchi, 2002).

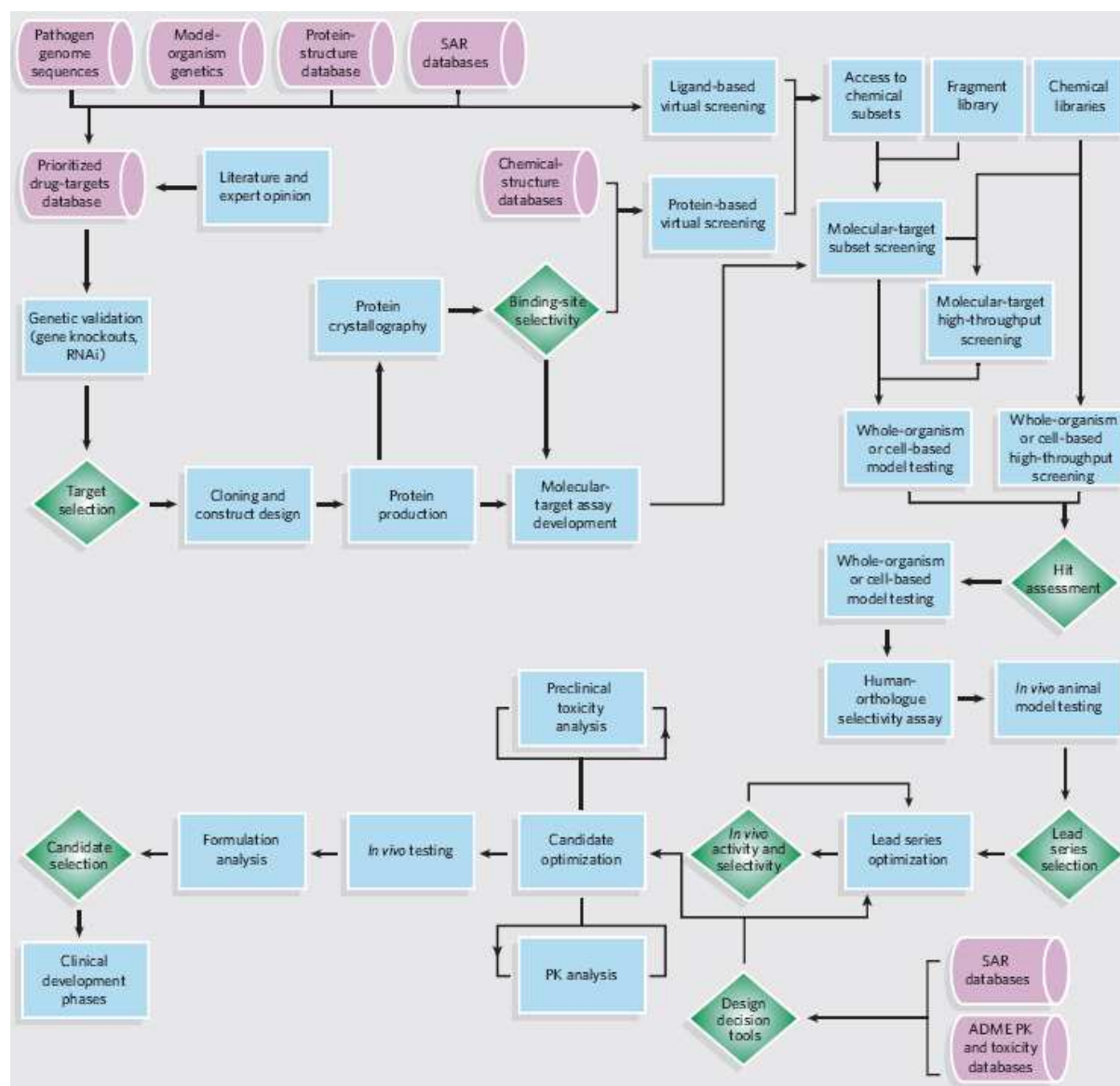


Figure 1.3: Drug Discovery: From the genome to the clinic. Drug discovery consists of discrete decision stages (diamonds), with iterative experiments at each step (boxes). The early stage involves identifying hits by exploiting information (cylinders) from the biochemical properties and genome sequence of an organism in a lead-discovery strategy. Taken from an article by Hopkins *et al.* (2007). PK, pharmacokinetics; RNAi, RNA interference; SAR, structure-activity relationships.

The validated drug target then enters the lead discovery stage which starts by solving the protein structure (Figure 1.3; decision stage: Binding-site selectivity). A compound can be defined as a lead when it demonstrates a desired biological activity at low micromolar concentrations ( $< 10 \mu M$ ) on a validated molecular target (Neamati and Barchi, 2002). The lead discovery stage can take place either *in vitro* and/or *in silico* with a com-



bination of the computational and experimental approaches having the most advantage. First a suitable molecular screening assay needs to be developed and subsequently the target undergoes molecular-screening by using subset (10-1 000 compound), fragment (1 000-2 000 compounds) and HTS (10 000-1 000 000 compounds) libraries (Hopkins *et al.*, 2007). Structure activity relationships (SAR) and quantitative SAR (QSAR) can be performed on known ligands derived from either *in vitro* screening experiments or data mining of literature. In the presence of a 3D structure, virtual screening of selective libraries can be performed and the best scoring compounds subsequently screened *in vitro*. Hits identified from screening are subsequently assessed for their efficacy, selectivity and druggability (Figure 1.3; decision stage: Hit assessment). Successful hit compounds are then clustered together to derive a lead series which enters the lead optimization stage of the drug discovery process (Figure 1.3; decision stage: Lead series selection). Lead optimization is then performed to obtain desirable pharmaceutical properties consisting of absorption, distribution, metabolism, excretion and toxicity (ADMET; Hopkins *et al.* 2007). Before a compound can be selected as a clinical candidate it goes through many iterative processes of decision making based on *in vivo* activity and selectivity as well as prediction from decision making tools which include ADMET and SAR (Hopkins *et al.*, 2007).

It is therefore clear that many approaches to fight malaria exist and that the discovery of antimalarials acting on new drug targets is desperately needed. Spermidine synthase, a protein in the polyamine pathway, is one such a target which has been selected for the discovery of novel antimalarials in this study. In the following section *P. falciparum* spermidine synthase (PfSpdSyn) and polyamine metabolism will be discussed, highlighting its importance to parasites.

## 1.2. Polyamines

Polyamines are ubiquitous aliphatic amines consisting of putrescine (1,4-butanediamine), spermidine (N<sup>1</sup>-(3-aminopropyl)-1,4-butanediamine) and spermine (N<sup>1</sup>,N<sup>4</sup>-bis(3-ami-

nopropyl)-1,4-butanediamine ;Table 1.2).

Table 1.2: The chemical structures of three main polyamines putrescine, spermidine and spermine.

Name	Chemical Structure
Putrescine	$\text{H}_3\text{N}^+ \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{NH}_3^+$
Spermidine	$\text{H}_3\text{N}^+ \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{NH}_2^+ \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{NH}_3^+$
Spermine	$\text{H}_3\text{N}^+ \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{NH}_2^+ \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{NH}_2^+ \text{---} \text{CH}_2 \text{---} \text{CH}_2 \text{---} \text{NH}_3^+$

The biosynthesis of these polyamines peaks during cell proliferation and differentiation (Panagiotidis *et al.*, 1995; Muller *et al.*, 2008). Although the absolute molecular functions of polyamines are still unclear, they are thought to play an important role in the stabilization of DNA and RNA, phospholipids and various proteins *in vivo* (Igarashi *et al.*, 1982; Tabor and Tabor, 1984; Wallace *et al.*, 1995). The stabilization of nucleic acids takes place via the binding of polyamines. Putrescine and spermidine to the minor groove of DNA (Ruiz-Chica *et al.*, 2001) whereas spermine prefers binding in the major groove but has also been shown to bind in the minor groove and the backbone of DNA (Pelta *et al.*, 1996). The binding of polyamines to both naked DNA and chromatin has been shown to cause condensation of DNA by acting as a clamp holding together two different molecules or two different parts of the molecule (Wemmer *et al.*, 1985; Bryson and Greenall, 2000). Both spermidine and spermine have been shown to act as free radical scavengers playing a protective role in cell health (Ha *et al.*, 1998; Khan *et al.*, 1992). Most interesting is that spermidine has been shown to increase DNA-polymerase activity six-fold in *P. falciparum* (Bachrach and Abu-Elheiga, 1990).

### 1.2.1. Polyamine metabolism in mammals and *P. falciparum*

The importance of polyamines are emphasized by the tight regulation of their cellular levels, which is handled in a very fast, precise and sensitive manner (Urdiales *et al.*, 2001). A tight regulation of polyamine pools is needed since it has been shown that

the depletion of polyamines causes cell death and that an excess is usually toxic (Davis, 1990). The tight control over the cellular polyamine pools can be exerted in four ways; *de novo* synthesis, interconversion (Caraglia *et al.*, 2000), terminal degradation (Rabellotti *et al.*, 1998), and transport (Seiler *et al.*, 1996).

The polyamine building blocks used in synthesis are arginine and methionine (Taylor *et al.*, 2008). Arginase (EC 3.5.3.1) catalyzes the hydrolysis of arginine to produce ornithine and urea. In this study emphasis will be placed on the synthesis of putrescine, spermidine and spermine (Figure 1.4). Ornithine gets decarboxylated to form putrescine and is catalyzed by ornithine decarboxylase (ODC; EC4.1.1.17). Putrescine acts as a scaffold accepting the transfer of an aminopropyl group from S-adenosyl-5'-deoxy-(5')-3-methylthiopropylamine (dcAdoMet) to produce spermidine mediated by spermidine synthase (SpdSyn; EC 2.5.1.16). The synthesis of spermine takes place by the addition of an aminopropyl group from dcAdoMet to spermidine catalyzed by spermine synthase (SpmSyn; EC 2.5.1.22). Both SpdSyn and SpmSyn are part of the aminopropyltransferases family of enzymes producing the respective polyamine and 5'-methylthioadenosine (MTA). dcAdoMet is produced by the decarboxylation of S-adenosylmethionine (AdoMet) and is catalyzed by S-adenosylmethionine decarboxylase (AdoMetDC; EC 4.1.1.50). Mammalian cells differ from *P. falciparum* in that an interconversion pathway exists (Figure 1.4). Mammalian cells therefore have the ability to inter-convert between spermine and spermidine which takes place by the acetylation with acetylCoA by means of spermine/spermidine acetyltransferase (cSAT, EC 2.3.1.57). The spermine and spermidine acetyl derivatives undergo oxidative splitting by polyamine oxidase (PAO, EC 1.4.3.4) to produce putrescine and spermidine, respectively, which can later be re-utilized in biosynthesis. The polyamine biosynthetic pathway of *P. falciparum* also differs from the mammalian cells in that no spermine synthase has been identified and that the *P. falciparum* SpdSyn has the ability to produce small quantities of spermine (Haider *et al.* 2005; Figure 1.4). A third difference between mammalian and *P. falciparum* cells is that AdoMetDC/ODC occurs as a bifunctional enzyme (Muller *et al.*, 2001). These differences raise hope for finding

possible drugs against parasitic protozoa (Muller *et al.*, 2001).

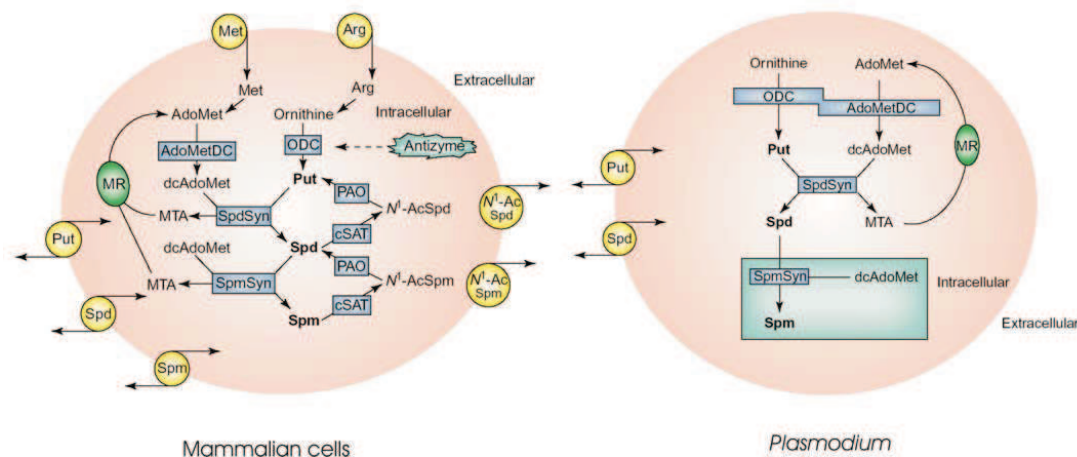


Figure 1.4: The polyamine pathway in mammals and *P. falciparum*. Abbreviations: AdoMet, S-adenosylmethionine; AdoMetDC, S-adenosylmethionine decarboxylase; cSAT, cytosolic N<sup>1</sup>-acetyltransferase specific for spermidine and spermine; dcAdoMet, decarboxylated S-adenosylmethionine; MR, Met recycling pathway; MTA, methylthioadenosine; N<sup>1</sup>-AcSpd, N<sup>1</sup>-acetyl spermidine; N<sup>1</sup>-AcSpm, N<sup>1</sup>-acetyl spermine; ODC, ornithine decarboxylase; PAO, polyamine oxidase; Put, putrescine; Spd, spermidine; SpdSyn, spermidine synthase; Spm, spermine; SpmSyn, spermine synthase. The polyamine pathway of *P. falciparum* does not have the capability of converting spermidine nor spermine to putrescine and spermidine, respectively. In the polyamine pathway of *P. falciparum* no spermine synthase has been identified. Adapted from Muller *et al.* 2001.

Polyamine metabolism has been extensively studied as a potential human cancer target due to its antiproliferative potential but, with moderate success since its inhibition generally results in cytostasis (Pegg *et al.*, 1995; Muller *et al.*, 2008). However, the use of polyamine enzyme inhibitors as chemopreventative agents has been shown to provide benefit in multiple cancer trials (Muller *et al.*, 2008). There is growing evidence that polyamines appears to play a role in apoptosis, however the relationship appear to be complex and is not fully understood (Flamigni *et al.*, 2007). Due to the cytostatic effect of polyamine inhibition it has been proposed to have potential in the treatment of cardiovascular disease, stem cell transplantation, arthritis and infections (Flamigni *et al.*, 2007). The lack of success of polyamine biosynthetic inhibitors can in most cases be explained by the supplementation of the depleted polyamine pools by transport from the extracellular environment (Seiler *et al.*, 1996; Reguera *et al.*, 2005). Polyamines

are usually taken up from polyamine rich erythrocytes via a single transporter able to transport putrescine, spermidine and spermine into the parasite (Gugliucci, 2004). It has been suggested that for a polyamine biosynthetic drugs to be successful in mammalian cells, effective polyamine transporter drugs are also needed (Seiler *et al.*, 1996).

### 1.2.2. Polyamine Metabolism Inhibitors and Parasitic Protozoa

The polyamine biosynthetic pathway and its control has widely been studied in a variety of parasitic protozoa including *Trypanosoma brucei*, *T. cruzi*, *Leishmania donovani*, *L. infantum*, *P. berghei* and *P. falciparum* (Heby *et al.*, 2007; Gonzalez *et al.*, 2001; Kaiser *et al.*, 2003; Haider *et al.*, 2005).

ODC is the most widely studied enzyme in the polyamine biosynthetic pathway. It is also the first enzyme in this pathway and its inhibition is expected to deplete cells of the three main polyamines putrescine, spermidine and spermine. D,L- $\alpha$ -difluoromethylornithine (DFMO) is an ODC inhibitor and is the most successful polyamine inhibitor to date. DFMO was originally designed to be an antitumor drug and currently is used in combination with other antitumor drugs and is being investigated as a prophylactic in cancer treatment (Muller *et al.*, 2008). DFMO is currently being used as an effective treatment for both the early and late stages of West-African trypanosomiasis also known as African sleeping sickness caused by *T. brucei gambiense* (van Nieuwenhove *et al.*, 1985; Heby *et al.*, 2003). The success of treatment with DFMO can be explained by the long half-life for ODC from *T. b. gambiense* of  $\sim 6$  hours, compared to less than 1 hour for the human ODC (Phillips and Wang, 1987). Implied in the stability of this enzyme is that the inhibited ODC of *T. b. gambiense* does not get replaced rapidly enough by newly synthesized active ODC to rescue the parasite. DFMO results in the depletion of spermidine, which prevents the parasite to synthesize trypanothione, a conjugate of spermidine and glutathione that is unique to trypanosomas and *Leishmanias*. DFMO is ineffective in the treatment of East African sleeping sickness caused by *Trypanosoma brucei rhodesiense*, which can not yet be explained. The treatment of *T. brucei* with DFMO is currently a success story for the treatment of parasitic protozoa by interfering

with the polyamine pathway and gives hope for the treatment of other parasitic protozoa.

The inhibition with an irreversible inhibitor of AdoMetDC 5'-{[(Z)-4-amino-2-butenyl]-methylamino}-5'deoxyadenosine (MDL73811; AbeAdo) has been shown to be effective against *T. b. brucei* infections in mice and rats (Byers *et al.*, 1991). Testing of this drug together with DFMO showed synergistic effects curing clinical isolates of *T. b. rhodesiense*, suggesting that inhibition of both enzymes may be used in chemotherapy against East African sleeping sickness (Heby *et al.*, 2003).

*T. cruzi*, the causative agent of Chaga's disease (American trypanosomiasis) lacks the ODC enzyme and relies on the uptake of exogenous putrescine for its survival (Persson *et al.*, 1998). Consequently, the treatment of this disease requires the development of drugs targeting enzymes downstream of ODC.

*P. falciparum* has been shown to be sensitive to the inhibition of ODC at several stages of its life-cycle. DFMO and structurally related compounds have been tested on malaria parasites to validate polyamine synthesis for use in chemotherapeutic intervention with results varying from modest to promising (Muller *et al.*, 2008). The effect of DFMO has been tested on the *P. falciparum* cultured erythrocytic stage, *Anopheles* and rodent model infections of *P. berghei* (Muller *et al.*, 2008). DFMO has been shown to block the development of trophozoites in erythrocytic schizogony and to have a cytostatic rather than a cytotoxic effect (Assaraf *et al.*, 1987). The cytostatic effect of DFMO inhibition could be explained by the depletion of polyamines. The supplementation of exogenous polyamines is able to compensate for the growth inhibition of *P. falciparum* cultures with exogenous putrescine and spermidine resulting in full recovery of parasites treated with DFMO. The effects of exogenous supplementation of spermine is still debated. Bitonti *et al.* (1987) showed that polyamine-depleted parasites can be recovered, however Wright *et al.* (1991) showed that they can not be recovered. The treatment of parasites with MDL 73811 an irreversible inhibitor of AdoMetDC showed an increase in putrescine levels (three-fold) and a decrease of spermidine levels by 67%, however the treated parasite could

not significantly be rescued by the exogenous spermidine (Wright *et al.*, 1991; Gupta *et al.*, 2005). It is therefore clear that interference with the polyamine biosynthetic pathway in *P. falciparum* shows potential for studies aimed at the development of new antimalarials.

### 1.2.3. Spermidine Synthase

Studies directed at polyamine biosynthesis as a drug target in *P. falciparum* have mainly been focused on ODC and AdoMetDC with attention only being paid to spermidine synthase (EC 2.5.1.16) in the last couple of years. SpdSyn is expressed during erythrocytic schizogony with both the mRNA and protein peaking at the mature trophozoite stage (Haider *et al.*, 2005). The latter coincides with the transcriptional rate of the bifunctional ODC/AdoMetDC (Muller *et al.*, 2000). SpdSyn is a cytoplasmic protein which consists of 321 amino acids and has a molecular weight of 36.6 kDa (Haider *et al.*, 2005). The removal of 29 amino acids from the N-terminal is needed for recombinant expression in *Escherichia coli*. The N-terminal extension is of unknown function and is believed to have a signal peptide-like character, also present in plant SpdSyn (Haider *et al.*, 2005). A number of nuclear genes from *P. falciparum* also shows a high homology to genes from plant species and it is suggested that these proteins were originally encoded by the genome of the apicoplast, a secondary endosymbiotic red algae (Foth *et al.*, 2003; Haider *et al.*, 2005).

PfSpdSyn occurs naturally as a dimer and is a member of the putrescine aminopropyltransferase family that generally consists of a small N-terminal domain and a large C-terminal domain (Rossmann-like fold). The N-terminal domain of a SpdSyn monomer consists of a six-stranded  $\beta$ -sheet ( $\beta$ 1-6) and a C-terminal domain of a seven-stranded  $\beta$ -sheet flanked by nine  $\alpha$ -helices. The gate-keeping loop can be found between residues Ser 197 and Glu 205 and is located between strand  $\beta$ 9 and the  $\alpha$ 5 helix.

SpdSyn catalyzes the reaction in which an aminopropyl group from dcAdoMet is added to putrescine to produce spermidine and MTA. Interestingly, PfSpdSyn has the ability to also produce low levels of spermine in the parasite, however no spermine synthase

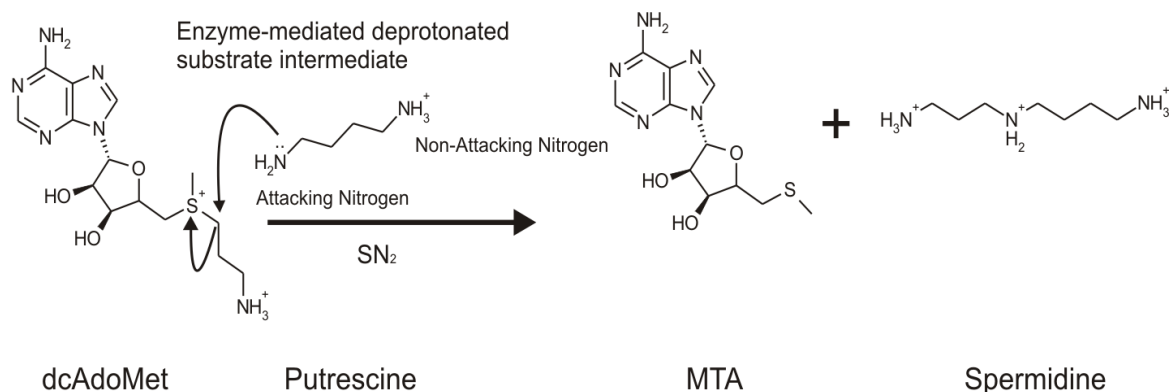


Figure 1.5: The mechanism of action of SpdSyn. Enzyme mediated deprotonation of the attacking nitrogen of putrescine results in a nucleophilic attack on an electrophilic carbon of dcAdoMet. This results in the formation of spermidine and MTA.

has been identified in the parasite (Haider *et al.*, 2005). It has been proposed that the mechanism of action of SpdSyn takes place via a SN<sub>2</sub> reaction resulting in the inverse configuration of the methylene carbon of dcAdoMet undergoing a nucleophilic attack by putrescine and is mediated by a gate-keeping loop (Figure 1.5; Ikeguchi *et al.* 2006). The ability of SpdSyn to produce spermine can be explained by the difference in substrate specificity of these enzymes between species. The human SpdSyn has a higher substrate specificity compared to those of other organisms such as *P. falciparum*, *E. coli* and *T. maritima*, which all produce small amounts of spermine (Korolev *et al.*, 2002; Bowman *et al.*, 1973; Haider *et al.*, 2005). The differences in the binding specificity between species can be explained by the differences in the gate-keeping loop residues, which affects the loop's mobility (flexibility) allowing for the binding of longer polyamines (Dufe *et al.*, 2007).

The essential nature of SpdSyn in the parasite is reflected in the importance of its product, spermidine. Spermidine contributes to the general role of polyamines by stabilizing DNA and RNA and was shown to increase the activity of DNA-polymerase in *P. falciparum* sixfold (Bachrach and Abu-Elheiga, 1990). Most importantly, spermidine is a precursor for the modification and activation of the eukaryotic translation initiation factor 5A (eIF-5A) and in trypanosomes for the biosynthesis of the glutathione mimic, trypanothione (Byers *et al.*, 1992; Muller *et al.*, 2003; Kaiser *et al.*, 2003). Some



effects of polyamine biosynthesis inhibitors have therefore been attributed to the accumulation of unmodified eIF-5A due to spermidine depletion, for example, in the suppression of multi-drug resistant human immunodeficiency virus type 1 (HIV-1) replication (Bachrach and Abu-Elheiga, 1990; Park *et al.*, 1993; Schafer *et al.*, 2006). Null mutants of SpdSyn of the protozoan *L. donovani* have also shown that this enzyme is absolutely essential for the survival of lower eukaryotes (Guo *et al.*, 1999; Roberts *et al.*, 2001; Jin *et al.*, 2002). Recently, SpdSyn was validated genetically to be a drug target in *T. brucei* (Taylor *et al.*, 2008). The inhibition of PfSpdSyn by a known inhibitor, dicyclohexylamine (cyclohexylamine) has been shown to block *P. falciparum* growth by depleting the endogenous polyamine pools (Kaiser *et al.*, 2001). A potent inhibitor of PfSpdSyn *trans*-4-methylcyclohexylamine (4MCHA), causes an 85% growth arrest after 48 hours (Haider *et al.*, 2005). These parasites could not be significantly rescued by the addition of exogenous spermidine (Haider *et al.*, 2005). Interestingly, the spermidine concentration in uninfected erythrocytes has been shown to be significantly lower than erythrocytes infected with the NF54 (chloroquine-susceptible; 10 times lower) or R (chloroquine-resistant; 35 times lower) strains (Kaiser *et al.*, 2001). Inhibition of the parasite cultures with the inhibitors, dicyclohexylamine (SpdSyn inhibitor), 1,7-diaminoheptane (homospermidine synthase inhibitor) and agmatine (ODC inhibitor) led to the depletion of spermidine in the erythrocytes, supporting the importance of spermidine to the parasite (Kaiser *et al.*, 2001). The reversal of inhibition of all three inhibitors was possible after the addition of high concentrations of spermidine (Kaiser *et al.*, 2001). The inhibitory effect of an ODC inhibitor, 3-aminooxy-1-aminopropane (APA) in *P. falciparum in vitro*, could be reversed by the addition of putrescine but not by spermidine. However, APA was also found to be a moderate inhibitor of spermidine synthase (IC<sub>50</sub> of 84 $\mu$ M) and had an IC<sub>50</sub> value in culture of 1 $\mu$ M. It has been suggested that the ability to rescue parasites treated with spermidine inhibitors with putrescine and not spermidine can be attributed to the inefficient uptake of spermidine in parasite infected erythrocytes. This, however, does not exclude the possibility of additional targets for SpdSyn inhibitors (Haider *et al.*, 2005).

The essential nature of the PfSpdSyn to the survival of the parasites warrants the further investigation into the design of novel inhibitors.

### 1.3. Aims

This study addresses the need to identify novel antimalarial drug classes directed at alternative drug targets to those targeted by conventional antimalarial drugs. This objective was pursued by developing a dynamic receptor-based pharmacophore model for PfSpdSyn. The specific aims were as follows:

- **Chapter 2 - Structural and Mechanistic Insights into the Action of *P. falciparum* Spermidine Synthase:** The aim was to gain mechanistic insights into the mode of action of PfSpdSyn and includes the homology modeling of the enzyme, which could be incorporated in the lead discovery process to follow.
- **Chapter 3 - The Development of a Dynamic Receptor-Based Pharmacophore Model of *P. falciparum* Spermidine Synthase:** The aim was to develop a dynamic receptor-based pharmacophore model for PfSpdSyn to identify novel lead compounds. Information gained from the latter methodology was incorporated into a knowledge-based rational drug design strategy to design novel lead compounds. The compounds identified and designed during the exploration of the biological space of PfSpdSyn were tested *in vitro* to validate specific binding characteristics.

Work from this study was published in the Journal of Bioorganic & Medicinal Chemistry entitled: Structural and mechanistic insights into the action of *Plasmodium falciparum* spermidine synthase (Burger *et al.*, 2007). This work was expected to contribute novel hit compounds specific for PfSpdSyn, which may lead to the development of a new drug class of antimalarials.

## Chapter 2

# Work leading up to Structure-Based Drug Design of Spermidine Synthase: Structural and Mechanistic Insights into the Action of *Plasmodium falciparum* Spermidine Synthase

### 2.1. Introduction

Significant efforts have been invested over the last decade to increase the number of available 3D protein structures, including projects such as for example the Protein Structure Initiative (PSI; <http://www.structuralgenomics.org/>) and the Structure Genomics Consortium (SCG; <http://www.sgc.utoronto.ca>). Despite these efforts many therapeutically relevant structures remain unresolved (Nayeem *et al.*, 2006). The time required for solving an average eukaryotic structure is estimated to take between one and three years and includes both protein expression, crystallization and solving the structure. This time is thought to increase significantly with high risk targets such as viruses, molecular machinery and membrane proteins (Stevens, 2004). It is estimated that the use of 3D structures in the early drug discovery process for finding high quality leads could save up to 50% of the total costs (Stevens, 2004). Solving of malaria protein structures is notoriously difficult especially in *P. falciparum* with the main reason being the high A + T composition of 80% in coding regions (Gardner *et al.*, 2002). The A + T composition can rise to as high as 90% in the intron and intergenic regions (Gardner *et al.*, 2002). This A + T bias in the genome composition alters the codon usage and therefore makes it

very difficult to express malarial proteins at sufficient quantities for crystallization in most heterologous systems (Withers-Martinez *et al.*, 1999). This challenge is currently being addressed by codon usage optimization and harmonization. Another factor complicating the expression of malarial proteins is the presence of low complexity segments and/or inserts. This can best be described as unstructured protein regions containing only a subset of the 20 amino acids, sometimes occurring as tandem repeat arrays (Xue and Forsdyke, 2003; Birkholtz *et al.*, 2008b).

Protein structure prediction (i.e. homology modeling) can be seen as an alternative method to get structural information where experimental techniques have failed (Nayeem *et al.*, 2006). Three widely accepted methods exist which can be used to predict 3D protein structures from the amino acid sequence including comparative modeling, threading and *ab initio* prediction (Bourne and Wiessig, 2003). Comparative modeling is the most common method used and utilizes predetermined protein structures to predict the protein conformation of unresolved protein structures with a similar amino acid sequence (Hillisch *et al.*, 2004). The comparative methods are based on the assumption that the 3D structures of proteins are more conserved than their primary structures or amino acid sequences (Chothia and Lesk, 1986). In other words proteins with a certain degree of evolutionary-related sequence identity, have similar structures. Threading is also known as fold recognition and entails that a protein sequence be scored against a 3D database of known protein folds. The best-fit structure is assumed to be the fold adopted by the protein sequence. This can then be used to find the structure of the protein of interest. Bourne and Wiessig (2003) described *ab initio* structure prediction as a mixture between science and engineering with no reliable methods at present. The science part can be described as how the 3D structure of a protein is attained and the engineering part the deduction of a 3D structure from a given sequence. Homology modeling is currently the most reliable and mature of the three methods and hence was used in this study.

The uses of homology models are versatile and models can be used in nearly all stages of the drug design process (Figure 2.1). The quality of models derived from homology

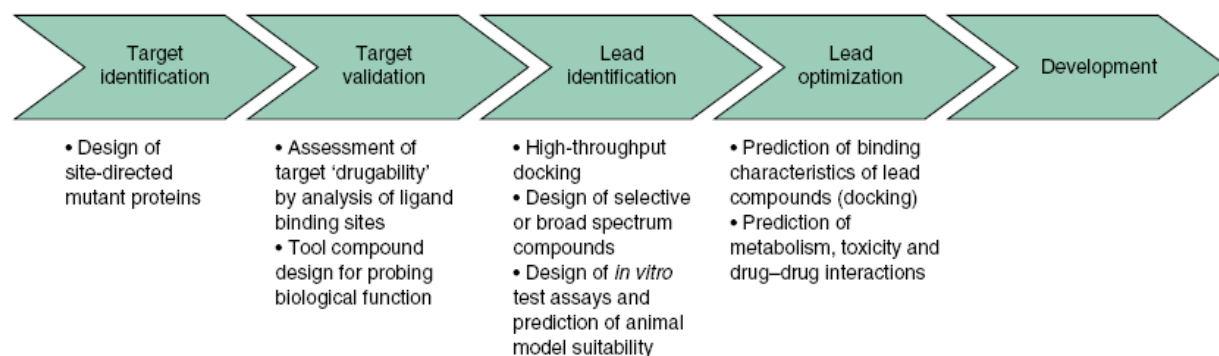


Figure 2.1: Application of homology models in the drug discovery process. The enormous amount of protein structure information currently available could not only support lead compound identification and optimization, but could also contribute to target identification and validation. Adapted from Hillisch *et al.* (2004).

modeling is dependent on the relationship between the protein sequence identities of the known structure and the target protein to be modeled (Chothia and Lesk, 1986). Homologous proteins can be described as proteins having a common ancestor and thus an evolutionary relationship with a high probability of sharing a common 3D structure. Homologous proteins with a sequence identity of 30% and higher can often be used to construct reliable homology models. The relationship between the target and template sequence identity and the information content of the resulting homology models is presented in Figure 2.2. Homology models based on a sequence identity of lower than 15% are thought to be speculative and could lead to misleading results. Models with a sequence identity of between 15 and 30% can be used in functional assignment of the target protein and to direct mutagenesis studies. However, sophisticated profile-based methods should be used for alignment of the target and template sequences. Proteins having between 30 and 50% sequence identity can be used in structure-based methods to predict the target druggability and the design of mutagenesis experiments and *in vitro* test assays. Hillisch *et al.* (2004) proposed that in the presence of a 50% or higher sequence identity between the target and template, the resulting models are frequently of a sufficient quality to be used in the prediction of protein-ligand interactions. This includes structure-based drug design and the prediction of preferred sites of metabolism for small molecules.

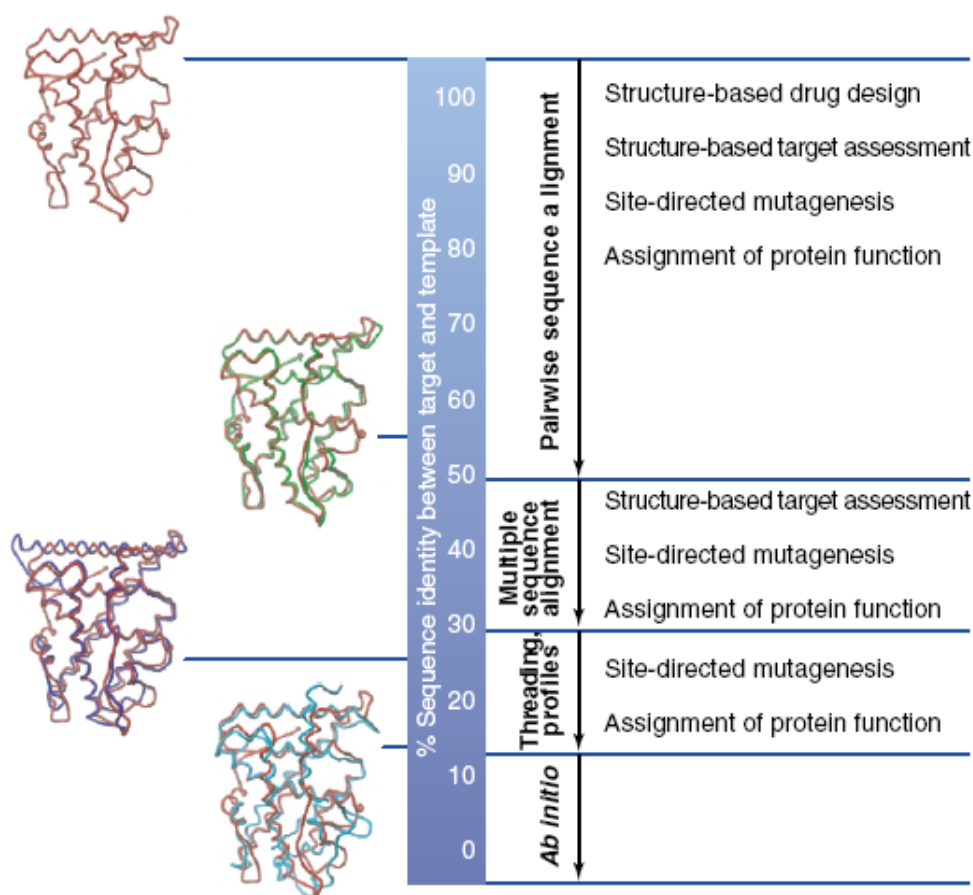


Figure 2.2: The relationship between target and template sequence identity and the information content of resulting homology models. The arrows indicate the methods that can be used to detect the sequence similarity between target and template sequences. Applications of the homology models in drug discovery are listed on the right. The higher the sequence identity, the more accurate the resulting structure information. Homology models that are built on sequence identities of  $>50\%$  can frequently be used for drug design purposes. Adapted from Hillisch *et al.* (2004).

The homology modeling process usually comprises of four steps which include fold assignment, sequence structure alignment, model building and model refinement (Figure 2.3). The first step, fold assignment, entails the template recognition and initial alignment. The second step involves the crucial step of sequence alignment. This is followed by model building which includes backbone generation, loop modeling and side-chain modeling. The final step involves the refinement or optimization of these models.

There are various software solutions which can be used to construct homology models including programs such as WHATIF (Vriend, 1990), MODELLER (Sali and Blundell,

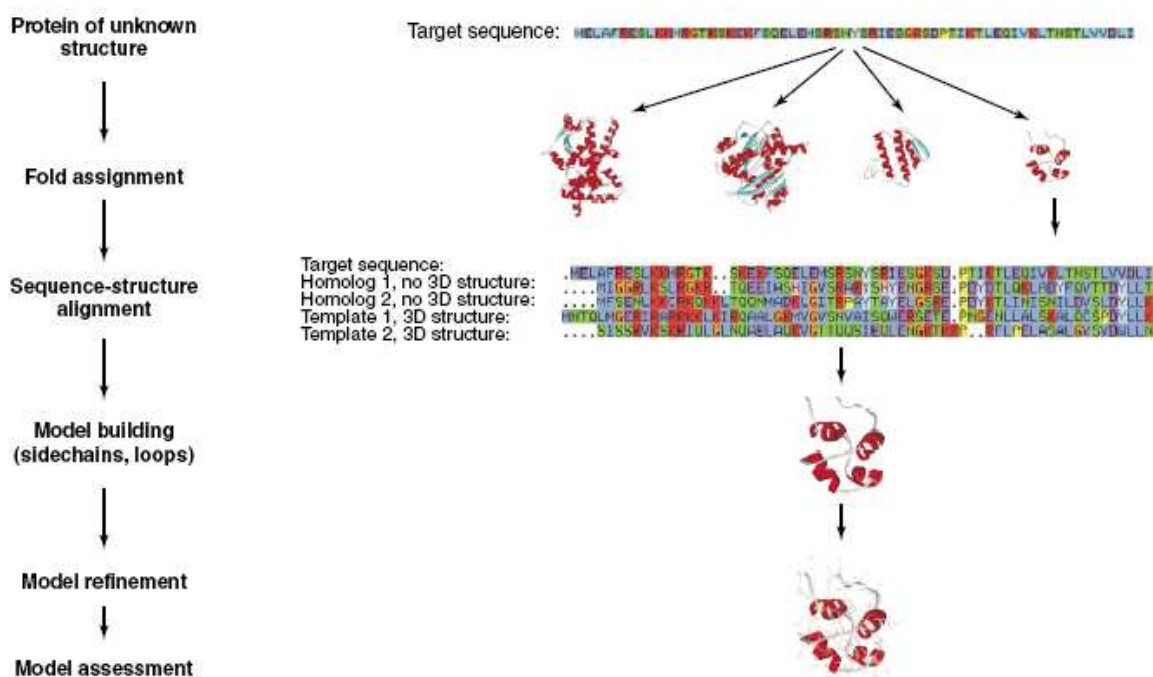


Figure 2.3: The steps involved in the prediction of protein structures following the homology model approach (Hillisch *et al.*, 2004).

1993), Swiss-Model (Gasteiger *et al.*, 2003), Profit (Martin, A.C.R. <http://www.bioinf.org.uk/software/profit/>), Prime (<http://www.schrodinger.com>), MOE ([www.chemcomp.com](http://www.chemcomp.com)) *etc.* MODELLER is a comparative protein modeling package which makes use of the satisfaction of spatial restraints to construct homology models and was used in this study. MODELLER proceeds by performing three steps to build a homology model (Sali and Blundell, 1993). These three steps include the alignment of target sequence with the template sequence, the extraction of spatial restraints and lastly the satisfaction of the spatial restraints. The protein sequence alignment is used to extract the distance and dihedral angles restraints for the model and is expressed as a probability density function which is used to describe the spatial restraints (Sali and Blundell, 1993). These restraints and energy terms which enforce proper stereochemistry are combined in an objective function. Optimization of the final model is performed by using methods of conjugate gradients and molecular dynamics with simulated annealing (Sali *et al.*, 1995).

The use of comparative protein modeling has been shown to be suitable for the predic-



tion of protein structures in *P. falciparum* and subsequently used in this study to predict the structure of PfSpdSyn (McKie *et al.*, 1998; Yuvaniyama *et al.*, 2003; Birkholtz *et al.*, 2004; de Beer *et al.*, 2006).

## 2.2. Methods

### 2.2.1. Homology Model Construction

Two crystal structures were used to construct the PfSpdSyn, which included the spermidine synthase crystal structures of *Thermotoga maritima* (1.80Å Resolution, PDBid: 1JQ3) and *Arabidopsis thaliana* (2.70Å Resolution, PDBid: 1XJ5). The spermidine synthase crystal structure from *Caenorhabditis elegans* was omitted as a template since the gate-keeping loops were not resolved (Blundell *et al.*, 2006). The spermidine synthase crystal structure from *T. maritima* (TmSpdSyn; PDBid 1JQ3) was co-crystallized with a combined substrate-product analogue, S-adenosyl-1,8-diamino-3-thiooctane (AdoDATO; Korolev *et al.* (2002)). In an attempt to find an optimal alignment between the template and target sequences, the T-Coffee package was used to align SpdSyn sequences retrieved from the UniProt and PDB databases, resulting in a protein family alignment (Notredame *et al.*, 2000; Apweiler *et al.*, 2004; Breman, 2001). The experimentally aligned protein sequences of the spermidine synthases of *T. maritima* and *A. thaliana* were used as templates to construct the homology model of *P. falciparum*. PfSpdSyn (UniProt entry: Q9FS5) was indicated to have a sequence identity of 32% and 49% with *T. maritima* and *A. thaliana*, respectively. Homology models of PfSpdSyn containing the substrate analogue, AdoDATO, were constructed using MODELLER 6v3 (Sali and Blundell, 1993). The PfSpdSyn models were subjected to stereochemical analysis, using PROCHECK, to evaluate the quality of the model (Laskowski *et al.*, 1993).

The PfSpdSyn homology models were further subjected to refinement by using the CHARMM (Chemistry at HARvard Molecular Mechanics) package (Brooks *et al.*, 1983). The partial charges used in the construction of the residue topology file of AdoDATO were computed using the MOPAC module within the InsightII (Accelrys; [www.accelrys.com](http://www.accelrys.com))

package. The PfSpdSyn models containing AdoDATO were subjected to 500 steps of steepest descent minimization followed by 50 steps of Adopted Basis Newton-Raphson (ABNR) minimization using the CHARMM27 all-atom empirical force field for proteins and nucleic acids (Igarashi *et al.*, 1982). The protein-ligand interactions between the PfSpdSyn and TmSpdSyn structures were then compared using LIGPLOT (Wallace *et al.*, 1995).

### 2.2.2. Binding Site Analysis

AdoDATO was removed from the PfSpdSyn model and docked back into the model using Cerius2 ([www.accelrys.com](http://www.accelrys.com)). The docked and build-in AdoDATO were compared with AdoDATO crystallized within TmSpdSyn. The comparison was made using LIGPLOT and visual inspection (Wallace *et al.*, 1995). The homology model of PfSpdSyn containing the build-in AdoDATO was used in further analysis. Evaluation of the binding cavity of PfSpdSyn was done using the LigandFit module of Cerius2 ([www.accelrys.com](http://www.accelrys.com)). Two binding cavities could be distinguished, one for binding of dcAdoMet and the other for putrescine binding.

### 2.2.3. Protein-substrate Interactions

Information obtained from the binding site analysis was subsequently used to elucidate protein-substrate interactions. The moieties of the substrate analogue AdoDATO were converted into dcAdoMet and putrescine using InsightII ([www.accelrys.com](http://www.accelrys.com)). The attacking nitrogen of putrescine was in the deprotonated state since it needs to perform a nucleophilic attack on the electrophilic carbon of dcAdoMet. The PfSpdSyn model containing the newly formed substrates were then subjected to 100 steps of steepest descent minimization using CHARMM (Brooks *et al.*, 1983). Putrescine adopted a strongly angular conformation as an artifact of the minimization conditions and was restored to a linear conformation using InsightII (Accelrys). The PfSpdSyn, AdoDATO complex was then further minimized for 400 steps using steepest descent minimization. To determine the protein-ligand interactions between PfSpdSyn and the protonated putrescine a further 100 minimization steps were performed with putrescine in the proto-

nated state. Protein-substrate interactions were evaluated using LIGPLOT and visual inspection (Wallace *et al.*, 1995).

#### 2.2.4. Molecular Dynamics

Molecular dynamics (MD) was performed on the homology models containing the substrate analogue AdoDATO as well as the substrates, putrescine and dcAdoMet. The protein was solvated with TIP3 water molecules. Molecular dynamics was started by 5 000 steps of steepest descent minimization followed by 200 steps of ABNR minimization. The system was then heated to 310K in steps of 5K every 100 steps and left to equilibrate for 10 picoseconds (ps). The molecular dynamics simulation was subsequently performed for 1 nanosecond (ns). VMD was used to visually inspect the molecular dynamics simulations of the homology models (Humphrey *et al.*, 1996). The site-directed mutagenesis models were also subjected to molecular dynamics and the same procedure as above was followed. Molecular dynamics simulations were also performed on models containing the products, spermidine and MTA as well as the SpdSyn inhibitor 4MCHA under the same conditions.

#### 2.2.5. Validation of homology model by site-directed mutagenesis

Three *in silico* mutants were generated for PfSpdSyn: Tyr102Ala, Asp196Asn and Ser197Ala. All mutants were constructed using the Biopolymer module in InsightII ([www.accelrys.com](http://www.accelrys.com)). The mutations were generated from PfSpdSyn models that contained both the natural substrates dcAdoMet and putrescine. The models were then subjected to molecular dynamics as described above.

#### 2.2.6. Site-directed mutagenesis and functional analysis of recombinant PfSpdSyn

The wild-type expression construct pTRCHisB: PfSPDS2 as described by Haider *et al.* (2005) was used as template in subsequent site-directed mutagenesis experiments. Mutations were created to change residues Tyr102 and Ser197 to Ala and Asp196 to Asn. Primers used in the mutagenesis reactions were as follows (5' to 3'):

Asp196Asn-Sense	TATGATGTTATTATCGTAAATAGTTCAGATCCAATAGGA
Asp196Asn-Antisense:	TCCTATTGGATCTGAACTATTTACGATAATAACATCATA
Ser197Ala-Sense	GATGTTATTATCGTAGATGCTTCAGATCCAATAGGACCA
Ser197Ala-Antisense	TGGTCCTATTGGATCTGAAGCATCTACGATAATAACATC
Tyr102Ala-Sense	GAAAAAGATGAATTTGCTGCTCATGAAATGATGACACAT
Tyr102Ala-Antisense	ATGTGTCATCATTTTCATGAGCAGCAAATTCATCTTTTTTC

Mutagenesis protocols were performed according to the methods as described previously (Birkholtz *et al.*, 2004; Wrenger *et al.*, 2001). Subsequently, wild-type as well as the three mutant proteins were expressed as His-tag fusion proteins in BLR(DE3) *E. coli* as described by Haider *et al.* (2005). Proteins were isolated by Ni-affinity chromatography and aminopropyltransferase activity was determined by measuring the formation of  $^{14}\text{C}$  labeled reaction products from  $1,4^{14}\text{C}$ -putrescine (Haider *et al.*, 2005). Enzyme activities are expressed as specific activity ( $\text{nmol}^{-1}\text{min}^{-1}\text{mol of protein}^{-1}$ ) and the results are the means  $\pm$  S.D. from three independent experiments performed in duplicate with the specific activity normalized to the wild-type activity.

## 2.3. Results and Discussion

### 2.3.1. Comparative Modeling of PfSpdSyn

Thirty-four spermidine synthase sequences were retrieved from the UniProt database, which were associated with four structures from PDB (Apweiler *et al.*, 2004; Breman, 2001). The T-Coffee package was used to construct a protein family alignment to optimize the target-template alignment (Notredame *et al.*, 2000). The crystal structures of *Arabidopsis thaliana* (AtSpdSyn) and *Thermotoga maritima* (TmSpdSyn) spermidine synthase were used as templates for modeling. The first 39 amino acids of the PfSpdSyn sequence were omitted due to insufficient template match (Figure 2.4). AtSpdSyn was used due to its high sequence identity and that of TmSpdSyn since it contained the substrate analogue, AdoDATO. Models were built using Modeller with and without AdoDATO (Sali and Blundell, 1993). The active site residues interacting with AdoDATO appear to be highly conserved between the human spermidine synthase (HsSpdSyn), Tm-

Table 2.1: A comparison between the active site residues of *Thematoga maritima* (TmSpdSyn), *Homo sapiens* (HsSPdSyn), *Arabidopsis thaliana* (AtSpdSyn) and *Plasmodium faciparum* (PfSpdSyn) spermidine synthases, indicating its conserved nature.

TmSpdSyn	HsSpdSyn	AtSpdSyn	PfSpdSyn
Gln <sub>46</sub> <sup>a</sup>	Gln <sub>49</sub>	Gln <sub>76</sub>	Gln <sub>72</sub> <sup>a</sup>
Leu <sub>62</sub>	Leu <sub>65</sub>	leu <sub>92</sub>	Leu <sub>88</sub>
Met <sub>67</sub>	Gln <sub>170</sub>	Gln <sub>97</sub>	Gln <sub>93</sub>
Tyr <sub>76</sub>	Tyr <sub>79</sub>	Tyr <sub>106</sub>	Tyr <sub>102</sub>
His <sub>77</sub> <sup>a</sup>	Gln <sub>80</sub>	Gln <sub>107</sub>	His <sub>103</sub> <sup>a</sup>
Gly <sub>98</sub>	Gly <sub>101</sub>	Gly <sub>128</sub>	Gly <sub>124</sub>
Gly <sub>99</sub>	Gly <sub>102</sub>	Gly <sub>129</sub>	Gly <sub>125</sub>
Asp <sub>101</sub> <sup>a</sup>	Asp <sub>104</sub>	Asp <sub>131</sub>	Asp <sub>127</sub> <sup>a</sup>
Glu <sub>121</sub> <sup>a</sup>	Glu <sub>124</sub>	Glu <sub>151</sub>	Glu <sub>147</sub> <sup>a</sup>
Val <sub>122</sub>	Ile <sub>125</sub>	Ile <sub>152</sub>	Ile <sub>148</sub> <sup>a</sup>
Gly <sub>151</sub>	Gly <sub>154</sub>	Gly <sub>181</sub>	Ala <sub>179</sub> <sup>a</sup>
Asn <sub>152</sub> <sup>a</sup>	Asp <sub>155</sub>	Asp <sub>182</sub>	Asp <sub>178</sub> <sup>a</sup>
Asp <sub>170</sub> <sup>a</sup>	Asp <sub>173</sub>	Asp <sub>201</sub>	Asp <sub>196</sub> <sup>a</sup>
Ser <sub>171</sub>	Ser <sub>174</sub>	Ser <sub>202</sub>	Ser <sub>197</sub>
Thr <sub>172</sub>	Ser <sub>175</sub>	Ser <sub>203</sub>	Ser <sub>198</sub>
Asp <sub>173</sub> <sup>a</sup>	Asp <sub>176</sub>	Asp <sub>204</sub>	Asp <sub>199</sub> <sup>a</sup>
Gln <sub>178</sub> <sup>a</sup>	Pro <sub>180</sub>	Pro <sub>208</sub>	Pro <sub>203</sub>
Leu <sub>182</sub>	Leu <sub>184</sub>	Leu <sub>212</sub>	Leu <sub>207</sub>
Tyr <sub>239</sub>	Tyr <sub>241</sub>	Tyr <sub>270</sub>	Tyr <sub>264</sub>
Trp <sub>244</sub>	Ile <sub>246</sub>	Ile <sub>275</sub>	Ile <sub>269</sub>

<sup>a</sup>Indicates residues forming hydrogen bonds with AdoDATO.

Residues in blue indicate non-identical residues between the four organisms.

SpdSyn, AtSpdSyn and PfSpdSyn (Table 2.1).

Spermidine synthases are members of the putrescine aminopropyltransferase family that generally consists of a small *N*-terminal domain and a large catalytic *C*-terminal domain (Rossmann-like fold). The *N*-terminal domain of the monomer of PfSpdSyn consists of a six-stranded  $\beta$ -sheet ( $\beta$ 1-6 on Figure 2.5a) and the *C*-terminal domain of a seven stranded  $\beta$ -sheet flanked by nine  $\alpha$ -helices ( $\beta$ 7-13,  $\alpha$ 1-9, Figure 2.4 and 2.5a). A Ramachandran plot of the PfSpdSyn model showed 87% of its residues to be in the most favourable region (Figure 2.5b), which was similar to the template structures of AtSpdSyn and TmSpdSyn. PROCHECK confirmed that all the parameters of the PfSpdSyn model were within normal ranges (Laskowski *et al.*, 1993).

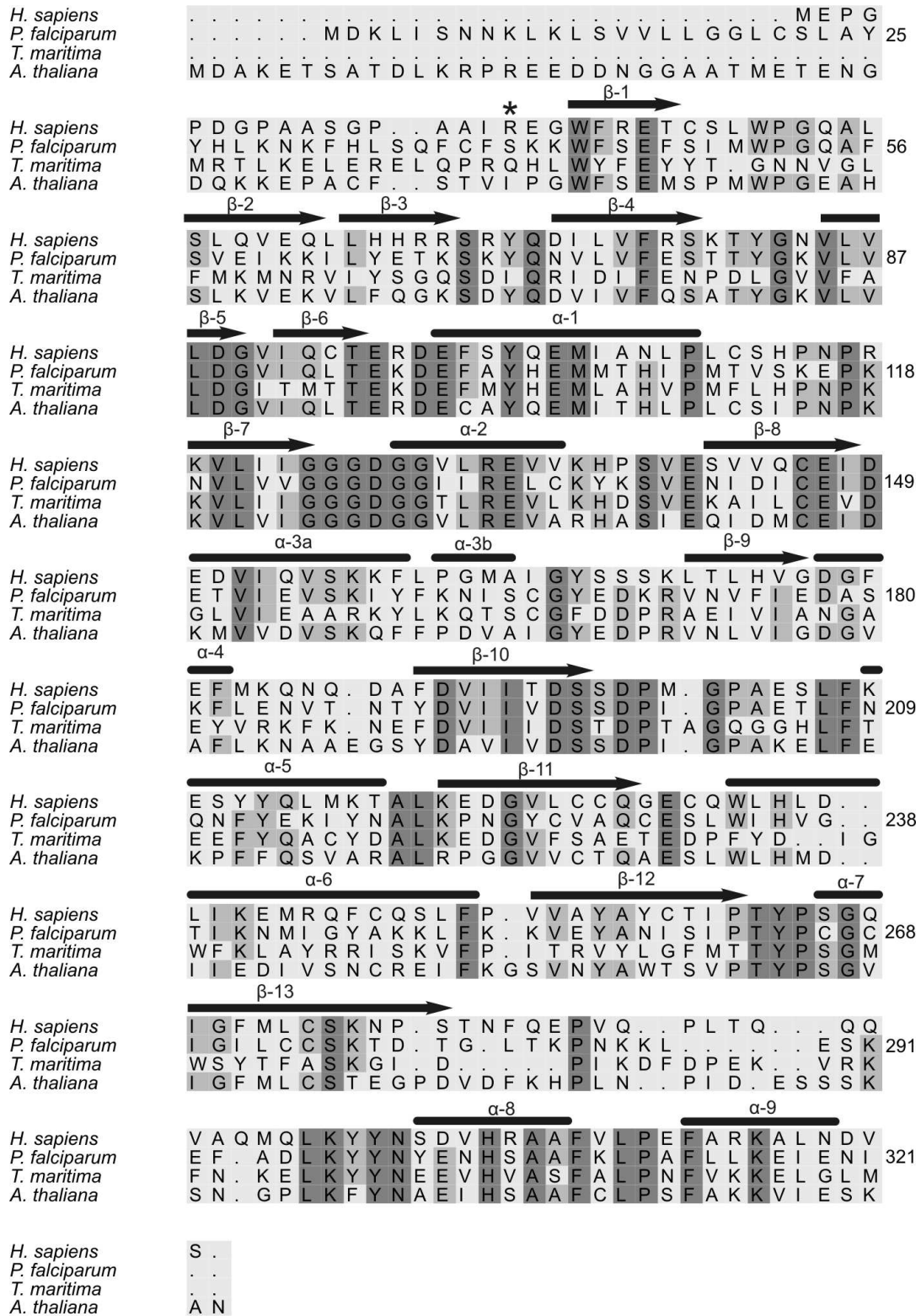


Figure 2.4: An alignment of the protein sequence of *P. falciparum* spermidine synthase and the spermidine synthases of *A. thaliana* and *T. maritima* used as templates during homology modeling. Also included in the alignment is the protein sequence of the human spermidine synthase. The cylinders indicate helices and the arrows indicate  $\beta$ -strands. The amino acids shaded light gray represent conservation between 50% and 80%, whereas the dark gray areas represent residues with conservation higher than 80%. The numbering used is in reference to the amino acid sequence of PfSpdSyn. \* indicates the start of the homology model.

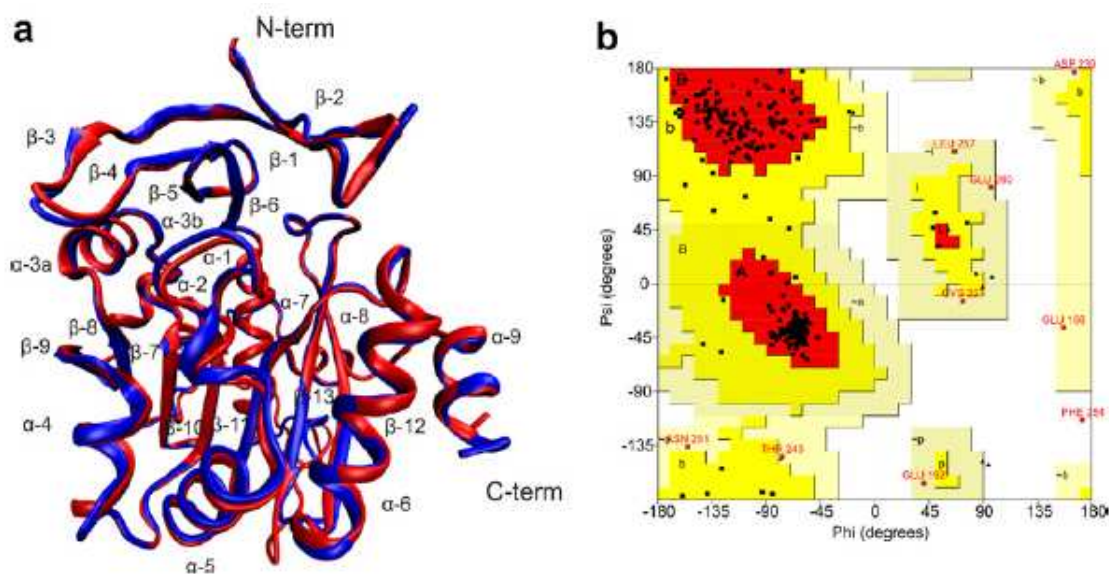


Figure 2.5: (a) A structural alignment of the PfSpdSyn model (red) and the crystal structure (blue; PDBid 2HTE). (b) Ramachandran plot for the model of PfSpdSyn as produced by PROCHECK. Eighty-seven percent of the residues were within favorable structural areas.

During this study a crystal structure of PfSpdSyn (PDBid: 2HTE) became available (<http://sgc.utoronto.ca>). Structural alignment revealed an excellent correlation between the C $\alpha$ -backbone of the crystal structure and model of the PfSpdSyn (root mean square deviation (RMSD) 0.594Å; Figure 2.5a). The active site residues identified to interact with AdoDATO, as illustrated in Table 2.1, were aligned and showed to be highly conserved (RMSD 0.476Å). The significant correlation between the crystal structure and homology model of PfSpdSyn provides support for the quality of the model used during this study.

### 2.3.2. Binding Cavity Analysis

From the PfSpdSyn model, the binding sites for putrescine and dcAdoMet were apparent (Figure 2.6a). The dcAdoMet binding cavity is represented by the residues surrounding the adenosyl-moiety of AdoDATO whereas the residues surrounding the polyamine moiety represents the putrescine binding cavity. The putrescine binding cavity in PfSpdSyn has a central hydrophobic region flanked by two negatively charged regions in agreement with suggestions by Korolev *et al.* (2002) and Shirahata *et al.* (1993). This

region is composed of Trp 51, Val 91, Tyr 102, Ile 235, Tyr 246, Pro 247 and Ile 269 (not shown; Korolev *et al.* 2002; Ikeguchi *et al.* 2006). The two negatively charged/electron donating regions consist of Gln 93, Tyr 102, Asp 196, Ser 197, Gln 229 and Glu 231, Asp 199, His 236, respectively (Figure 2.6a).

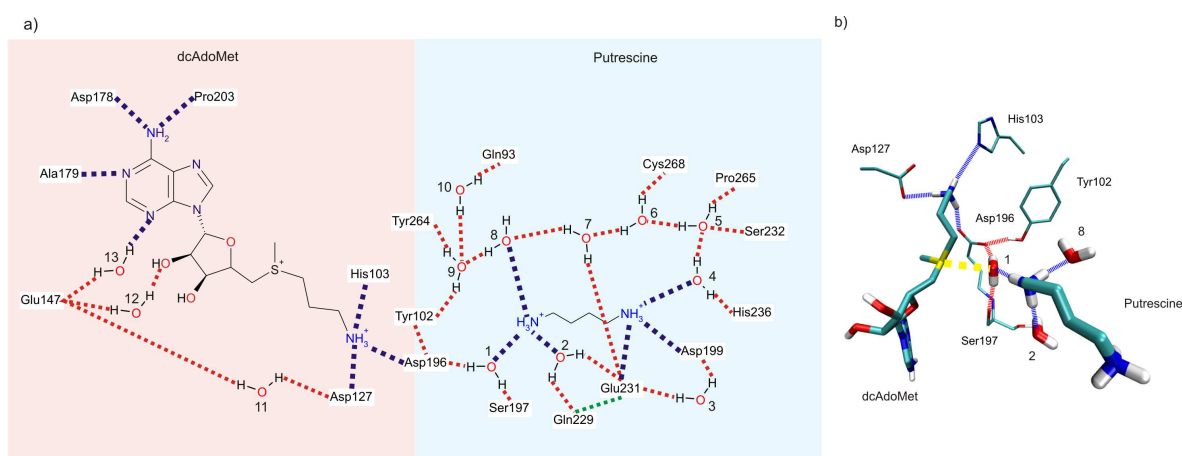


Figure 2.6: (a) A 2D representation of the interactions between PfSpdSyn and its substrates. The dcAdoMet binding cavity is represented in the apricot shaded area to the left and the putrescine binding cavity in the blue area to the right. Water molecules thought to anchor and orient putrescine (via hydrogen bonds represented as dashed red lines) are indicated and labeled numerically. Dashed blue lines indicate hydrogen bonds with nitrogen atoms. A protein-protein interacting hydrogen bond between Gln 229 and Glu 231 is represented in green. (b) A 3D representation of interactions showing the most important interactions for substrate binding and catalysis. Dashed lines in red and blue represent hydrogen bonds with the substrates. The polar interaction between water molecule 1 and the positively charged sulphur of dcAdoMet is colored in yellow.

AdoDATO was converted *in silico* to dcAdoMet and putrescine as an indirect means of determining protein-substrate interactions and the protein-substrate complex was subsequently minimized. Evaluation of the interactions between the substrates and PfSpdSyn revealed that Tyr 102, Asp 196 and Ser 197 interact with the attacking nitrogen of putrescine. These residues have been suggested by Korolev *et al.* (2002) to play a central role in the deprotonation of putrescine. Interactions with dcAdoMet were conserved to those of the dcAdoMet moiety of AdoDATO (see below).



### 2.3.3. Dynamic protein-substrate interaction analyses

The PfSpdSyn model containing both substrates was subsequently solvated with water (TIP3) and subjected to molecular dynamic analysis to investigate the protein-substrate interactions in the presence of water. A simulation was performed with the attacking nitrogen of putrescine in the unprotonated state to ensure the correct orientation of the substrates in the protein. After about 20ps a protein state was captured and the attacking nitrogen of putrescine protonated. The model was again subjected to molecular dynamics and a network of water molecules identified that potentially interact with the substrates (Figure 2.6a).

Eight hydrogen bonds were predicted between PfSpdSyn and dcAdoMet. Hydrogen bonds were directly formed between dcAdoMet and Asp 127, Asp 178, Ala 179, Asp 196, His 103 and Pro 203 (Figure 2.6a). Glu 147 mediates two hydrogen bonds via water molecules 12 and 13 and an additional hydrogen bond with water molecule 11, which in turn forms a hydrogen bond with Asp 127. Asp 127 together with His 103 and Asp 196 forms hydrogen bonds with the aminopropyl chain of dcAdoMet (Figure 2.6a and b). It is proposed that these three hydrogen bonds are necessary to orient the aminopropyl chain in order to present the electrophilic carbon for a nucleophilic attack by putrescine.

Only two direct hydrogen bonds are formed via residues Asp 199 and Glu 231 between PfSpdSyn and putrescine. However, an extensive network of ten water molecules was found around putrescine, which is thought to play a role in anchoring and orienting putrescine in the active site in such a way that catalysis can take place (Figure 2.6a). The attacking nitrogen forms hydrogen bonds with water molecules 1, 2 and 8. The non-attacking nitrogen of putrescine forms three hydrogen bonds with water molecule 4, Asp199 and Glu231 (Figure 2.6a and b).

Water molecule 1 was of most interest since it was anchored between Asp 196 and the carbonyl group of Ser 197 most of the time during the 1ns simulation (Figure 2.6b). It forms hydrogen bonds with Asp 196 and Ser 197 and is oriented in such a way that

a hydrogen bond is formed with the attacking nitrogen of putrescine as well as a polar interaction with the positively charged sulphur of dcAdoMet. It therefore appears that water molecule 1 plays an important role during catalysis by bringing and holding together both the substrates. The importance of this water molecule was further supported by a 1ns simulation of the PfSpdSyn model without substrates, which showed the presence of a water molecule between Asp 196 and Ser 197 throughout the simulation. Although this water molecule might get exchanged for another during the simulation, a water molecule was consistently found to occupy this molecular space. This water molecule also alternated its hydrogen bonds between Asp 196, Ser 197 and Gln 229 in the absence of substrates. Similar water molecules were also evident in the crystal structures of PfSpdSyn (HOH 8) and AtSpdSyn (HOH 125). The crystal structure of TmSpdSyn without AdoDATO, (PDBid 1INL) showed a similar water molecule as in the above mentioned structures in chain C only (HOH 687), whereas in the presence of AdoDATO (PDBid 1JQ3), no such water molecule was found. In the AdoDATO co-crystallised quarternary TmSpdSyn structure, one of the monomers has an unresolved gate-keeping loop and the other three have AdoDATO crystallized within it, therefore not allowing the anchoring of the water molecule because the aminopentyl chain (putrescine moiety) of AdoDATO would displace it when it binds.

Hydrogen bonds corresponding to the bonds formed between Tyr 102 and Asp 196 in the PfSpdSyn model are also conserved within the crystal structures of TmSpdSyn (without AdoDATO), PfSpdSyn, AtSpdSyn and HsSpdSyn. It is proposed that the hydrogen bond formed between Tyr 102 and Asp 196 orients Asp 196 in such a way that a hydrogen bond is formed with the aminopropyl chain of dcAdoMet, which together with His 103 and Asp 127 anchors this chain. This further supports the important role of Asp 196 in the anchoring of the aminopropyl chain and the orientation of the electrophilic carbon, allowing for a nucleophilic attack by putrescine.

### 2.3.4. Proposed mechanism of action of PfSpdSyn mediated by a gate-keeping loop

Based on the results presented here, water molecule 1 (which anchors residues Asp 196 and Ser 197) is therefore proposed to play an important role during the aminopropyl transfer reaction by facilitating the deprotonation of putrescine to allow an electrophilic attack on dcAdoMet (Figure 2.7a). It is possible that Asp 196 removes a hydrogen from water molecule 1, which in turn deprotonates the attacking nitrogen of putrescine. The deprotonated nitrogen consequently attacks the electrophilic carbon of dcAdoMet resulting in the addition of the aminopropyl chain to putrescine.

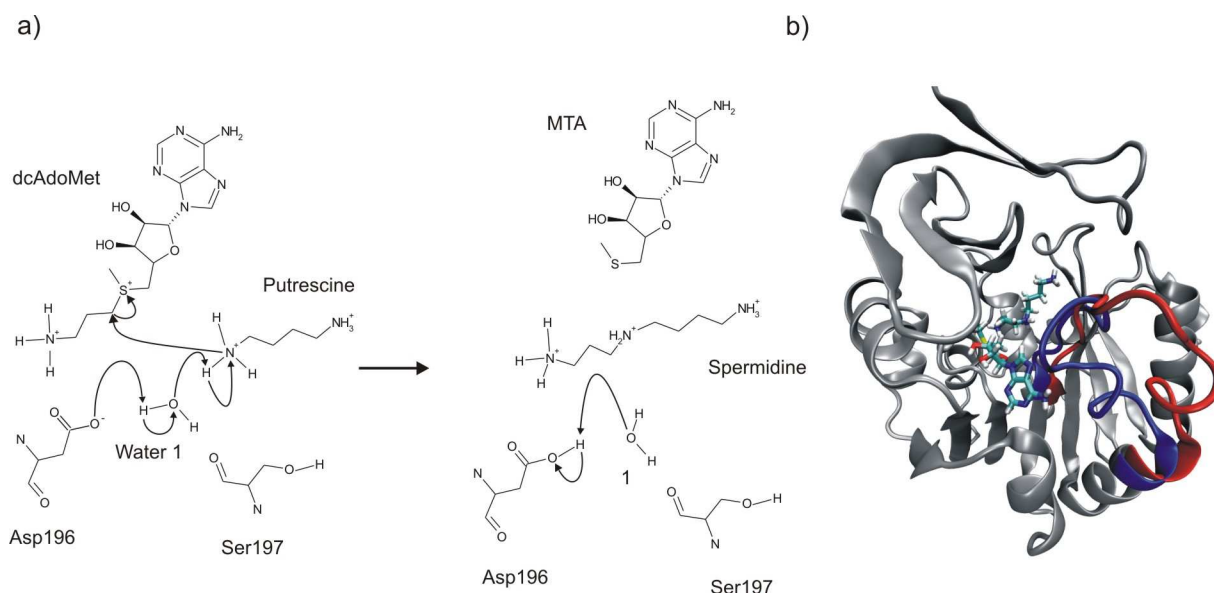


Figure 2.7: (a) A more detailed mechanism of action is presented for PfSpdSyn. Water molecule 1 is attacked by Asp196 removing the hydrogen from it. The attacking nitrogen of putrescine is in turn deprotonated by the water molecule to enable a nucleophilic attack on the electrophilic carbon of dcAdoMet and the formation of spermidine and MTA. Asp 196 is deprotonated by water molecule 1 in order to regenerate the active form of PfSpdSyn for further catalysis to take place. (b) Loop movement reported during molecular dynamics of the PfSpdSyn containing the products of MTA and spermidine. The gate-keeping loop of PfSpdSyn in both an open (red) and closed conformation (blue) is represented by a cartoon.

The results are the formation of the two products, spermidine and MTA. After the formation of the products, water molecule 1 removes the hydrogen from Asp 196 releasing it into the system and by so doing, regenerates the normal state of the protein.

Aminopropyltransferase reactions are proposed to be mediated by a gate-keeping loop (Ikeguchi *et al.*, 2006; Korolev *et al.*, 2002). The PfSpdSyn model containing both reaction products (spermidine and MTA) was created and evaluated with molecular dynamics. It was observed that during this simulation, the gate-keeping loop covering entry to the active site is mobile and opens up (Figure 2.7b). Subsequently, MTA started to move out of the active site as the gate-keeping loop opens, after which spermidine can exit. Therefore, this dynamic visualization of the gate-keeping loop mobility provides for the first time evidence of gatekeeper activity in the spermidine synthase active site.

### 2.3.5. *In vitro* validation by site-directed mutagenesis

Site-directed mutagenesis was utilized to test the proposed importance of residues Tyr 102, Asp 196 and Ser 197. Tyr 102 and Ser 197 were mutated to Ala and Asp196 was mutated to Asn. The Asp196Asn mutation showed an 89% loss of activity, the Tyr102Ala change resulted in 91% loss of activity whereas the Ser197Ala mutant had only 24% loss in activity (Figure 2.8a). These mutant forms of the protein were also analysed *in silico* on the PfSpdSyn model that contained both the natural substrates.

As shown above (Figure 2.6), one of the oxygens of Asp 196 protrudes into the aminopropyl binding cavity, anchoring the aminopropyl group of dcAdoMet, whereas the other is oriented towards the putrescine binding cavity. Structural analysis of the Asp196Asn mutant PfSpdSyn model evaluated by molecular dynamics indicated that the replacement of the carboxylic acid of Asp by the amide group of Asn results in two different orientations of the amide group (Figures 2.8b and c). In the first case, the nitrogen of the amide group protrudes into the aminopropyl binding cavity and in the second case it is oriented towards the putrescine binding cavity. Simulation of the first scenario suggested that the anchoring of the aminopropyl chain of dcAdoMet was weakened by the loss of its hydrogen bond with Asp 196 and it consequently moved out of its binding pocket into the catalytic space. This in turn caused putrescine to move deeper into its binding cavity with the attacking nitrogen of putrescine now interacting with Gln 229 and the backbone oxygen of Ser 197 (Figure 2.8b). The hydrogen bond between Asp 196 and Tyr

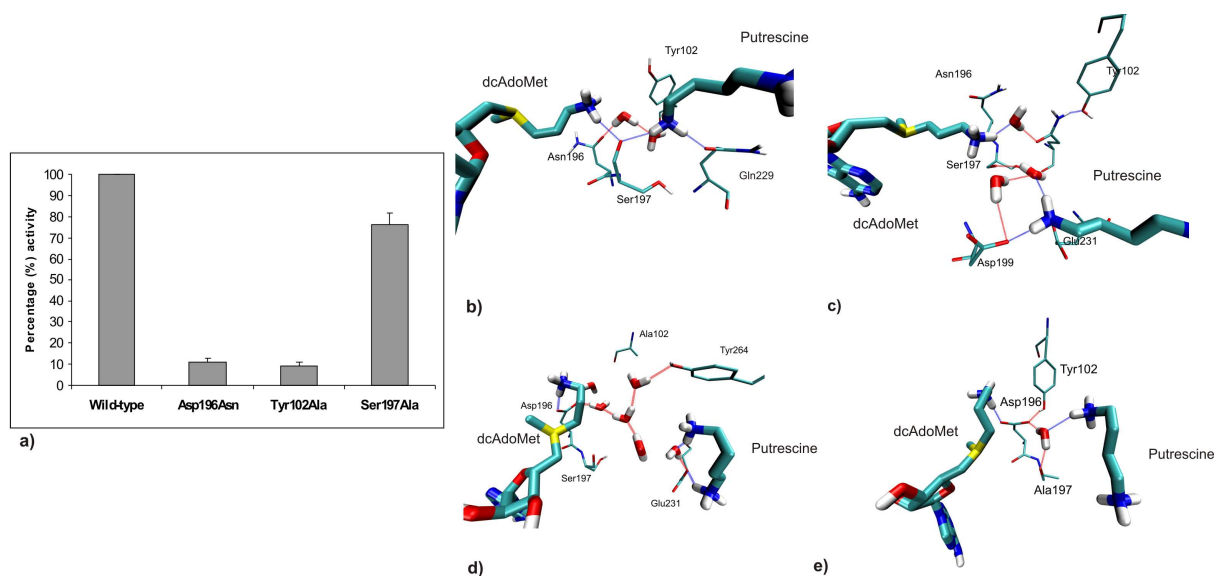


Figure 2.8: Graphical presentation of specific activity after point mutations performed on PfSpdSyn. (a) Effect of the point mutations Asp196Asn, Ser197Ala and Tyr102Ala on the specific activity of PfSpdSyn compared to the wild-type. (b) The Asp196Asn point mutation with the amine group of Asn pointing into the aminopropyl binding cavity (Scenario 1). (c) The Asp196Asn point mutation with the carbonyl oxygen of Asn pointing into the aminopropyl binding cavity (Scenario 2). (d) Point mutation Tyr102Ala with the most important hydrogen bond formation indicated. (e) Point mutation Ser197Ala with the most important hydrogen bond formation indicated.

102 is also lost after the introduction of the mutation. It is therefore neither possible for the protein to anchor the aminopropyl chain of dcAdoMet nor to anchor the attacking nitrogen of putrescine, which explains the drastic reduction in catalysis. In the second scenario, a hydrogen bond with Tyr 102 is not formed resulting in the rotation of the amide group of Asn 196. This allowed the aminopropyl group of dcAdoMet to move out of its binding pocket and consequently also out of the catalytic center of the protein. Putrescine subsequently moved deeper into its binding cavity with its attacking nitrogen now forming hydrogen bonds with Glu 231, Asp 199 and a water molecule mediated by the backbone oxygen of Ser 197 (Figure 2.8c). Both these simulated scenarios indicated that the Asp196Asn mutation distorted the anchoring and orientation of putrescine and the aminopropyl chain of dcAdoMet and hence caused the marked (89%) reduction in activity.

Simulation of the Tyr102Ala mutation on the PfSpdSyn model indicated that this

mutation also drastically altered the environment of the active site, inducing a loss of the hydrogen bond between Tyr 102 and Asp 196 (Figure 2.8d). The hydrogen bond loss allowed Asp 196 to move slightly towards the aminopropyl binding cavity of dcAdoMet, preventing the anchoring of a water molecule between Asp 196 and Ser 197. This in turn prevented the anchoring of the attacking nitrogen of putrescine. The attacking nitrogen of putrescine initially formed a novel hydrogen bond with a water molecule mediated by Gln 93, which was later replaced by hydrogen bonds formed randomly with water molecules occupying the molecular space between putrescine and dcAdoMet (Figure 2.8d). The hydrogen bond formation of the non-attacking nitrogen of putrescine was unchanged compared to the wild-type. The hydrogen bonds formed with the aminopropyl chain of dcAdoMet, Asp 127, Asp 196 and His 103 were retained. Since Asp 196 was not anchored by the hydrogen bond formed with Tyr 102 it was allowed to move, causing a shift in the orientation of the aminopropyl chain of dcAdoMet, which resulted in an unfavourable orientation between dcAdoMet and putrescine. The importance of Tyr 102 in anchoring Asp 196 for the proper orientation of both the aminopropyl chain of dcAdoMet and the attacking nitrogen is therefore reflected in the 91% loss in activity.

The Ser197Ala mutation resulted in a 24% loss in activity. Ser 197 together with Asp 196 is thought to be important since they anchor water molecule 1 which in turn interacts with the attacking nitrogen of putrescine and the positively charged sulphur of dcAdoMet. During the simulation of the PfSpdSyn model containing the substrates, it was shown that water molecule 1 interacts with the backbone carbonyl group of Ser 197, which was also true for the molecular dynamic simulation without any substrates (Figure 2.6). It is therefore proposed that the interaction formed with the carbonyl group of Ala instead of the carbonyl group of Ser 197 with water molecule 1, allowed catalysis to take place only 76% of the time (Figure 2.8e). The main change induced in the chemical environment of the active site was the loss of the hydrogen bond between Ser 197 and Gln 229, thought to give rigidity to the gate-keeping loop. The Ser197Ala mutation is expected to make the loop more flexible and thus reducing the activity of the enzyme. Since this study was not intended to investigate loop movement, further analysis is however needed to confirm this

hypothesis. As with the simulation of PfSpdSyn containing both the substrates, a water molecule was identified during the PfSpdSyn Ser197Ala mutation simulation to play a role in the anchoring and orientation of both putrescine and the aminopropyl chain of dcAdoMet (Figure 2.8e).

In summary, the mutation and molecular dynamic studies indicated that Tyr 102, Asp 196 and Ser 197 play important roles in catalysis of PfSpdSyn. Tyr 102 is essential in the orientation of Asp 196 to allow this residue to anchor both the aminopropyl chain of dcAdoMet and putrescine during catalysis. Ser 197 was found to interact with its carbonyl backbone with water molecule 1 to facilitate the anchoring of putrescine. This data therefore validates the proposed mechanism of catalysis as predicted by the dynamic visualization of aminopropyl transfer in the PfSpdSyn model.

### 2.3.6. Inhibitor studies

As mentioned previously, investigations of seven known SpdSyn inhibitors have shown 4MCHA to be the most potent inhibitor of PfSpdSyn (Haider *et al.*, 2005). This inhibitor was further investigated in this study to obtain a dynamic view of the aminopropyltransferase inhibition of PfSpdSyn using molecular dynamics. 4MCHA was overlaid with putrescine in its binding cavity and two different scenarios were simulated. The amine group of 4MCHA was aligned with either the attacking nitrogen of putrescine or the non-attacking nitrogen of putrescine. Both molecular dynamics simulations revealed that 4MCHA binds with its amine group in the putrescine binding cavity, partially filling it, whereas the hydrophobic moiety of the 4MCHA molecule protrudes into the adjacent hydrophobic cavity constituted by residues Met 50, Trp 51, Phe 56, Val 91, Ile 92, Tyr 102, Ile 201, Ile 235, Tyr 264 and Pro 265 (Figure 2.9). This binding pattern correlates with previous reports, which proposed that the related inhibitor, cyclohexylamine, binds partially in the putrescine binding cavity and the adjacent hydrophobic cavity (Shirahata *et al.*, 1993; Ikeguchi *et al.*, 2006; Pegg *et al.*, 1995; Goda *et al.*, 2004). Since the only difference between cyclohexylamine and 4MCHA is the addition of a methyl group

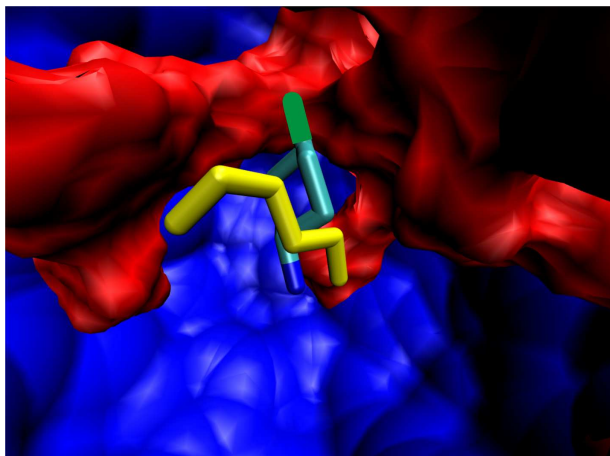


Figure 2.9: 4MCHA (cyan) was found to bind partially in the putrescine binding cavity with its nitrogen group (blue). The methylcyclohexyl group was found to bind in an adjacent hydrophobic cavity with the methyl group (green) lodged into the roof of the cavity. The hydrophobic regions are represented in red and the rest of the protein in blue. Putrescine is represented in yellow indicating the orientation of 4MCHA.

to 4MCHA, the PfSpdSyn simulations containing 4MCHA can additionally be used as evidence to support the proposed binding of cyclohexylamine.

## 2.4. Conclusion

At the start of this study no crystal structure data was available for the PfSpdSyn. Since then eight crystal structures have been obtained in either the apo form (without ligand) or in complex with ligands. The analysis performed on the homology model of PfSpdSyn in this study was supported by an excellent correlation with the PfSpdSyn crystal structure, which became available during the preparation of this chapter. A crystal structure of the human SpdSyn co-crystallized with putrescine and MTA was released as the work in this chapter was being published at the end of 2006 (PDBid 2O06; 2.00Å). The interactions of putrescine could be determined from the HsSpdSyn structure, which differ from the interactions proposed in this study in that it had no water molecule present which interacts with the attacking nitrogen of putrescine (Wu *et al.*, 2007). The attacking nitrogen of putrescine was shown to bind in the highly negatively charged environment provided by Asp 173, Ser 174 (backbone), Tyr 79 and Tyr 241 (Figure 2.10). In the present study a water molecule was found to occupy this negatively charged cavity which



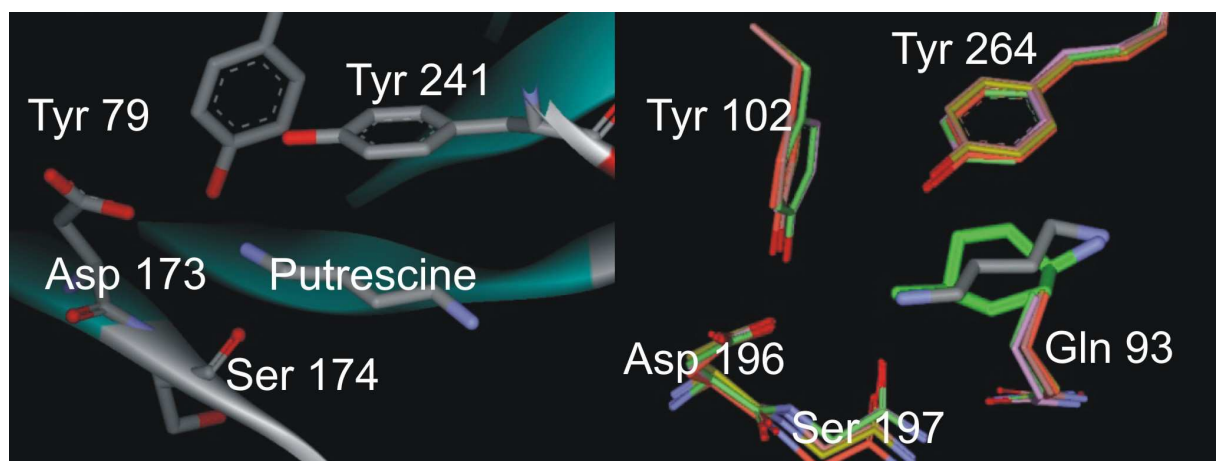


Figure 2.10: A comparison of the negatively charged binding pocket which binds the attacking nitrogen of putrescine of HsSpdSyn (PDBid 2O06) and PfSpdSyn (PDBid 2PT9). (*Left*) It is proposed by Wu *et al.* (2007) that Tyr 79, Asp 173, Ser 174 and Tyr 241 plays a role in the deprotonation of the attacking nitrogen of putrescine. (*Right*) Residues corresponding to the ones proposed by Wu *et al.* (2007) for the PfSpdSyn are shown on the right and include Tyr 102, Asp196, Ser 197 and Tyr 264, also shown is an overlay of 4MCHA (green) and putrescine (gray) within the PfSpdSyn active site.

is constituted of Asp 196, Ser 197 (backbone), Tyr 102 and Tyr 264 in PfSpdSyn (Figure 2.10).

A structural comparison between the PfSpdSyn crystal structure (PDBid 2PT9) and the HsSpdSyn was made which shows that the active sites are very similar (Figure 2.10). The PfSpdSyn (PDBid 2PT9) was crystallized in complex with 4MCHA and dcAdoMet which were superimposed on the HsSpdSyn crystal structure containing putrescine and MTA. The binding poses of putrescine and 4MCHA are very similar (Figure 2.10). This called for a reevaluation of our findings regarding the protein substrate binding.

Since putrescine could not be docked to PfSpdSyn and the conversion of AdoDATO had to be used to generate the two naturally occurring substrates within their respective binding pockets, errors might have been incorporated into the system. It was assumed that the conversion of AdoDATO to putrescine and dcAdoMet resulted in the correct binding poses. However during the solvation of the protein the active site was flooded by water molecules. No water molecules were shown within the crystal structures of either

the PfSpdSyn or HsSpdSyn. As was discussed at length in the above sections, the presence of water molecule 1 was supported by its presence in the apo-structures of PfSpdSyn, AtSpdSyn and TmSpdSyn (PDBid 1INL). However, it appears that this water molecule is replaced upon binding of putrescine to SpdSyn. After the conversion of AdoDATO to putrescine and dcAdoMet the PfSpdSyn were subjected to a MD simulation. The orientation of putrescine within its binding cavity was different from that proposed by the HsSpdSyn crystal structure, which resulted in a change in the orientation of Gln 229 during the MD simulation. After overlaying of 4MCHA and putrescine this Gln 229 conformation resulted in a strong repulsion of the cyclohexyl ring of 4MCHA forcing it to adapt the orientation proposed in Figure 2.9. The importance of the correct orientation of Gln 229 will become more apparent in Chapter 3 where it is discussed at length.

It can therefore be concluded that differences in the results derived from the PfSpdSyn homology model and the crystal structures can be contributed to problems during optimization of the PfSpdSyn system and not due to the model. Similar problems were encountered during further studies of PfSpdSyn and are discussed in detail in Chapter 3.