Aminoacyl-tRNA Synthetases as Targets for Structure Guided Drug Design (SGDD) Against Pathogenic Protozoa and Bacteria

Ximena Barros Alvarez

A dissertation
submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

University of Washington
2018

Reading Committee:
Wilhelmus G.J. Hol, Chair
Christophe L.M.J. Verlinde
Frederick S. Buckner

Program Authorized to Offer Degree:
Biochemistry

University of Washington

Abstract

Aminoacyl-tRNA Synthetases as Targets for Structure Guided Drug Design (SGDD) Against Pathogenic Protozoa and Bacteria

Ximena Barros Alvarez

Chair of the Supervisory Committee:

Professor Wilhelmus G.J. Hol

Biochemistry

Tuberculosis (TB) and neglected tropical diseases (NTDs) caused by trypanosomatids are devastating diseases affecting millions of people around the globe. *Mycobacterium tuberculosis* causes TB, while the trypanosomatids *Trypanosoma brucei*, *Trypanosoma cruzi* and parasites of the genus *Leishmania*, cause sleeping sickness (or human African trypanosomiasis (HAT)), Chagas disease (or American trypanosomiasis) and the leishmaniases in tropical and subtropical areas of the world. Visceral leishmaniasis (VL), the deadliest form of the disease, is caused by *L. infantum* and *L. donovani*. For some of these diseases there is no vaccine or cure. For others, vaccine protection and treatment efficiency are limited. In some cases, development of resistance to available drugs has made useless otherwise successful treatments. New drugs as well as new drug targets are desperately needed. The essential aminoacyl-tRNA synthetase (aaRS) enzymes

provide the charged tRNAs required for protein synthesis. AaRS have been previously pursued as drug targets in bacteria and fungi and have been validated as drug targets in protozoa. The structural work presented as part of this dissertation has been part of collaborative structure guided drug design (SGDD) projects among various research groups, most of them within University of Washington, for the discovery and iterative optimization of inhibitors targeting aaRSs of parasitic protozoa and bacteria. The selection of methionyl-tRNA synthetase (MetRS) and tyrosyl-tRNA synthetase (TyrRS) as drug targets was done based on their predicted feasibility of developing selective inhibitors. Crystal structures of *M. tuberculosis* and *T. brucei* MetRS (*Mtub*MetRS and *Tbru*MetRS) and *L. donovani* TyrRS (*Ldon*TyrRS) were solved in the presence of different compounds to assist in the iterative SGDD development of drugs against TB, HAT and VL, respectively. Structural information contributed in different stages in the SGDD process, from the description of new protein structures of the essential pathogenic aaRSs to the assistance in the optimization and design of novel inhibitors. In an example of early steps in the SGDD process, the crystal structure of *Mtub*MetRS in complex with the catalytic intermediate Met-AMP was solved at 2.6 Å resolution. Differences with other MetRSs including the human counterparts were revealed and could potentially be useful in the chemotherapeutic development against TB. The use of nanobodies as crystallization chaperones and of the tyrosyl adenylate analog TyrSA was crucial for obtaining well diffracting crystals that lead to solving the crystal structure of *Ldon*TyrRS at 2.75 Å resolution. The presence of an extra pocket (EP) was revealed that is not present in the human counterparts, but is shared with other pathogens, and could be exploited in seeking for a cure for VL and other infectious diseases. As an example of the value of the contribution of structural information in later stages in the SGDD process, a total of 57 crystal structures obtained upon soaking of *Tbru*MetRS with multiple compounds and inhibitors served as platform to assist

in the discovery and optimization of new lead compounds against the causative agent of HAT. Promising compounds generated through the utilization of collaborative SGDD strategies as the described in this dissertation should eventually facilitate the development of inhibitors targeting homologous aaRS across the related protozoa and bacteria affecting the lives of the most underprivileged human populations *worldwide*.

Table of Contents

Chapter 1

Introduction

Infectious diseases are caused by a vast spectrum of pathogenic microorganisms, such as bacteria, viruses, parasites and fungi, that can spread, directly or indirectly, from one person to another [1]. HIV, tuberculosis (TB), malaria and tropical neglected diseases (NTDs) are some of the devastating infectious diseases affecting millions of people in poverty in developing countries, who can often suffer from more than one infection at a time [2]. For many of these illnesses there is no cure and for others the available treatments are suboptimal and drug resistance has undermined some highly effective medicines in the past and is a continuing concern. Effective novel treatments against infectious diseases would benefit from the discovery of new drugs for known targets as well as from the identification new drug targets and the subsequent development of drugs against them.

In this work, we have focused on the essential aminoacyl-tRNA synthetases as chemotherapeutic targets in a structure guided drug design (SGDD) effort as a strategy in the fight against some of the pathogenic bacteria and parasites causing millions of deaths *worldwide*.

1.1. The targeted diseases

1.1.1. Tuberculosis

Tuberculosis (TB) is one of the top 10 causes of death globally [3]. Usually affecting the lungs, TB is caused by the bacillus *Mycobacterium tuberculosis* and is frequently fatal if left untreated [4]. In 2016 not less than 10.4 million new TB cases and 1.7 million deaths were reported. Of those cases, children accounted for 1 million new infections and 250,000 deaths [3]. 0.4 million deaths in 2016 were reported among HIV infected individuals, who are especially susceptible to TB [3]. The Bacillus Calmette-Guérin (BCG) vaccine is the only vaccine in use for prevention of TB and confers partial protection especially to young children in countries with high TB incidence [5]. People can live with a latent TB infection and never develop the disease. TB disease occurs when the bacteria become active (i.e. multiplies within the human body). This happens when the hosts' immune system is not able to control the infection. TB disease in the lungs or throat can be infectious. *M. tuberculosis* is spread from one person to another through the air when a person with TB in their lungs or throat coughs, speaks or sings and people nearby breathe the air containing bacteria [4]. In the lungs, alveolar macrophages phagocytize *M. tuberculosis* and the bacterium is able to modulate the immune response promoting its intra-macrophage survival [6]. The formidable cell wall of *M. tuberculosis*, with an unusual mycolyl-arabinogalactan–peptidoglycan (mAGP) complex core [7], contributes to its virulence, immune suppression and prolonged survival [8]. A persistent *M. tuberculosis* infection in the lungs stimulates the generation of organized aggregates of immune cells called granulomas which are pathological hallmarks of TB [9]. The animal model of choice and its generation of granulomas in experimental *M. tuberculosis* infections has shown to be challenging when reproducing the human infection [6]. Moreover, getting

compounds across the highly impermeable cell wall is also of concern when developing chemotherapies against TB [6].

Ethambutol, isoniazid, pyrazinamide and rifampicin are the four first-line drugs recommended for drug-susceptible TB treatment that are to be taken as part of a 6-month regimen [3]. Associated side-effects such as hepatotoxicity, CNS toxicity, exanthema, arthralgia and nausea are frequent causes of discontinuation [10]. Even though millions of lives have been saved through the available TB treatments, multidrug-resistant TB (MDR-TB) is a major global health concern. In the order of 600,000 patients with rifampicin-resistant TB (RR-TB) were reported in 2016, of which 490,000 had MDR-TB [3]. Additionally, some of the current TB drugs were shown to interfere with HIV drugs or other chronic disease medications [6]. Therefore, there is an urgent need for new drug candidates against TB.

1.1.2. Neglected tropical diseases caused by trypanosomatid protozoa

The so-called neglected tropical diseases (NTDs) are caused by a variety of pathogens. Three of the etiological agents of NTDs are the hemoflagelated protozoa *Leishmania* spp., *Trypanosoma brucei* and *Trypanosoma cruzi*. These eukaryotic parasites belong to the kingdom Protista, phylum Sarcomastigophora, sub-phylum Mastigophora, class Zoomastigophora, group Euglenozoa, order Kinetoplastida, suborder Trypanosomatina, family Trypanosomatidae [11].

The order Kinetoplastida is characterized by the presence of a kinetoplast, a structure formed by densely packed mitochondrial DNA present in a single branched mitochondrion, close to the basal body of the flagellum [11, 12]. The generation of functional messenger RNAs (mRNA) of mitochondrial genes involves an RNA editing process that is unique to kinetoplastids [7, 13]. Other outstanding differences with the mammalian hosts make kinetoplastids very attractive to study. Some of these characteristics are: their nuclear genes are transcribed as long multicistronic RNA molecules that are then processed to mRNA by trans-splicing; the apparent absence of RNA polymerase II promoters; and a unique form of metabolic compartmentalization, whose most peculiar aspect is that part of the glycolytic pathway (the first 6-7 depending on the species and stage of the parasite) and other enzymatic systems are found within peroxisome-like organelles called glycosomes [11, 14].

Trypanosomatid parasites go through a series of stages of development during their digenetic life cycle (Figures 2, 3 and 4) in the insect vector and in the human host (among other mammals). The developmental stages adapt rapidly to extremely distinct and challenging environments and are microscopically distinguishable according to the characteristics of the flagellum and the relative position of the kinetoplast with respect to the cell nucleus [15].

Depending on the species, some of the stages within the human host are intracellular and others extracellular, and trypanosomatids have evolved sophisticated ways to avoid the human immune response which complicate greatly the generation of vaccines against them [16]. Surprisingly, like *M. tuberculosis*, *Leishmania* organisms are able to live inside macrophages, which are actually main mediators in the hosts' immune system. *T. brucei* possesses an incredible capacity to evade the host immune system involving a 15 nm thick variable surface glycoprotein (VSG) coat. The VSG coat, formed approximately by 10 million identical molecules, goes through a complete and successive turnover once an antibody response has been mounted. The surviving parasites display a brand-new coat made up from a new set of a next generation of VSGs [17]. *T. cruzi* uses molecular mimicry strategies, hides for many years in various human tissues and evades the human complement system [18-20].

Due to their complicated administration regimens, low efficacy for parasite elimination, poor safety profile, occurrence of drug resistance [21], treatment failure, and the fact that most of the available drugs are not oral, new oral drug candidates are desperately needed to go into clinical development [22].

We focused the structural work for this thesis on tRNA synthetases of *Leishmania donovani* and *Trypanosoma brucei*. However, a successful chemotherapeutic agent targeting one of these kinetoplastid parasites, could potentially translate to its close relatives.

The **leishmaniases** are caused by more than 20 species of *Leishmania* genus parasites [23]. These protozoans are transmitted to humans through the bites of infected female phlebotomine sandflies (Figure 1) [15]. Even though only a fraction of those infected by *Leishmania* parasites develop the disease, from 0.7 to 1.2 million new cases and 20,000 to 30,000 deaths occur annually distributed

in the tropics and sub-tropics where approximately 350 million people are at risk of infection [22, 23].

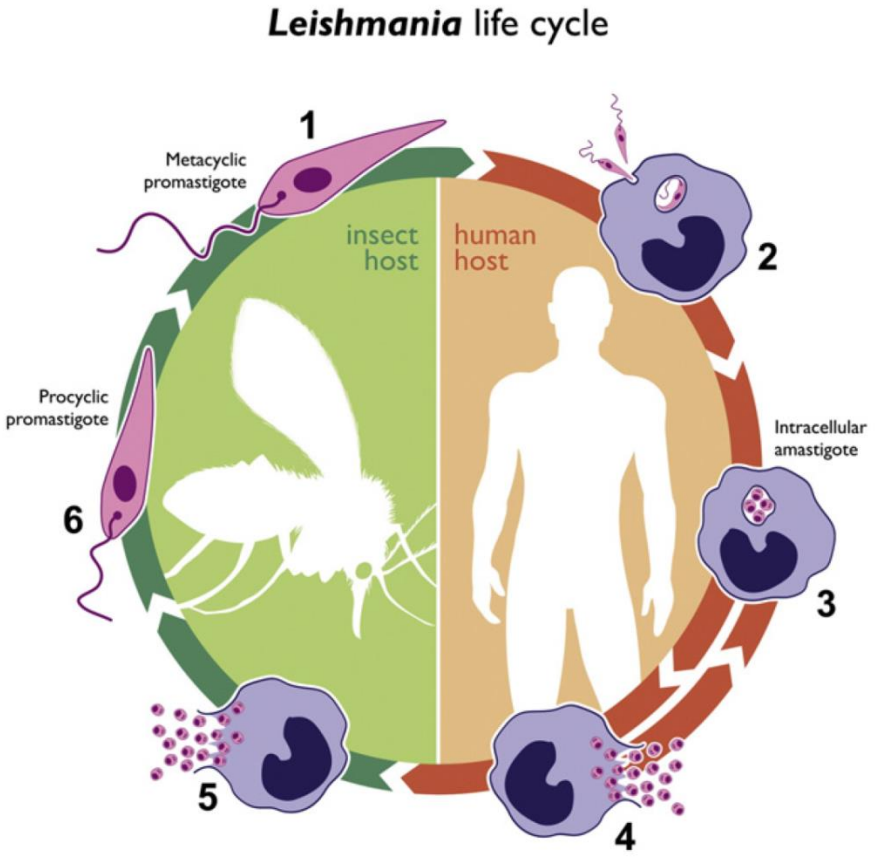


Figure 1. ***Leishmania* life cycle.** Figure from Cuervo *et al.* (2010) [15]. Metacyclic promastigotes (1) are the infective form of *Leishmania* to humans and other mammals. An infected female phlebotomine sandfly can transmit metacyclic promastigotes through its bite during a blood meal. Macrophages phagocyte the metacyclic promastigotes (2) and a phagolysosome forms. In the phagolysosome the parasites transform into amastigotes (3). Amastigotes multiply, lyse macrophages and infect other mononuclear phagocytic cells (4). A sandfly ingests infected macrophages (5) with a blood meal. Amastigotes transform into procyclic promastigotes (6) which is the proliferative form in the insect gut and migrate to the insect proboscis as metacyclic promastigotes (1), continuing the cycle [15].

One of three main forms of leishmaniasis can develop in patients, depending on the species involved. Cutaneous leishmaniasis (CL) (caused by e.g. *L. major* and *L. mexicana*) is the most

common form of leishmaniasis and is characterized by skin lesions, mainly ulcers, leaving life-long scars and serious disability. Visceral leishmaniasis (VL) or kala-azar (*L. donovani* and *L. infantum*) is fatal if left untreated. It is characterized by irregular bouts of fever, weight loss, enlargement of the spleen and liver, and anemia. About 200,000 to 400,000 new cases of VL occur each year and children are the most severely affected group [23, 24]. Mucocutaneous leishmaniasis (e.g. *L. brasiliensis*) leads to partial or total destruction of mucous membranes of the nose, mouth and throat [23, 25]. Available drugs used for the treatment of the leishmaniases are mainly pentavalent antimonial complexes, amphotericin B, the aminoglycoside paromomycin, and the alkylphosphocholine miltefosine, either as single or combination treatments [21, 26] and have shown problems related to efficacy, toxicity and cost [27].

Sleeping sickness or Human African Trypanosomiasis (HAT), is caused by two sub-species of *Trypanosoma brucei* and it is usually fatal if untreated. HAT is endemic in 36 sub-Saharan Africa countries where it is transmitted by the bite of a *Glossina* genus insect, commonly known as the tsetse fly [28]. Sustained control efforts have lowered the number of new cases per year from 300,000 estimated in 1995 to 7,139 cases recorded in 2010 and 2,804 cases in 2015 [28]. However, since the disease is not easily diagnosed and mostly occurs in remote rural areas, the estimated number of actual cases might be approximately 15,000 [28]. The estimated population at risk is 65 million [29].

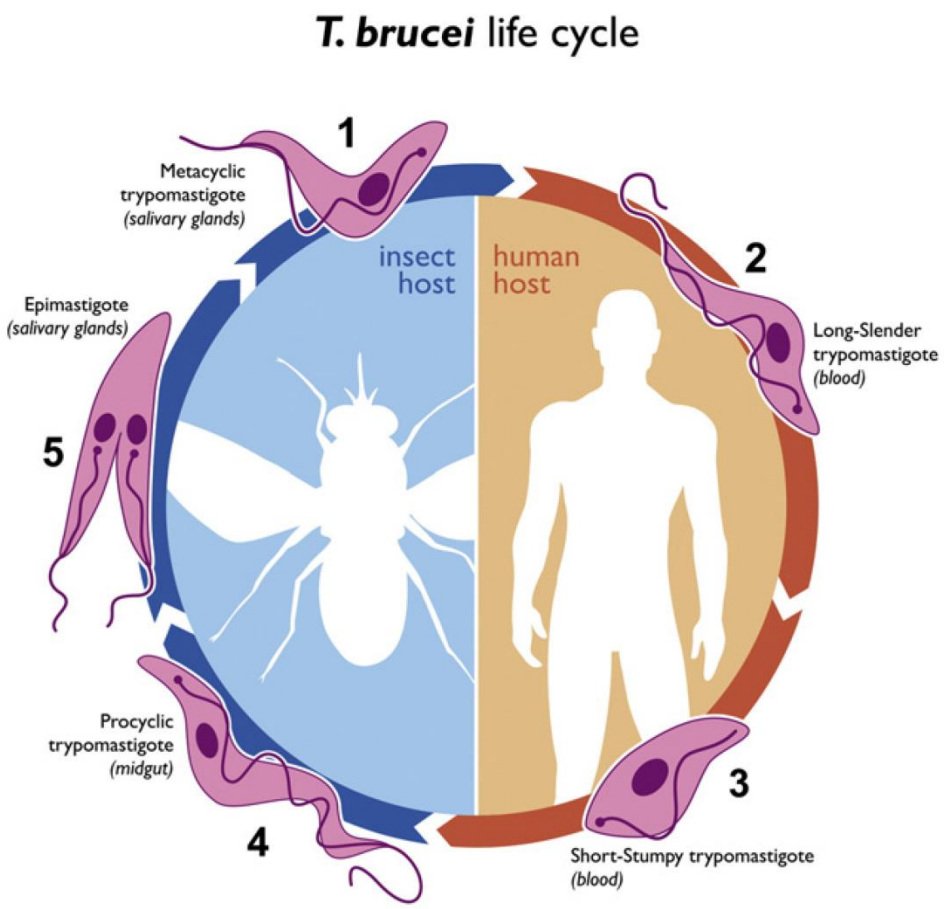


Figure 2. ***Trypanosoma brucei* life cycle.** Figure from Cuervo *et al.* (2010) [15]. An infected tse-tse fly can transmit the infective metacyclic trypomastigotes (1) to a mammalian host through its bite during a blood meal. Within the host, metacyclic trypomastigotes transform into long slender trypomastigotes (2) that multiply in the blood and tissue fluids. Upon differentiation, the parasites transform into non-proliferative short stumpy trypomastigotes (3) that can be taken up by an insect vector with a blood meal. In the insect midgut the parasites differentiate into procyclic trypomastigotes (4) and later to epimastigotes that proliferate in the salivary glands (5). The cycle continues when the parasites differentiate into metacyclic trypomastigotes (1) in the salivary glands and they can be transmitted to the human host in a new blood meal [15].

T. brucei parasites have the remarkable property to be able to invade the central nervous system (CNS) by crossing the blood-brain barrier. The development of the disease depends on the sub-species involved. *Trypanosoma brucei rhodesiense* HAT represents under 2% of reported

cases and causes an acute infection. Once the parasites have reached the brain the disease develops rapidly. *Trypanosoma brucei gambiense* accounts for more than 98% of reported cases of sleeping sickness and causes a chronic infection [28, 30]. The disease is normally asymptomatic until the neurological phase develops upon parasite CNS invasion. The neurological phase is when the most obvious and terrible signs and symptoms of the disease appear, including behavioral changes, confusion, sensory disturbances, lack of coordination and ultimately death. Sleep cycle disorders, which give the disease its name, are an important characteristic of the second stage [28].

Part of the fall in the case load of HAT discussed above is due to the replacement of previously very toxic treatments by the combination therapy Nifurtimox-Eflornithine (NECT), developed in 2009, to treat chronic HAT patients infected with *T. b. gambiense.* However, the patients must be hospitalized and undergo a complex and painful lumbar puncture to determine the stage of the disease before the treatment administration which remains problematic because of difficulties in shipment, storage and administration of the drug [29].

Chagas Disease or American Trypanosomiasis, caused by *Trypanosoma cruzi*, can be a lifelong infection and a life threatening illness if untreated. About 12,000 deaths occur annually and it is estimated that 6-8 million people are currently infected, with 70 million at risk of infection [31]. Endemic in 21 Latin American countries from Mexico to Argentina, Chagas disease has also been increasingly detected in the United States of America, Canada, many European and some Western Pacific countries, primarily due to population mobility [32, 33].

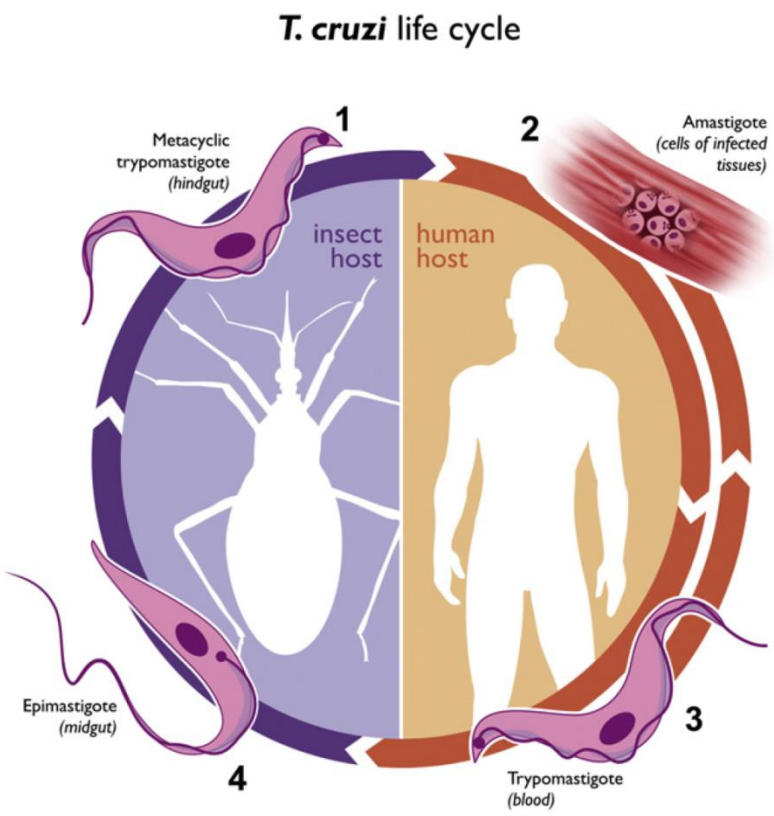


Figure 3. ***Trypanosoma cruzi* life cycle.** Figure from Cuervo *et al.* (2010) [15]. An infected reduviid insect vector can deposit the infective metacyclic trypomastigotes (1) with its feces while feeding from a mammalian host blood. In the host, metacyclic trypomastigotes can infect various cell types including cardiac cells. Within the cells, trypomastigotes transform into amastigotes (2) which are the proliferative forms in the vertebrate host. Upon differentiation, the parasites transform into non-proliferative trypomastigotes (3) that can invade new neighbor cells or be released into the bloodstream where they can continue invasion of new tissues or are taken by an insect vector in a blood meal. In the insect midgut the parasites differentiate into epimastigotes (4) that proliferate and migrate to the hindgut where transform into metacyclic trypomastigotes (5). The cycle continues when the metacyclic trypomastigotes (1) are transmitted by the insect vector with a new blood meal [15].

American trypanosomiasis is transmitted in 90 % of the cases by triatomine bugs (also known as "kissing bugs"). Non-vector-borne transmission includes oral route, blood transfusions, congenital route, organ transplants and accidental transmission in laboratories [32]. Chagas disease

is considered the third most important parasitic disease worldwide in terms of morbidity and mortality, after malaria and schistosomiasis. Three phases are recognized in Chagas disease: a short acute phase and a long-lasting chronic phase, separated by a clinically asymptomatic phase called the indeterminate phase [34]. The acute phase of Chagas disease lasts around 2 months and the symptoms are mild and nonspecific of the disease, which is why it is generally not diagnosed in this stage [32]. The inflammation located in the entrance site of *T. cruzi* is called chagoma. The eye chagoma (Romaña's sign) is one of the specific gateway signs of the disease, but it occurs only in a small percentage of those infected [35, 36]. The indeterminate phase begins about 8-10 weeks after the acute phase, it is characterized by the absence of symptoms and may last several years or indefinitely. During this indeterminate stage, most patients are not aware of being infected with *T. cruzi*. In a third of the infected population, the indeterminate stage develops in the chronic phase in which cardiac complications develop due to cardiomyopathy by heart enlargement and sudden death (due to cardiac arrest), and/or intestinal complications by enlargement of esophagus (megaesophagus) or colon (megacolon) [32, 33].

Current available treatments are more than 40 years old, and while they show good efficacy in the acute phase, they need to be used in long regimens and cause significant side effects [31]. Moreover, Chagas disease treatment efficacy diminishes the longer a person has been infected. New drugs are therefore in particular needed to treat chronic Chagas patients [31, 32].

1.2. Protein biosynthesis and aminoacyl-tRNA synthetases as drug targets

Protein translation has been widely validated as target in the development of drugs against infectious diseases and especially for anti-bacterials, as demonstrated by the action of antibiotics such as aminoglycosides, tetracyclines, macrolides, streptogramins and phenicols [37, 38]. Protein synthesis is essential for trypanosomatid life cycle stages within the human host. The fact that many parasites depend on abundant protein translation by fast growing cells makes them sensitive to disruptions in their translation machinery [39]. Most available antibiotics that target protein synthesis in bacteria bind to the microorganisms' ribosomes. However, other essential molecules involved in pathogen protein biosynthesis can be targeted to develop novel antibiotics, as well as other new anti-infectives. Such compounds would have an advantage in that these would be unlikely to show cross-resistance to other classical antibiotics. One class of such targets is formed by the essential aminoacyl-tRNA synthetases (aaRS) [39, 40].

1.2.1. Aminoacyl-tRNA synthetases

The reaction catalyzed by aaRS is the ATP dependent attachment of a particular amino acid onto the ribose 2'- or 3'-hydroxyl group on the 3' end of its cognate tRNA. The reaction occurs in two steps:

(1) amino acid + ATP → aminoacyl-AMP + PP_i

(2) aminoacyl-AMP + tRNA → aminoacyl-tRNA + AMP

First, ATP and amino acid bind at the active site. The α-carboxylate of the amino acid performs a nucleophilic attack on the α-phosphate of ATP which leads to the formation of an

enzyme-bound aminoacyl-adenylate and an inorganic pyrophosphate leaving group (reaction 1, shown above).

In the second step of the reaction, the 2'- or 3'-OH group of the 3' terminal adenosine of the tRNA (Figure 4) carries out a nucleophilic attack on the α-carbonyl of the aminoacyl-adenylate which results in the 3'-esterification of the tRNA with the amino acid moiety and generation of AMP as a leaving group (reaction 2, shown above), followed by product release [41].

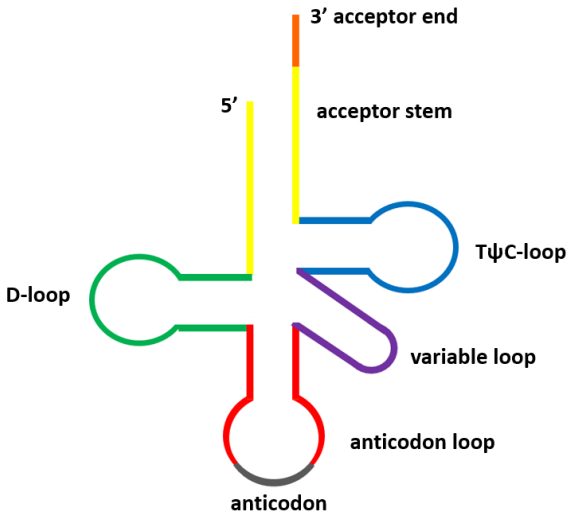


Figure 4. **Schematic structure of a canonical tRNA molecule.** Acceptor stem in yellow, D-loop in green, anticodon loop in red, variable loop in purple, TψC-loop in blue, 3' acceptor end in orange, anticodon in grey.

Based on structural characteristics, the aaRS are divided into two classes (I and II). The active site of class I aaRSs is located within a domain with a Rossmann fold and binds ATP in an extended conformation. The active site of class II aaRSs is formed by an antiparallel β-fold and binds ATP in a bent conformation [41]. Furthermore, class I enzymes approach the acceptor stem of tRNA from the minor groove side, with the tRNA variable loop facing the solvent, attaching the amino acid to the 2'-OH group of the terminal tRNA adenosine (as an exception, 3'-OH is the site

of aminoacylation for CysRS which is a class I aaRS) [42, 43]. In contrast, class II enzymes interact with the tRNA from the major groove side of the acceptor stem, with the variable loop facing the enzyme [43], and attach the amino acid to the 3'-OH group of the terminal tRNA adenosine (with the exception of PheRS whose aminoacylation site is the tRNA terminal adenosine 2'-OH) [42]. Each class can be further divided into three subclasses, designated a, b, and c, on the basis of structural elements, characteristic motifs and sequence homology [40, 44].

1.2.2. Targeting aminoacyl-tRNA synthetases

Target selection is a crucial step in the drug development process. Some of the characteristics of aaRS that make them promising drug targets are:

(1) *Essentiality.* Drug targets need to be essential so that blocking their function is detrimental to the infectious organism [16]. AaRSs are essential for survival. The overall fidelity of protein synthesis is dependent on the accuracy of both codon-anticodon recognition and aminoacyl-tRNA synthesis [41]. AaRS carry out the precise pairing (or charging) of an amino acid with its cognate tRNA. AaRS inhibition blocks the charging of tRNAs leading to interruption or poisoning of the translation process [39].

(2) *Non-redundancy.* There are 20 essential aaRSs, each representing a potential drug target [45]. Human cells have two separate sets of aaRS for cytosolic and mitochondrial aminoacyl-tRNA synthesis for most amino acids [46-48]. In contrast, with the exception of three amino acids (aspartate, lysine and tryptophan), trypanosomatids contain only a single set of aaRS genes. Mitochondrial tRNA genes have apparently been lost on the mitochondrial genome, and the mitochondrial tRNAs are encoded on the nuclear DNA and imported from the cytosol into the mitochondrion [49-51]. The non-redundant nature of most trypanosomatid aaRS is a positive

aspect when considering these enzymes as possible drug targets against kinetoplastids since inhibition of any of the 17 aaRSs functioning in both subcellular locations could have a detrimental effect on protein synthesis in both cytosol and mitochondrion.

(3) *Diversity*. AaRSs display differences among species that have been exploited to design species-specific enzyme inhibitors [40]. In spite of their conserved mechanisms of catalysis, aaRSs are structurally diverse in molecular weight, oligomeric state, primary sequence, and three-dimensional structure [42, 43]. Furthermore, aminoacyl-tRNA synthesis diversity not only resides on the aaRS enzyme. Another level of diversity is shown by the tRNA molecules. For example, cytosolic and mitochondrial $tRNA^{Trp}$ in trypanosomatids (which have two TrpRSs, as stated before) differ due to post-transcriptional modifications that are ultimately relevant for recognition by the aaRS [49]. Even though homologous aaRSs are found in the human host, a detailed study of the differences at the protein sequence and tridimensional structure level between the pathogenic and human aaRSs is extremely important. Those differences could be exploited to arrive to highly selective inhibitors [16]. Depending on the exact pathogen and enzyme, the grade of divergence can vary. In the present work we have focused on pathogenic aaRSs with the largest differences with the human homologs. The targeted aaRSs are ~30-40 % identical to the human homologs in sequence and also differ in residues forming part of essential substrate binding sites.

(4) *Druggability*. AaRS have shown to be "druggable" targets since available structures have revealed exploitable pockets (Figure 5) that are suitable to be used in the design of inhibitors satisfying the Lipinski Rules-of-Five seeking the development of drugs with good oral bioavailability [52]. AaRS inhibitors have been identified from natural sources or by screening of synthetic libraries [40, 53, 54], most of them binding to the ATP or/and amino acid binding sites,

in many cases as analogues of ATP, amino acids, or aminoacyl-adenylate (Aa-AMP) intermediates [55]. Moreover, aaRS inhibitors can block the tRNA recognition, bind to an allosteric site, or to a secondary editing site [56]. The editing site is present in some aaRSs (e.g. LeuRS, IleRS and ValRS) that have an editing domain and are able to hydrolyze a mis-charged aa-tRNA and correct the misacylation [57, 58]. These domains have also been targeted for drug development.

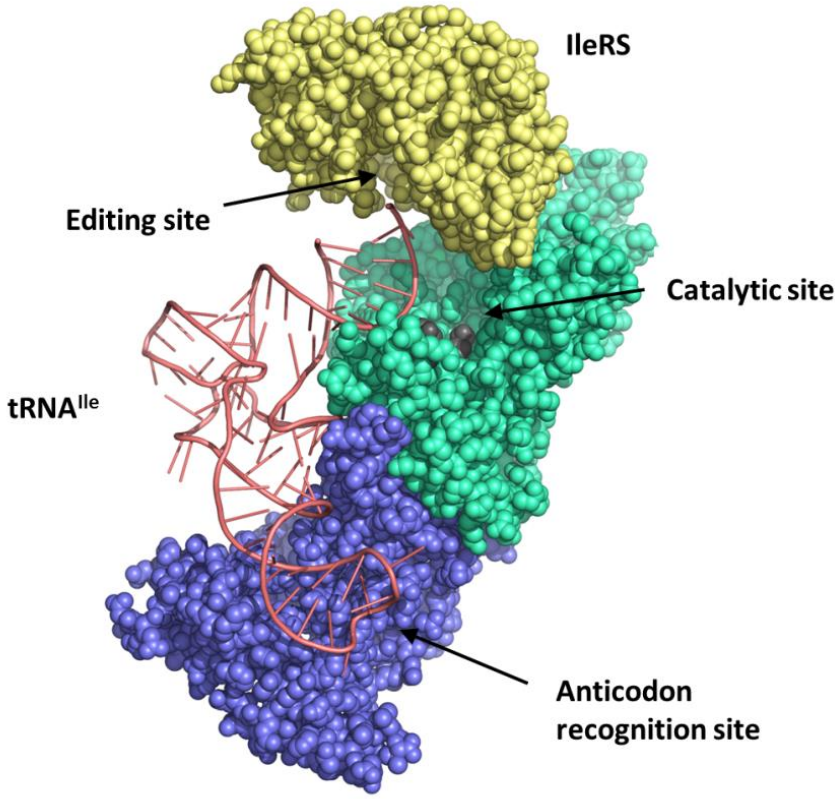


Figure 5. **Schematic representation of *Staphylococcus aureus* IleRS with bound tRNA^{Ile}** (PDB ID: 1FFY [59]). IleRS and mupirocin are shown as spheres and tRNA^{Ile} molecule as sticks. IleRS catalytic domain is shown in cyan, anticodon binding domain in blue and editing domain in yellow. tRNA^{Ile} is shown in red and mupirocin in grey.

(7) *Feasibility*. Successes in targeting bacterial aaRS are encouraging, with at least one topical drug, mupirocin (Bactroban®), in clinical use against *Staphylococcus aureus* infection. Mupirocin is a natural product, originally isolated from *Pseudomonas fluorescens*, and inhibits the isoleucyl-tRNA synthetase (IleRS) of gram positive bacteria including MRSA (Methicillin-resistant *S. aureus*). The structural basis of the recognition is known and it acts by binding both the ATP and Ile binding sites [60]. Also Tavaborole (Kerydin®), an antifungal topical medication for the treatment of onychomycosis that inhibits fungal leucyl-tRNA (LeuRS) by binding to its editing site [58], is in clinical use. Some other anti-microbial aaRS inhibitors are in clinical trials [39]. Compounds inhibiting aaRS of important parasitic protozoa have been shown to interfere with their growth. Several protozoan parasite aaRSs are currently studied as targets. For example, *Plasmodium falciparum* prolyl-tRNA synthetase (ProRS), lysyl-tRNA synthetase (LysRS) and threonyl-tRNA (ThrRS) are inhibited by the natural products halofuginone (Pro and ATP binding sites), cladosporin (ATP binding site) and borrelidin (Thr and tRNA binding sites), respectively [56, 61, 62].

(8) *Availability of structural information.* The first aaRS structure solved at high resolution in 1989 was *Bacillus stearothermophilus* tyrosyl-tRNA (TyrRS) at 2.3 Å [63]. Since then, there is a large and still growing body of structural information that can be used for rational drug design. The Protein Data Bank contains currently about 800 entries of aaRS structures (including apo, compound bound and tRNA bound) from more than 30 species. In fact, for example, the crystal structure of the parasite *P. falciparum* LysRS in complex with cladosporin is available for structure-guided development of compounds [61]. The available aaRS crystal structures in complex with natural and synthetic inhibitors can be useful as starting points to arrive at better compounds in terms of efficacy, specificity, and pharmacological properties [40].

(9) *No cross resistance to other major antibiotics*. Drug resistance can be a terrible occurrence that diminishes effectiveness of drug therapies as exemplified by the anti-malarials chloroquine and pyrimethamine which became ineffective in large regions of the world [16, 64]. A very attractive feature of aaRS inhibitors is that these compounds do not show cross-resistance with other major antibiotics [40], opening the possibility for multicompound strategies that could contribute avoiding resistance development. However, there are some examples of developed resistance against aaRS inhibitors that need consideration. For example, two levels of mupirocin resistance can occur in staphylococci: a high level of resistance (IC_{50} increases 1000 fold) due to the acquisition by the bacteria of a plasmid encoding an exogenous IleRS gene and a low level of resistance (IC_{50} increases 3 fold) due to a chromosomal IleRS mutation [65]. In another example, GSK2251052 (benzoxaborol compound), that binds to the LeuRS editing site and is active against anaerobic and Gram-negative bacteria, was discontinued in phase II studies. Mutations in the editing site occurred rapidly. One of the mutations found in clinical isolates, T247I, lead to a 512 times higher MIC for GSK2251052 in isogenic *E. coli* mutants with respect to that of their parental strains [66].

1.3. Structure guided drug design

Structure guided drug design (SGDD) has been recognized by the pharmaceutical industry as being able to make important contributions to multiple stages of the drug development process, as sketched in Figure 6. SGDD implies that medicinal chemistry for the design of new drugs is done with the guidance of structural details of the target protein [67]. Structural insights of target biomacromolecules have been crucial in the development of several drugs available on the market and of drug candidates in clinical development [68, 69]. We prefer the term SGDD over the commonly used term structure based drug design (SBDD) because other factors also influence greatly the drug design process. Detailed structural information of protein targets is usually obtained through high resolution x-ray crystallography, nuclear magnetic resonance (NMR) and, more recently, by single particle cryo-electron microscopy (Cryo-EM). Such structural information can impact different stages of drug development from target selection to affinity optimization of lead compounds to decreasing drug resistance [68]. Structural knowledge of the ligand-binding pocket of the drug target enables the *de novo* design of initial inhibitor scaffolds and the iterative structure-guided ligand optimization to rationally improve early leads. Also, by means of fragment screening, crystal structures of small molecules (called "fragments") in complex with the protein target can be used to design novel inhibitor scaffolds [68]. SGDD can be also involved in the design of inhibitors that are less susceptible to the development of resistance, either early in the drug development process or, in much later stages, when resistance has already been generated in the field [16, 70]. Importantly, drug development requires the simultaneous optimization of other critical proprieties such as stability, solubility, toxicity, bioavailability, intestinal absorption, tissue distribution, metabolic stability, plasma protein binding and elimination [16, 68].

In the case of parasitic and bacterial infectious diseases additional considerations need to be taken into account in the drug development process since many of these infectious agents are intracellular which means that the drug must cross one or more extra membranes to reach them. Also, when the targeted infectious diseases are prevalent in rural areas where health care centers are often unable to provide complex treatments, oral administration of compounds is essential. In these settings, drug stability in hot and humid conditions can be also crucial to avoid the need of cold chain. Additionally, the costs of the therapy have to be very low so it is available to underprivileged patients [16, 71].

Structural guidance has served in the design of many well-known drugs. These include the neuraminidase inhibitors zanamivir (Relenza®) and oseltamivir phosphate (Tamiflu®) [72], as well as darunavir (Prezista®), a nonpeptidic protease inhibitor taken as part of the anti-HIV treatment [73, 74]. These compounds were developed through a classical SGDD approach. *P. falciparum* dihydro-orotate dehydrogenase (DHODH) is an example in which the structural knowledge served as guidance for the further development of selective and potent inhibitors against the parasitic protozoa. *P. falciparum* DHODH inhibitor DSM265 is in phase I human clinical trials against malaria [75, 76]. In addition, *P. falciparum* dihydrofolate reductase (DHFR) inhibitor P218 in preclinical studies is an example of an effort to overcome resistance that benefited from structural information of the target enzyme [77].

While structural information can be very informative, many successful drugs, including ribosome-targeting anti-bacterials [78] as well as the anti-malarials chloroquine [79] and artemisinin [80, 81], were developed without the use of structural information, and have been critical in the control of major infectious diseases. Moreover, targets of drugs in use for decades can be identified later, and structural information of the target can be then used to make next generation antimicrobials.

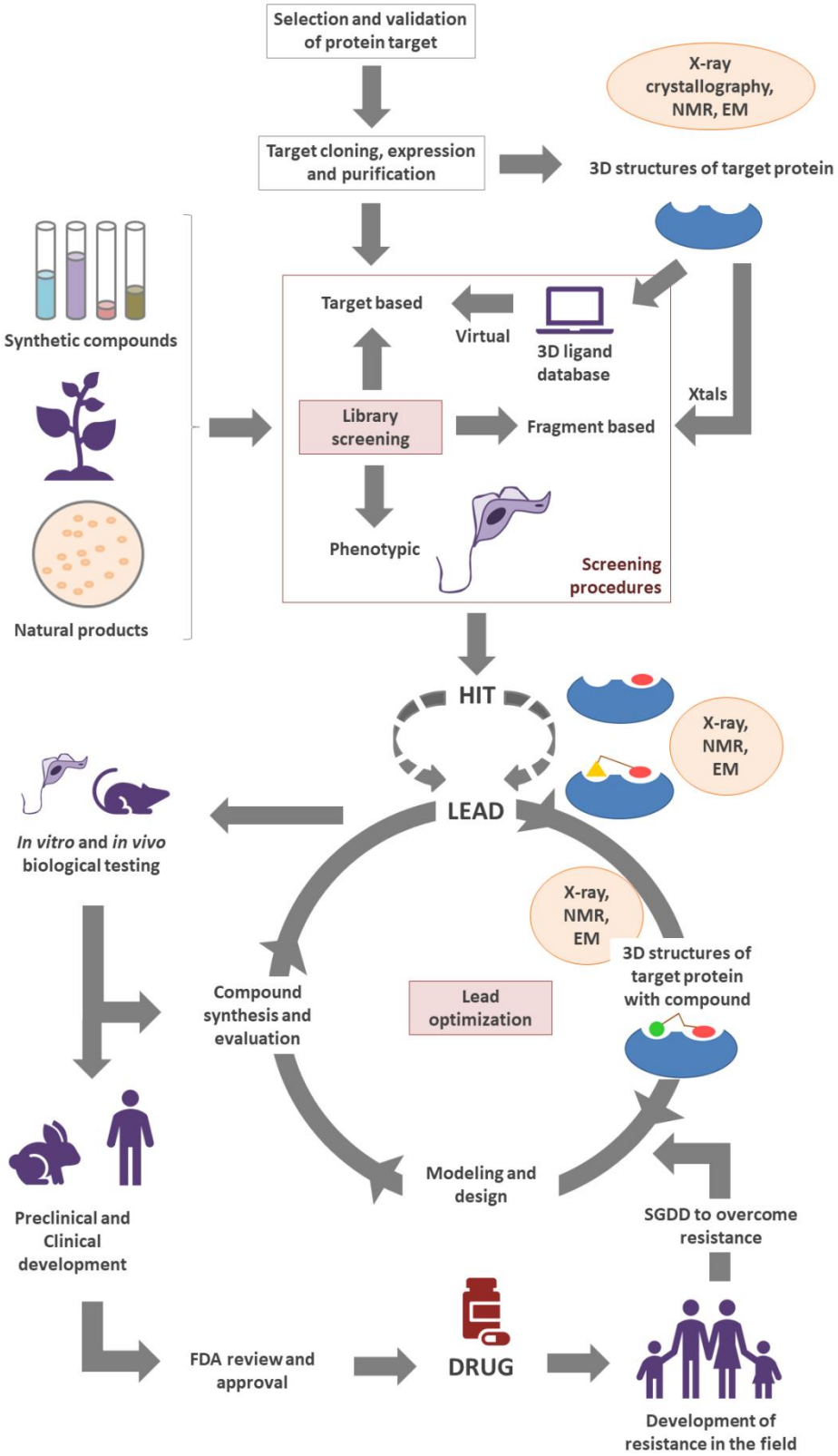


Figure 6. **A schematic representation of ways in which structure guided drug design (SGDD) can contribute to the drug development process**. The classical SGDD pathway starts with the selection and subsequent 3D structure

determination of the protein target and of the protein target in complex with inhibitors that feeds the design of lead compounds (e.g. Influenza virus neuraminidase inhibitors [72]). However, there are other ways in which structures can be very useful in the drug development pipeline. Promising active inhibitors can be discovered in compound library screening efforts against the enzyme of interest and use structural information in the optimization stage (e.g. *P. falciparum* DHODH inhibitor DSM1[82]). The screening can also be virtual when the 3D structure of the target is known (or potentially even predicted) and a 3D ligand database is accessible. If protein crystals (or pure protein) are available, a fragment based screen is an alternative. Phenotypic screening against the infectious agent has shown to be powerful in the discovery of compound hits. Structures of the protein target with hits obtained from the various screening methods can now guide the lead optimization cycle where potency, toxicity and various pharmacological properties are improved until arrival to a promising compound that can move to pre-clinical trials. Verification of the structure of the protein-lead complex is routinely done. After this stage, knowledge of critical protein-ligand interactions can still be useful to direct modifications in the drug candidate to improve pharmacokinetic properties [83]. However, it is also possible to develop a drug with minimal or no input from structural information. After a drug has been approved, structural studies can again become crucial in later stages of drug development when resistance has developed in the field and the need exists to develop new compounds less prone to resistance (e.g. HIV non-nucleoside reverse-transcriptase inhibitors (NNRTIS) [84]; *P. falciparum* DHFR inhibitor P218 [77]).

1.4. The targeted enzymes

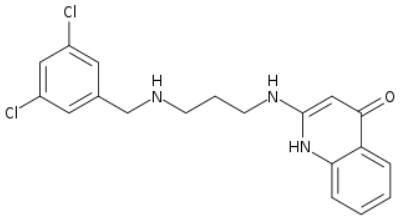
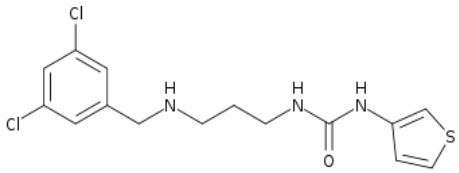
Trypanosomatid tyrosyl-tRNA synthetases (TyrRS) and methionyl-tRNA synthetase (MetRS) and *Mycobacterium tuberculosis* MetRS have been chosen in this work because of sequence differences in the active sites compared with the mammalian orthologs, which holds promise for the development of highly potent yet selective inhibitors.

Compounds targeting bacterial MetRS enzymes were initially developed for treating infections caused by methicillin-resistant *Staphylococcus aureus* and *Clostridium difficile* [85, 86]. Although some of these compounds inhibit the recombinant human mitochondrial MetRS, there is no apparent cytotoxicity of mammalian cell cultures [87]. Structural information of *M. tuberculosis* MetRS (*Mtub*MetRS) could be of considerable benefit to support the discovery of novel anti-tuberculosis inhibitors targeting MetRS (Chapter 2).

In Trypanosomatids, MetRS has been also genetically (by RNAi in *T. brucei*) and chemically validated as drug target[88]. Subsequently, the *L. major* MetRS crystal structure was elucidated and revealed differences near the active site when compared with its human homologs that encourages the development of specific inhibitors against the leishmania parasite [89]. Next, analysis of *T. brucei* MetRS (*Tbru*MetRS) crystal structures in complex with methionine, methionyl adenylate (Met-AMP) and inhibitors concluded that the enzyme undergoes extensive conformational changes upon inhibitor binding by conformational selection [90]. Furthermore, crystal structures of *Tbru*MetRS in complex with aminoquilonone and urea-based inhibitors [90, 91] (Table 1) have guided the design of new anti-trypanosomatid compounds with low nanomolar inhibitory effects on enzyme and on parasites [88, 92, 93]. Promisingly, no resistance mutations were found after culturing *T. brucei* parasites in the presence of *Tbru*MetRS inhibitors. Our collaborators in the group of Dr. Fred Buckner were, however, able to induce resistance in

T. brucei parasites by long-term *in vitro* culture at sub-lethal doses of the compounds in Table 1, caused by overexpression of the target (35 fold for 1433) [94]. The mechanism of resistance by amplification of the *T. brucei metRS* gene appeared to have a fitness cost to the parasites, as shown by the reduced parasitemia produced by these parasites in the mouse model [94].

Table 1. Representative compounds inhibiting *Tbru*MetRS.

Compound series	**Structure**	**Reference**	**Enzyme assay**	***T. brucei* EC50 (nM)**
Aminoquinolone derivatives. (1312 shown)		Compound 1 in Shibata *et al.* (2011) [88]	*Tbru*MetRS ΔTm (°C) 12.9	4
Urea-based inhibitors (1433 shown)		Compound 26 in Shibata *et al.* (2012)[92]	*Tbru*MetRS IC_{50} (nM) 28.0	220

Inhibitors like those presented in Table 1 are also potent against *S. aureus* MetRS (*Saur*MetRS) and other Gram-positive bacteria MetRSs, with minimum inhibitory concentrations (MICs) of less than 0.02 μg/mL [95]. Nevertheless, available inhibitors against both trypanosomatid and bacterial MetRSs need further improvement in diverse pharmacological proprieties. Structure determinations of the target enzymes in complex with newly discovered inhibitors can make important contributions to the iterative design of compounds with better drug-like properties leading to novel candidate drugs against these infectious diseases.

TyrRSs have been successfully targeted both with natural compounds, as SB-219383 [96-99], and with synthetized inhibitors [100-103] in bacteria. However, with respect to parasitic TyrRSs there have been no significant advances in drug development [39]. The crystal structure of *P. falciparum* TyrRS (*Pfal*TyrRS) in complex with tyrosyl-adenylate [104] has been reported. The crystal structure of *L. major* TyrRS (*Lmaj*TyrRS) is available [105] in order to assist additional structure determinations and guide anti-TyrRS drug design to arrive at compounds useful for the treatment of leishmaniasis. The enzyme appeared to be an asymmetric pseudo-dimer formed by one polypeptide chain and its structure differs importantly from the human counterpart. Interestingly, while the human enzyme is a homodimer, the trypanosomatid enzyme contains only one functional active site (contributed by the N-terminal half) and one functional anticodon recognition site (contributed by the C-terminal half). A structure guided drug development project for the treatment of visceral leishmaniasis would greatly benefit from a careful analysis of new leishmanial TyrRSs structures, in particular from *L. donovani* and *L. infantum*. Therefore, a structure determination of *L. donovani* TyrRS (*Ldon*TyrRS), in complex with a nanobody and the catalytic intermediate analog TyrSA was undertaken (Chapter 3).

1.5. Outline of the dissertation

Collaborative projects among various research groups, most of them within University of Washington, in which different procedures including *in silico*, *in vitro* and *in vivo* techniques are constantly being used, have provided a stimulating environment for our iterative SGDD process. The research done as part of this dissertation has focused on providing a structural platform to assist collaborative anti-aaRS drug development against parasitic protozoa and bacteria, solving new protein structures of the essential pathogenic aaRSs, as well as new inhibitor-bound structures, and collaborating in the design of novel inhibitors. It is of high priority to obtain high quality diffracting crystals that can lead to the structure determination of the selected drug targets. Given tremendous difficulties encountered in obtaining crystals of some of these proteins, innovative approaches with the use of crystallization chaperones, ligands and aaRS specific inhibitors, as well as combinations of them, were employed. This dissertation reflects my participation in the different stages (Figure 6) on which structural information of the target biomacromolecule is valuable in the SGDD process. Structural work in this dissertation has contributed to the early steps in the determination of new protein structures, as well as to later in the description of the binding mode of inhibitors that were identified from compound library screening and from previous compound optimization cycles. 3D structures generated as part of this collaborative work have also assisted in the design of new compounds through modelling and chemical intuition.

The research described in the second chapter of this dissertation is an example of an early-on protein target structure determination in the SGDD process. The crystal structure of *Mtub*MetRS in complex with the catalytic intermediate Met-AMP was solved at 2.6 Å resolution [106] and the work is described in in chapter 2. The structure reveals differences with other

MetRSs including the human counterparts that could be of use in the chemotherapeutic development against tuberculosis.

The crystal structure of *Ldon*TyrRS was determined at 2.75 Å resolution [107] and the work is described in the third chapter of this dissertation. Well diffracting crystals of this enzyme were only obtained in complex with nanobodies as crystallization chaperones and with the tyrosyl adenylate analog TyrSA. The structure revealed the presence of an extra pocket that seems to be shared with other pathogens but is not present in the human counterparts and could be used in the design of specific compounds for the treatment of visceral leishmaniasis and other infectious diseases. The research described in chapter 3 is an example of both an early-on protein target structure determination and an analysis of the potentially exploitable structural differences with the human enzymes which could lead to modeling and design of selective compounds against various pathogenic parasites.

A total of 57 crystal structures obtained upon soaking of *Tbru*MetRS with multiple compounds were solved by the current writer as part of a collaborative drug design project. The research described in chapter 4 is an example of later stages in the SGDD process where analysis of crystal structures served in the description of a compound binding mode [108] that was identified from a compound library screening and in the optimization and assistance in the design of new lead compounds [109, 110]. The compilation of the manuscripts that were published as products of these collaborative efforts is presented in the fourth chapter of this dissertation. The list of *Tbru*MetRS•Inhibitor crystal structures that were solved as part of this work but not included in publications is presented in the appendix A.

Chapter 2

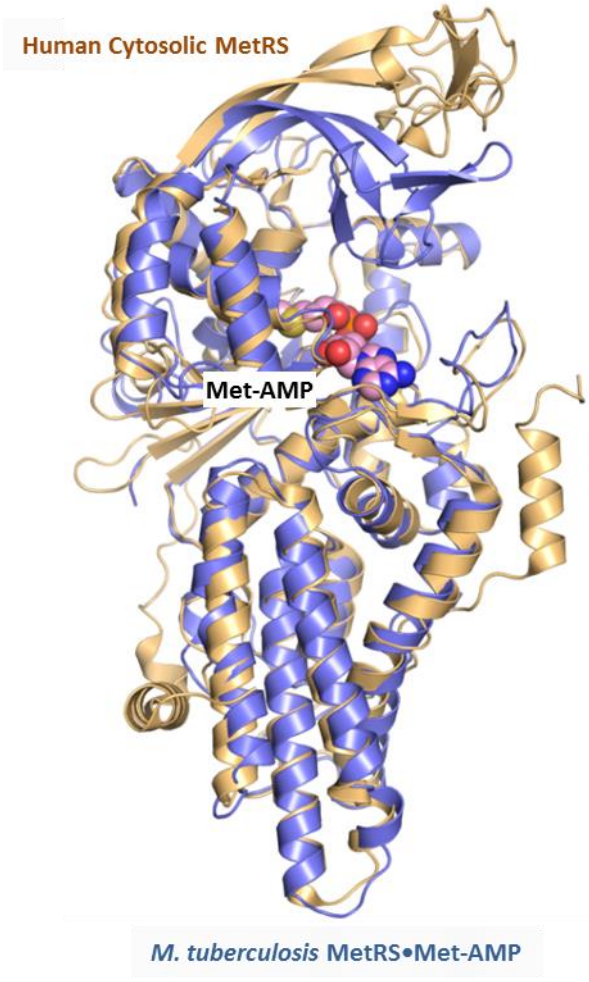
The crystal structure of the drug target *Mycobacterium tuberculosis* methionyl-tRNA synthetase in complex with a catalytic intermediate

A modified version of this chapter was previously published in Barros-Álvarez *et al*., 2018 [106].

ABSTRACT

Mycobacterium tuberculosis is a pathogenic bacterial infectious agent responsible for approximately 1.5 million human deaths annually. Current treatment requires a long-term administration of multiple medicines with substantial side effects. Lack of compliance, together with other factors, has resulted in a worrisome increase of resistance. New treatment options are therefore urgently needed. Here we report the crystal structure of methionyl-tRNA synthetase (MetRS), an enzyme critical for protein biosynthesis and therefore a drug target, in complex with its catalytic intermediate methionyl-adenylate. Phenylalanine 292 from the *M. tuberculosis* enzyme is in an "out" conformation and barely contacting the adenine ring, in contrast to other MetRS structures where ring stacking occurs between the adenine and a protein side chain ring in the "in" conformation. A comparison with human cytosolic MetRS reveals substantial differences in the active site as well as regarding the position of the connective peptide sub-domain 1 (CP1) near the active site which bodes well for arriving at selective inhibitors. Comparison with the human mitochondrial enzyme at the amino acid sequence level suggests that arriving at inhibitors with higher affinity for the mycobacterial enzyme than for the mitochondrial enzyme might be achievable.

Graphical abstract



2.1. Introduction

Caused by the bacillus *Mycobacterium tuberculosis*, tuberculosis (TB) is an infectious disease that typically affects the lungs (pulmonary TB) and is fatal if untreated [4]. TB is one of the top 10 causes of death worldwide [3]. In 2016, there were an estimated 10.4 million new TB cases. Of those cases, approximately 1 million were children. People infected with HIV, as well as other immunocompromised individuals, are especially susceptible to TB, and accounted for 0.4 million deaths in 2016. Globally, 1.7 million people died from tuberculosis in 2016 [3]. The recommended treatment for drug-susceptible TB is a 6-month regimen starting with two months of four first-line drugs: ethambutol, isoniazid, pyrazinamide and rifampicin [3]. The treatment is associated with significant side-effects such as hepatotoxicity, CNS toxicity, exanthema, arthralgia and nausea, important factors in discontinuation [10]. Although TB treatment has saved millions of lives, the crisis of multidrug-resistant TB (MDR-TB) treatment is a major global health concern. In 2016 there were an estimated 600,000 patients with rifampicin-resistant TB (RR-TB) of which 490,000 had MDR-TB [3]. On top of this, there are drug-drug interactions between anti-HIV drugs or other chronic disease medications with some of the current TB drugs [6]. Hence, there is an urgent need for new drug candidates to fight the TB scourge.

As demonstrated by the action of aminoglycosides, tetracyclines, macrolides, streptogramins and phenicols, interfering with bacterial protein synthesis has been shown to be extremely effective for the development of antibiotics [37, 38].

Aminoacyl-tRNA synthetases (aaRSs) are a major group of proteins essential for protein synthesis. AaRSs catalyze a two-step esterification reaction in which each amino acid is paired with its cognate tRNA forming aminoacyl-tRNA [41, 111]:

$$E + aa + ATP \leftrightarrows E{\bullet}aa{\sim}AMP + PP_i \quad (1)$$

$$E{\bullet}aa{\sim}AMP + tRNA \leftrightarrows E + aa\text{-}tRNA + AMP \quad (2)$$

In the first ATP-dependent step a highly reactive aminoacyl adenylate (aa~AMP) intermediate is formed and an inorganic pyrophosphate released. In the second step the amino acid moiety is transferred from the activated intermediate to the 3'-end of the tRNA, yielding aminoacyl-tRNA as the product and AMP as the leaving group [41].

AaRSs have been genetically and chemically validated as antimicrobial drug targets [45, 88, 112]. Anti-aaRS compounds can block the ATP and/or amino acid binding site, the tRNA recognition site, an allosteric site, or an editing site [56]. The widely used antibiotic mupirocin is in clinical topical use against *Staphylococcus aureus* infections and binds both the ATP and Ile binding sites of *S. aureus* isoleucyl-tRNA synthetase (IleRS) [60]. However, *M. tuberculosis* IleRS is not inhibited by mupirocin and therefore *M. tuberculosis* is naturally resistant to this antibiotic [113].

Methionyl-tRNA synthetase (MetRS) is particularly interesting since it recognizes both the initiator $tRNA_f^{Met}$ and the elongator $tRNA^{Met}$ delivering methionine during translation. Depending on their structural features and conserved motifs, aaRSs can be separated into two classes [114]. MetRS belongs to the class I aaRSs, characterized by the amino acid sequence motifs HIGH and KMSKS and a classical Rossmann fold as part of the catalytic domain. Together with arginyl-,

cysteinyl- isoleucyl-, leucyl- and valyl- tRNA synthetases, MetRS belongs to subclass Ia with the HIGH and KMSKS motifs corresponding to ^{18}HVGH^{21} and ^{299}KMSKS^{303} in *M. tuberculosis* MetRS (*Mtub*MetRS). MetRSs can be categorized in two major forms, MetRS1 and MetRS2, depending on structural features and susceptibility to inhibitors. Structurally, MetRS1 orthologs contain one "knuckle" motif with either one or no Zn^{2+} ion bound. MetRS2 enzymes contain two knuckle motifs with either one or two Zn^{2+} ions bound. Most Gram-positive bacteria genera harbor the MetRS1 form, while most Gram-negative genera contain the MetRS2 form [87, 115]. A few bacteria harbor both forms [115, 116]. Eukaryotic organisms contain both forms, with the mitochondrial one exhibiting MetRS1 features, while the cytosolic enzyme being of the MetRS2 form [87].

Highly effective compounds inhibiting MetRSs from pathogenic trypanosomatid protozoa have been reported by our group [88, 92, 117]. These reports focused mainly on MetRS from *Trypanosoma brucei*, the causative agent of Human African trypanosomiasis (HAT), also known as sleeping sickness. Crystal structures of *T. brucei* MetRS synthetase (*Tbru*MetRS) in complex with various inhibitors [90, 91, 93, 109, 110] have guided the design of new anti-trypanosomatid compounds with low nanomolar inhibitory potency on enzyme and on parasites, and a steadily improved pharmacokinetic profile. This series of inhibitors has also been optimized and tested for their effect on *Staphylococcus aureus* MetRS (*Saur*MetRS) and other Gram-positive bacteria, with minimum inhibitory concentrations (MICs) for some compounds of less than 0.02 μg/mL [95].

A major structural biology effort has been made by many investigators to create a platform of crystal structures of key *M. tuberculosis* proteins to assist in the development of new therapeutic agents to treat tuberculosis patients (e.g. [118-120]. Here we describe the crystal structure of *M. tuberculosis* MetRS (*Mtub*MetRS) at 2.6 Å resolution in complex with the intermediate

methionyl-adenylate (Met-AMP). Various structural features are compared with those of other MetRSs with known structure. Striking conformational changes near the active site are evident for both the KMSKS loop and the connective peptide sub-domain 1 (CP1) (Table 2 provides information on the MetRS structures used for comparisons in this paper). In the Met-AMP binding region of the human cytosolic and mitochondrial enzymes several amino acid differences occur compared to *M. tuberculosis* MetRS, indicating that there are opportunities to arrive at inhibitors which have a higher affinity for the *M. tuberculosis* enzyme than for the human homologs.

Table 2. Structures of methionyl -tRNA synthetases referred to in this study.

PDB ID	Organism	Ligand(s)	Crystallized MetRS	Reference
6AX8	*Mycobacterium tuberculosis*	Met-AMP	Full length	This publication
2X1M	*Mycobacterium smegmatis*	L-Methionine	Full length	[121]
2X1L	*Mycobacterium smegmatis*	L-Methionine and adenosine	Full length	[121]
2CT8	*Aquifex aeolicus*	Met-AMP and $tRNA^{Met}$	Full length	[122]
4EG3	*Trypanosoma brucei*	Met-AMP	237–773	[90]
5GL7	*Homo sapiens* Cytosolic	None	221-834	(H. Y. Cho, H. J. Lee & B. S. Kang (unpublished work))
3KFL	*Leishmania major*	Met-AMP and PPi	206-747	[89]

2.2. Materials and methods

2.2.1. *Mtub*MetRS expression and purification

*Mtub*MetRS was cloned into the AVA0421 vector [123-125] and expressed in *Escherichia coli* for subsequent purification. The protein was purified through a Ni-NTA affinity column (Qiagen, Valencia, CA) followed by size-exclusion chromatography (SEC) on a Superdex 75 column (Amersham Pharmacia Biotech) using SEC buffer (20 mM HEPES at pH 7.5, 500 mM NaCl, 5% glycerol and 2 mM DTT). *Mtub*MetRS purity was assessed through SDS-PAGE and the protein concentration determined using the Bio-Rad protein assay dye reagent (Bio-Rad), based on the Bradford method [126]. A final yield of 1.3 mg of pure *Mtub*MetRS per liter of *E. coli* culture was obtained and concentrated to about 8 mg/mL for crystallization.

2.2.2. *Mtub*MetRS crystallization

Purified *Mtub*MetRS was prepared for crystallization with the addition of 10 mM MgATP, 10 mM methionine and 1 mM TCEP (protein solution). The crystals were obtained after one month at room temperature by vapor diffusion using sitting drops equilibrated against a reservoir containing 30-34 % PEG 8,000, 150-200 mM ammonium sulfate and 100 mM sodium cacodylate pH 6.5. The drops contained 1 µL of *Mtub*MetRS protein solution and 1 µL of reservoir solution. After growth, crystals were picked up by a loop [127], flash frozen in liquid nitrogen [128] in cryo-solution (30% glycerol in reservoir solution) and stored until data collection.

2.2.3. Data Collection and Structure Determination

Data were collected under cryogenic conditions at the Stanford Synchrotron Radiation Lightsource (SSRL) using beamline 12-2 at a wavelength of 0.98 Å. Many crystals exhibited a considerable

degree of twinning. Only a single crystal was not twinned and diffracted to higher resolution than all others. This crystal allowed the eventual structure determination. HKL2000 [129] was used for data processing. The structure of *M. smegmatis* MetRS ([121]; PDB: 2X1L) was used as a model to obtain initial phases by molecular replacement using Phaser [130]. Iterations of manual building and rebuilding using Coot [131] were alternated with refinement of the structure with REFMAC5 [132]. MolProbity [133] was used throughout the process for structure validation. Data collection and crystallographic refinement statistics are given in Table 3. Figures were created with Pymol [134]. Coordinates and structure factors of the *Mtub*MetRS•Met-AMP complex have been deposited in the Protein Data Bank under the PDB ID: 6AX8.

Table 3. Crystallographic data collection and refinement statistics*.

PDB ID	**6AX8**
Data collection	
Space group	H_3
Cell dimensions	
a, *b*, *c* (Å)	196.96, 196.96, 39.18
Resolution (Å)	38.18 – 2.60 (2.72 – 2.60)
R_{merge}	0.198 (1.085)
R_{pim}	0.081 (0.526)
Observed reflections	91700 (11269)
Unique reflections	17391 (2110)
Mean $I / \sigma I$	7.3 (2.3)
Multiplicity	5.3 (5.3)
Wilson B factor	33.9 Å^2
Completeness (%)	99.9 (99.9)
$CC_{1/2}$	0.989 (0.709)
Refinement	
Resolution (Å)	38.18 - 2.60
Reflections used	16441
R_{work} / R_{free}	0.237 / 0.246
Number of atoms	
Protein	3980
Met-AMP	31
Water	38
Number of residues	509
Average *B*-factors (Å^2)	
All atoms	60.7
Protein	60.5
Met-AMP	95.6
Water	51.8
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.29
Ramachandran plot[#]	
Favored (%)	97
Outlier (%)	0
Met-AMP	
LLDF[@]	1.13
RSR[^]	0.24

*Values in parentheses are for the highest-resolution shell

[#]Ramachandran Plot statistics as reported by the wwPDB validation report

[@]Local ligand density fit as reported by the wwPDB validation report

[^]Real space R value as reported by the wwPDB validation report

2.3. Results and discussion

2.3.1. The structure of the *Mycobacterium tuberculosis* MetRS•Met-AMP binary complex

We solved the structure of *Mycobacterium tuberculosis* MetRS (*Mtub*MetRS), a 58 kDa monomeric enzyme, in complex with its intermediate product methionyl-adenylate (Met-AMP). Despite major efforts, no co-crystals of *Mtub*MetRS in complex with our available MetRS inhibitors could be obtained. The purified enzyme was incubated with its substrates methionine and ATP prior to setting up crystallization trays. The electron density map revealed the presence of Met-AMP in the active site of the enzyme. The *Mtub*Met*RS*•Met-AMP complex crystal structure, where the symbol "•" indicates a non-covalent complex, was determined at 2.6 Å resolution (Table 3). There is one molecule in the asymmetric unit (ASU) and, except for the final 10 residues at the C-terminus, every residue could be built in well-defined electron density.

Several structural elements can be distinguished in *Mtub*MetRS (Figure 7): (1) a Rossmann fold catalytic domain (CD) with (2) an inserted connective peptide domain (CP); (3) a stem-contact fold domain (SCF) containing the conserved KMSKS motif; and (4) the anticodon binding domain (ABD). The SCF and ABD domains are connected by a linking π1-helix between helices α14 and α15.

The *Mtub*MetRS CD is predominantly an α/β domain that extends from the N-terminal end of the protein up to residue His290 and is interrupted by the CP that comprises residues Ile116 to Tyr226. The CP domain comprises two sub-domains: sub-domain CP1 formed by the anti-parallel β-strands β4 and β8, and sub-domain CP2, formed by helices α5 and α6. Sub-domain CP1 in *Mtub*MetRS, is in the closed conformation as seen previously in the *Msme*MetRS and other MetRS structures [121]. CP1 harbors one "knuckle" devoid of residues able to coordinate Zn^{2+} which identifies *Mtub*MetRS as a member of the MetRS1 form [87, 115].

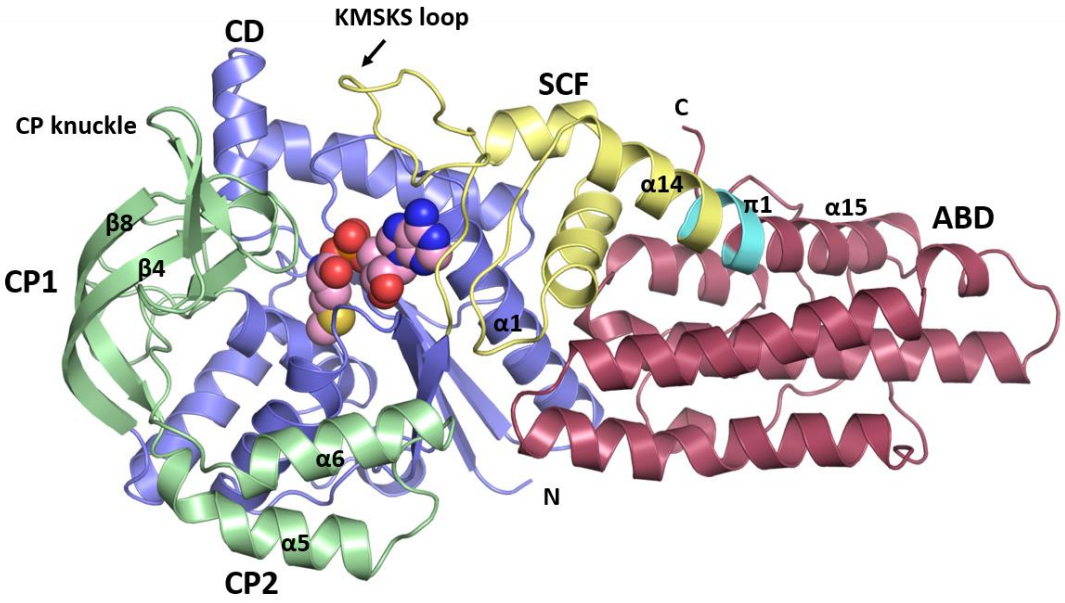


Figure 7. **Domain organization of *Mtub*Met*RS*.** The Rossmann fold catalytic domain (CD) is shown in blue with the connective peptide domain (CP) in green. CP sub-domains CP1 and CP2 are labeled. The stem-contact fold domain (SCF) containing the KMSKS loop is shown in yellow. In cyan is depicted the π-helix connecting the SCF and the anticodon binding domain (ABD). The ABD is depicted in red. Methionyl-adenylate (Met-AMP) is shown as spheres with carbon atoms in pink, nitrogens blue, oxygens red, phosphorus orange and sulfur yellow.

The SCF domain connects the CD and the ABD domains. The SCF domain harbors the KMSKS loop which contains, using *M. tuberculosis* sequence numbers, the conserved motif ^{299}KMSKS^{303}. Interestingly, when superimposing the *Mtub*Met*RS* structure with the structure of the homologous *M. smegmatis* MetRS (*Msme*MetRS) which share 76 % overall sequence identity and 94 % identity in the Met-AMP binding site, the KMSKS loop adopts dramatically different conformations in these two closely related tRNA synthetases. The KMSKS loop in *Msme*MetRS, in the presence of either methionine (PDB ID: 2X1M [121]), or of methionine and adenosine (PDB ID: 2X1L [121]), is in a wide open conformation (Figure 8). In contrast, the same loop in *Mtub*Met*RS* is rotated by ~90° adopting a conformation which brings the motif in proximity to the adenosine

ring of Met-AMP. This is an intermediate position when compared to the structure of *Leishmania major* MetRS (*Lmaj*MetRS) in complex with Met-AMP and pyrophosphate (PDB ID: 3KFL [89]) where the KMSKS loop adopts a closed conformation towards the active site as a result of a ~30° movement of the loop further towards the Met-AMP binding pocket. The *Hsap*CytoMetRS loop adopts essentially the same conformation as the one in *Mtub*MetRS.

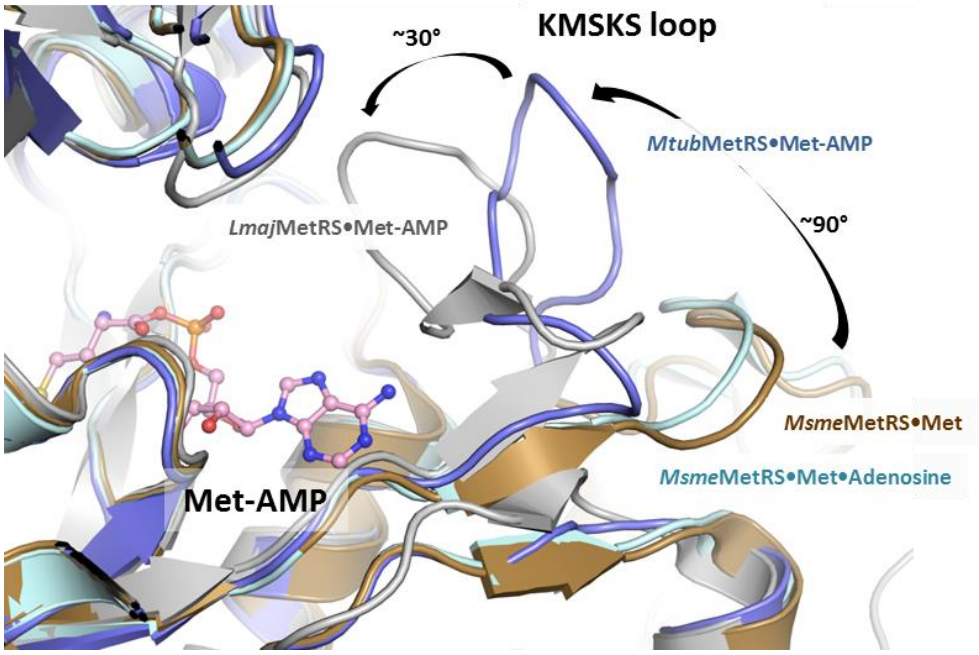


Figure 8. **Variability of KMSKS loop conformations among MetRS structures.** The *Mtub*MetRS•Met-AMP structure shown in blue, that of *Lmaj*MetRS•Met-AMP•PPi (PDB ID: 3KFL [89]) in grey, *Msme*MetRS•Met (PDB ID: 2X1M [121]) in brown and *Msme*MetRS•Met•Adenosine (PDB ID: 2X1L [121]) in cyan. Methionyl-adenylate (Met-AMP) bound to *Mtub*MetRS is depicted, with carbon atoms in pink, nitrogens blue, oxygens red, phosphorus orange and sulfur yellow.

Based on the various positions of the KMSKS loop in these structures it is not easy to find a correlation between ligands bound and KMSKS loop conformation. For instance, the presence of Met-AMP in both the *L. major* and *M. tuberculosis* enzymes resulted in intermediate and closed KMSKS loop conformations, respectively. Amino acid sequence differences and crystallization

conditions might obviously also play a role in determining which conformation the flexible KMSKS loop adopts.

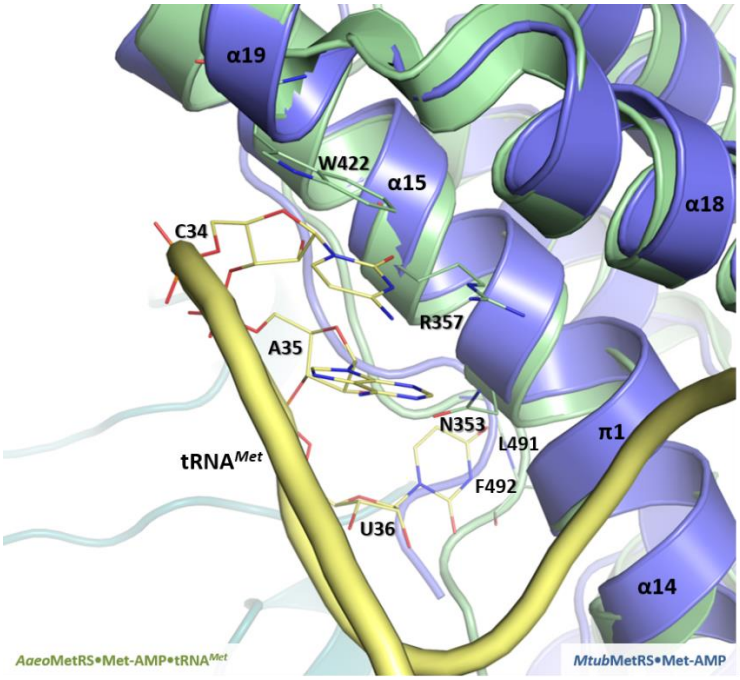


Figure 9. **Comparison of anti-codon binding domains (ABDs) of *Aquifex aeolicus* MetRS (*Aaeo*MetRS) and *Mtub*MetRS.** *Mtub*MetRS•Met-AMP in blue, *Aaeo*MetRS•Met-AMP•tRNA^{Met} (PDB ID: 2CT8 [122]) in light green (protein) and yellow (tRNA). Residues involved in *Aaeo*MetRS binding to tRNA^{Met} and the anticodon triplet ^{34}CAU^{36} are shown in sticks and labeled. Carbon atoms are colored green (*Aaeo*MetRS) or yellow (tRNA^{Met}), nitrogens blue and oxygens red. By homology, the residues expected to bind tRNA^{Met} in *Mtub*MetRS would be N357 in helix π1, R361 in helix α15, W431 in helix α19 and residues V506 and F507at *Mtub*MetRS C-terminal end. The labels of secondary structure elements are those of the *Mtub*MetRS•Met-AMP crystal structure.

Regarding tRNA^{Met} binding, the mainly helical ABD of *Mtub*MetRS superimposes with an r.m.s.d of 1.00 Å for the C$^{\alpha}$ atoms onto the equivalent domain of *Aquifex aeolicus* MetRS (*Aaeo*MetRS) in complex with tRNA^{Met} (PDB ID: 2CT8 [122]). Four out of the five residues, N353, R357, W422, L491 and F492, involved in the recognition of the anticodon triplet ^{34}CAU^{36} of tRNA^{Met} in *Aaeo*MetRS are conserved in *Mtub*MetRS (N357, R361, W431, V506 and F507,

respectively), while L491 of the mycobacterial enzyme is very similar to V506 of *Aaeo*MetRS. Therefore, $tRNA^{Met}$ anti-codon recognition by *Mtub*MetRS is expected to be similar as by *Aaeo*MetRS (Figure 9).

2.3.2. Interactions of *Mtub*MetRS with Met-AMP

Methionine and ATP were essential for obtaining *Mtub*MetRS crystals. Their reaction led to the formation of Met-AMP which was bound to the enzyme's active site, as seen in the electron density map (Figure 10 A). Two pockets can be defined in the active site: the methionine binding pocket (MBP), where the methionyl group of Met-AMP is situated, and the adenine binding pocket (ABP) where the adenine moiety of Met-AMP binds (Figure 10 B).

Figure 11 summarizes *Mtub*MetRS residues binding to the methionyl, phosphoryl, ribosyl and adenylyl groups of Met-AMP along with a sequence alignment with MetRS from other species mentioned in the paper. Most of the interactions between the Met-AMP methionyl group and the enzyme in the MBP are hydrophobic and are mediated by residues A9, W228, A231, L232, Y235, I264 and H268. The nitrogen of the methionyl moiety makes hydrogen bonds with the carbonyl of I10 and with the side chain carboxylate of D49. One oxygen atom of the phosphoryl group of Met-AMP forms hydrogen bonds with the main chain amino group of residue Y12 as well as with the imidazole ring of H21. *Mtub*MetRS residues forming hydrogen bonding with the ribose ring are E24, G261 and D263. Specifically, the 3' hydroxyl of the ribose is engaged in hydrogen bonds with the carboxylate of E24 and with the main chain amino group of G261. At the same time, the 2' hydroxyl hydrogen bonds with the carboxylate of D263. In the ABP, the Met-AMP adenine moiety atom N_7 interacts with the imidazole ring of residue H18, while atoms N_1 and N_6 are engaged in hydrogen bonds with the L293 main chain carbonyl and amino groups, respectively.

Additionally, residues G20 and F292 make hydrophobic contacts with the Met-AMP adenine ring in the ABP.

In summary, extensive interactions, both hydrophobic and hydrophilic (Figure 10 C), mediate the *Mtub*MetRS•Met-AMP complex formation with a buried surface area of 919 $Å^2$ as calculated by PISA [135].

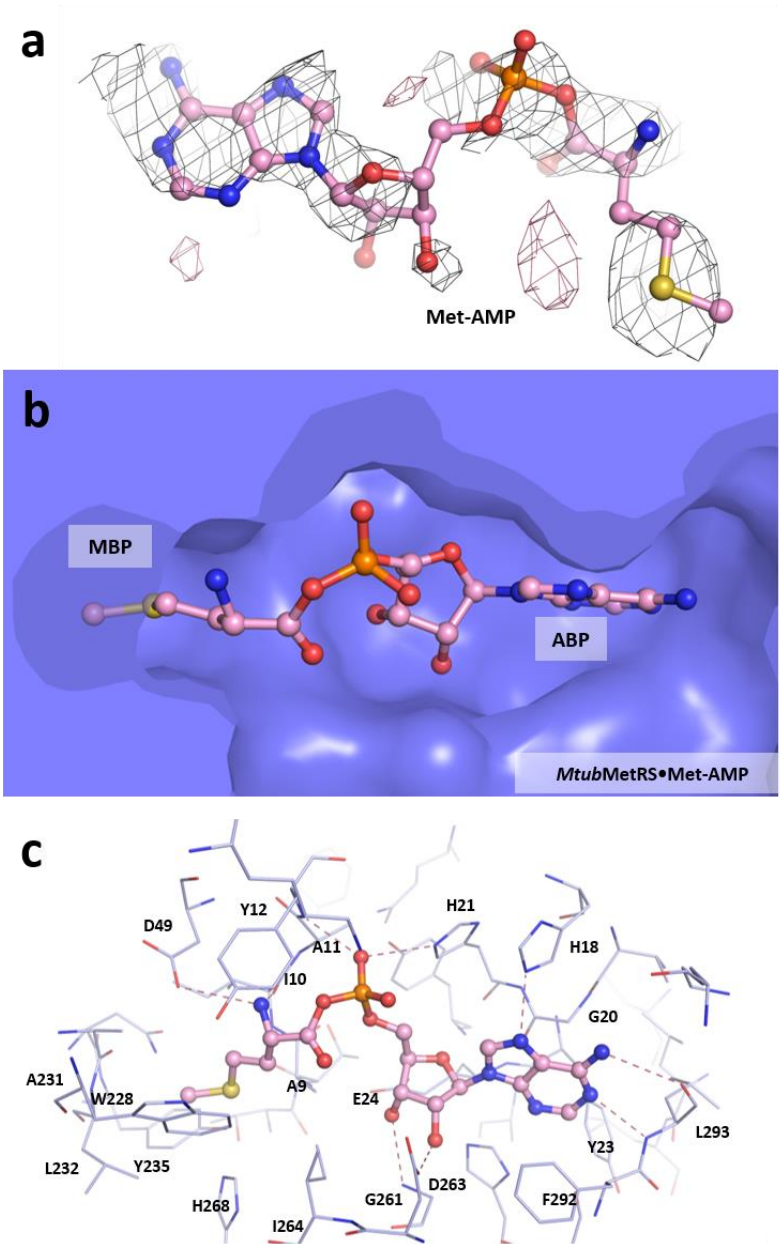


Figure 10. **Met-AMP binding to *Mtub*MetRS.** A) Methionyl-adenylate (Met-AMP) difference electron density map calculated by omitting the molecule, contoured at 2σ (positive density in grey, negative density in red). B) General features of the Met-AMP binding mode. Shown are the protein surface (in light blue), bound Met-AMP in sticks, and the two pockets, the methionine binding pocket (MBP) and the adenine binding pocket (ABP). C) Hydrogen bond network in the *Mtub*MetRS / Met-AMP interaction is shown in dotted red lines. *Mtub*MetRS residues interacting with Met-AMP are labeled. Met-AMP carbon atoms are shown in pink, nitrogens blue, oxygens red, phosphorus orange and sulfur yellow.

Res. No.	9	10	11	12	18	20	21	24	49	51	52	53	54	211	214	215	224	225	226	227	228	230	231	232	234	235	261	263	264	267	268	292	293	OI%	MI%	II%
Met-AMP binding pocket	MBP	MBP		MBP P	ABP	ABP	P	R	MBP												MBP		MBP	MBP		MBP	R	R	MBP		MBP	ABP	ABP			
Inhibitor binding pocket	EMP	EMP	EMP	LR					AP	AP	AP	AP	AP	AP	AP	AP	AP	AP	AP	AP	EMP	AP	EMP	EMP	EMP	EMP			EMP	EMP	EMP					
M.tuberculosis	A	I	A	Y	H	G	H	E	D	H	G	L	K	F	G	V	V	M	Y	V	W	D	A	L	N	Y	G	D	I	F	H	F	L	N/A	N/A	N/A
M.smegmatis	A	I	A	Y	H	G	H	E	D	H	G	Q	K	F	G	V	V	M	Y	V	W	D	A	L	N	Y	G	D	I	F	H	W	L	76	94	96
B.melitensis	A	I	A	Y	H	G	H	E	D	H	G	I	K	F	G	I	V	M	Y	V	W	D	A	L	N	Y	G	D	I	F	H	F	L	61	100	92
S.aureus	P	I	Y	Y	H	G	H	S	D	H	G	Q	K	F	G	V	V	V	Y	V	W	D	A	L	N	Y	A	E	I	F	H	W	I	41	67	84
T.brucei	P	I	Y	Y	H	G	H	S	D	H	G	Q	K	L	A	I	C	V	Y	V	W	D	A	L	N	Y	G	D	I	F	H	W	W	42	78	68
H.sapiens mito	P	I	F	Y	H	G	H	S	D	H	G	L	K	L	G	I	T	I	Y	V	W	D	A	L	N	Y	G	D	I	F	H	H	W	38	78	76
H.sapiens cyto	A	L	P	Y	H	G	N	C	D	Y	G	T	A	L	G	T	V	F	Y	V	W	D	A	T	G	Y	A	D	N	F	H	Y	L	26	56	56

Figure 11. **Amino-acid sequence alignment of residues forming the Met-AMP-binding and inhibitor-binding pockets in different species.** Met-AMP-binding pocket: MBP, methionine-binding pocket; ABP, adenine-binding pocket; P, Met-AMP phosphate-binding site; R, Met-AMP ribose-binding site. Inhibitor binding pocket: EMP, enlarged methionine pocket; LR, linker region; AP, auxiliary pocket. Residues indicated as part of the inhibitor-binding pocket are predicted to bind the MetRS inhibitors based on previous MetRS•inhibitor crystal (e.g. *Tbru*MetRS [90, 91, 93, 109, 110], *Bmel*MetRS [136]), *Saur*MetRS (PDB 4QRD, 4QRE)). OI, percentage overall protein sequence identity; MI, percentage identity for Met-AMP-binding pocket residues; II, percentage identity for inhibitor-binding pocket residues.

Figure 12 shows the temperature factor (B factor) distribution for the Met-AMP molecule in the crystal. The methionyl moiety has a lower B factor value than the adenosine group of Met-AMP. This might be due to the fact that the methionyl end of the Met-AMP molecule is better anchored to the protein than the adenosine end. This idea corresponds with the fact, discussed below, that the side chain of F292 side near the adenine ring has the highest B factor value in the *Mtub*MetRS Met-AMP binding site (Figure 12). Alternatively, it is possible that Met-AMP is not present in full occupancy but, instead, the *Mtub*MetRS crystal used is a combination of a mixture of *Mtub*MetRS•Met-AMP and *Mtub*MetRS•Met complexes (possibly even with partial occupancy). This might be somewhat similar to the structure of *Msme*MetRS•Met•Adenosine

(PDB ID: 2X1L [121]) where the adenosine molecule presents B factors more than 3 times higher than the methionine molecule.

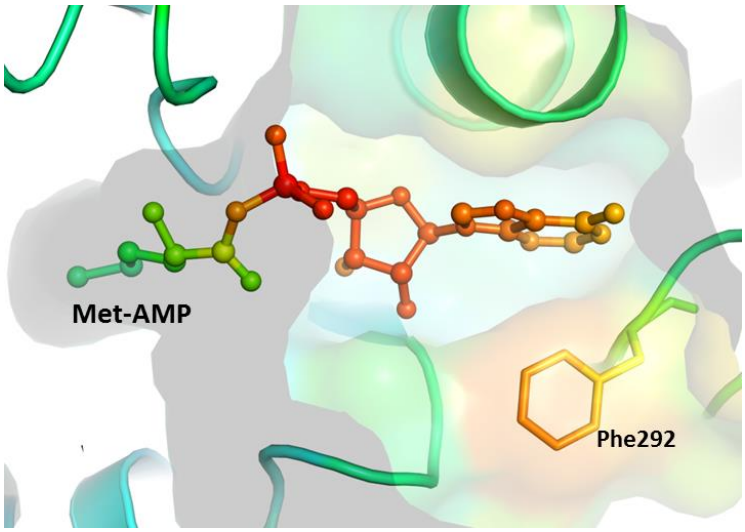


Figure 12. **Temperature factors of Met-AMP and the catalytic pocket.** The *Mtub*MetRS catalytic pocket and Met-AMP are shown colored according to B-factors, with cyan the lowest and red the highest values. From the residues binding Met-AMP in the *Mtub*MetRS•Met-AMP structure, F292 has the highest B-factor value (around 100 $Å^2$) which corresponds with high B-factors of the adenosyl group of the Met-AMP molecule.

2.3.3. The Met-AMP binding pocket among methionyl tRNA synthetases

There are a considerable number of MetRS structures with an adenine ring occupying the adenine binding pocket. Here we compare adenine binding in a number of pathogen structures and the only human MetRS structure so far, that of the human cytosolic enzyme. Figure 11 shows the sequence alignment of *Mtub*MetRS with *Msme*MetRS (its closest homolog with known structure [121]) and additional selected MetRS enzymes. F292 is the only residue interacting directly with Met-AMP that is different in these two mycobacterial enzymes. The side chain of F292 in *Mtub*MetRS exhibits an intriguing difference in conformation with respect to the equivalent W294 in *Msme*MetRS. While in *Msme*MetRS•Met•Adenosine the indole ring of W294 adopts an inward

conformation ("in"), resulting in a stacking hydrophobic interaction with the adenine ring (Figure 13 B) of the adenosine molecule (PDB ID: 2X1L [121]), the phenyl side chain group of F292 is in an "out" conformation in *Mtub*MetRS•Met-AMP, with a single atom of F292 in contact with the adenine ring, at a distance of 3.9 Å. The high B factor value of the F292 side chain (Figures 12 and 13 A) suggests that the side chain of F292 adopts multiple conformations. In the structure of *Tbru*MetRS•Met-AMP (PDB ID: 4EG3 [90]), the equivalent W547 is interacting with the adenine ring of Met-AMP with an "in" conformation (Figure 13 C). In the structure of a truncated version of the human cytosolic apo MetRS (*Hsap*CytoMetRS), comprising residues 221-834, the residue corresponding to *Mtub*MetRS F292 is Y586 and its side chain adopts an "out" conformation (Figure 13 D). In several MetRS structures with a filled adenine pocket, a homologous residue corresponding with *Mtub*MetRS F292 is present which in some instances is in an "in" conformation (e.g. *Aaeo*MetRS•Met-AMP•tRNA, PDB ID: 2CT8,) and in other cases in an "out" conformation (e.g. *Lmaj*MetRS•Met-AMP• $PP_{i,}$ PDB ID: 3KFL). Hence, which conformation the residue equivalent to *Mtub*MetRS F292 adopts in related MetRSs is clearly highly variable. The human mitochondrial enzyme presents residue H340 in the same position and it is obviously hard to predict what conformation this residue could adopt. Nevertheless, the structural and sequence differences near the adenine binding pocket of *Mtub*MetRS vs. the human homologs might be of help in future anti-tuberculosis chemotherapeutic development.

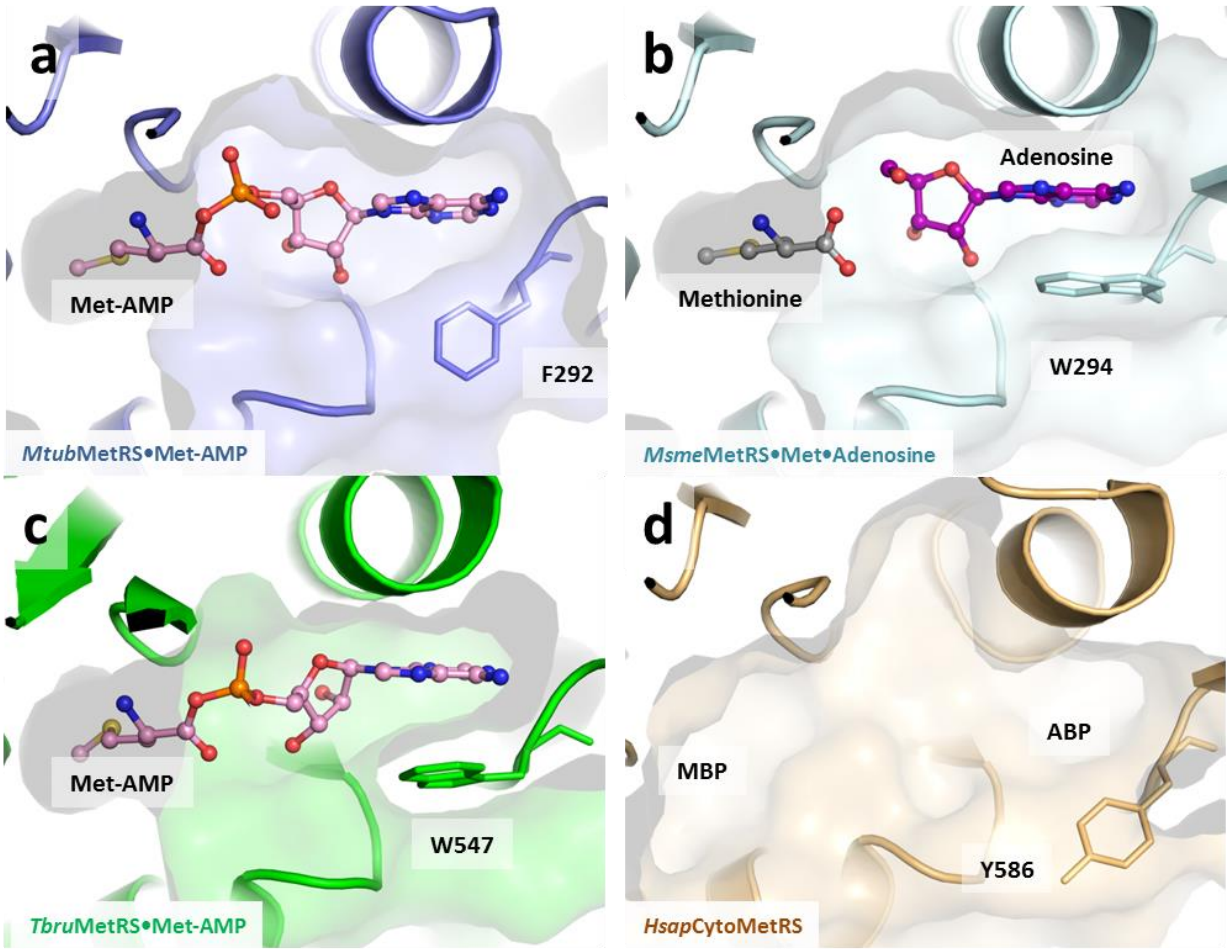


Figure 13. **Interactions of residue F292 from *Mtub*MetRS with the Met-AMP adenine ring and comparison with selected MetRS structures.** A) Met-AMP adenine ring interaction with F292 (sticks) in the *Mtub*MetRS•Met-AMP structure. *Mtub*MetRS is shown in blue with the Met-AMP carbon atoms in pink. B) In the *Msme*MetRS•Met•Adenosine structure (PDB ID: 2X1L [121]) the adenine ring of adenosine engages in a hydrophobic stacking interaction with W294 which is equivalent to F292 of *Mtub*MetRS. W294 shown in sticks. *Msme*MetRS colored cyan, methionine carbon atoms grey and adenosine carbon atoms purple. C) Likewise, in the *Tbru*MetRS•Met-AMP structure (PDB ID: 4EG3 [90]) the adenine ring of Met-AMP engages in a hydrophobic stacking interaction with W547 (shown as sticks; equivalent to F292 of *Mtub*MetRS). *Tbru*MetRS colored green with Met-AMP carbons pink, nitrogens blue, oxygens red, phosphorus orange and sulfur yellow. D) Surface representation of the Met-AMP binding pocket of the apo-*Hsap*CytoMetRS (PDB ID: 5GL7) structure. *Hsap*CytoMetRS orange. Residue Y586, equivalent to F292 of *Mtub*MetRS, is depicted as sticks. The empty MBP and ABP pockets are labeled.

2.3.4. M. tuberculosis MetRS versus its human homologs

It is imperative to compare the *Mtub*MetRS structure with its human homologs for evaluating the potential of arriving at selective inhibitors. Superposition of *Mtub*MetRS onto the available human cytosolic apo MetRS (*Hsap*CytoMetRS) crystal structure (PDB ID: 5GL7), yields an overall r.m.s.d. of 1.26 Å for the C^{α} atoms. The C^{α} atoms of the catalytic and anticodon binding domains of these two structures superpose quite well, with r.m.s.d values of 1.07 and 1.11 Å, respectively. When superposing the catalytic domains of both structures, the connective peptide CP1 sub-domain of *Mtub*MetRS is moved considerably "downwards" towards the active site, adopting a closed conformation (Figure 14 A) as described before for *Msme*MetRS [121], different from the conformation of CP1 in human cytosolic MetRS and in other MetRSs [137, 138]. Especially interesting is the fact that the same sub-domain changes conformation upon binding of inhibitors in *Tbru*MetRS [90, 91, 93, 109, 110]. In the *Tbru*MetRS•Met-AMP structure, CP1 is in the closed conformation, similar to the one described here for *Mtub*MetRS•Met-AMP, but adopts an open conformation when a *Tbru*MetRS inhibitor is bound [90]. In the case of the human apo cytosolic enzyme structure, CP1 is in a much more open conformation (Figure 14 A). The structural differences regarding CP1 might be exploitable when aiming for selectivity of compounds against *Mtub*MetRS since CP1 residues could potentially interact with a compound designed to bind specifically to the mycobacterial enzyme, keeping the CP1 knuckle in the open conformation and thereby prevent catalytic action.

Overall, *Mtub*MetRS and the human mitochondrial and cytosolic counterparts are 38% and 26% identical in sequence, respectively, with 78% and 56% identity, respectively, when comparing residues binding Met-AMP in *Mtub*MetRS with *Hsap*MitoMetRS and *Hsap*CytoMetRS (Figure 11). When comparing the available *Hsap*CytoMetRS structure with that

of *Mtub*MetRS, there is a significant difference between the two enzymes close to the Met-AMP ribose binding site in *Mtub*MetRS. The homologous segment to helix α1 of *Mtub*MetRS forms in *Hsap*CytoMetRS two helices: α2 and α3 (Figures 14 B, 14 C and 15) separated by a hinge region. This secondary structure difference is accompanied by the presence in *Hsap*CytoMetRS of a "human cytosolic extra pocket" (HCEP) harboring four water molecules (Figure 14 D). Due to both the difference in the secondary structure of helix α1 of *Mtub*MetRS vs. helices α2 and α3 in *Hsap*CytoMetRS, as well as the positioning of the pocket filler side chain of residue H21 in *Mtub*MetRS, the bacterial pocket is smaller, harboring one water molecule instead of four (Figure 14 E). When aligning the bacterial and human sequences, the reason for this difference appears to be the insertion of an extra glycine, G286 in the human cytosolic protein, exactly in the connection between both helices (Figure 15). In the case of *Hsap*MitoMetRS, whose structure is not available, the amino acid sequence lacks the extra glycine residue present in *Hsap*CytoMetRS between helices α2 and α3. Therefore, the hinge would be expected to be absent, probably resulting in a secondary structure similar to helix α1 in *Mtub*MetRS and the concomitant absence of the HCEP pocket in the human mitochondrial enzyme.

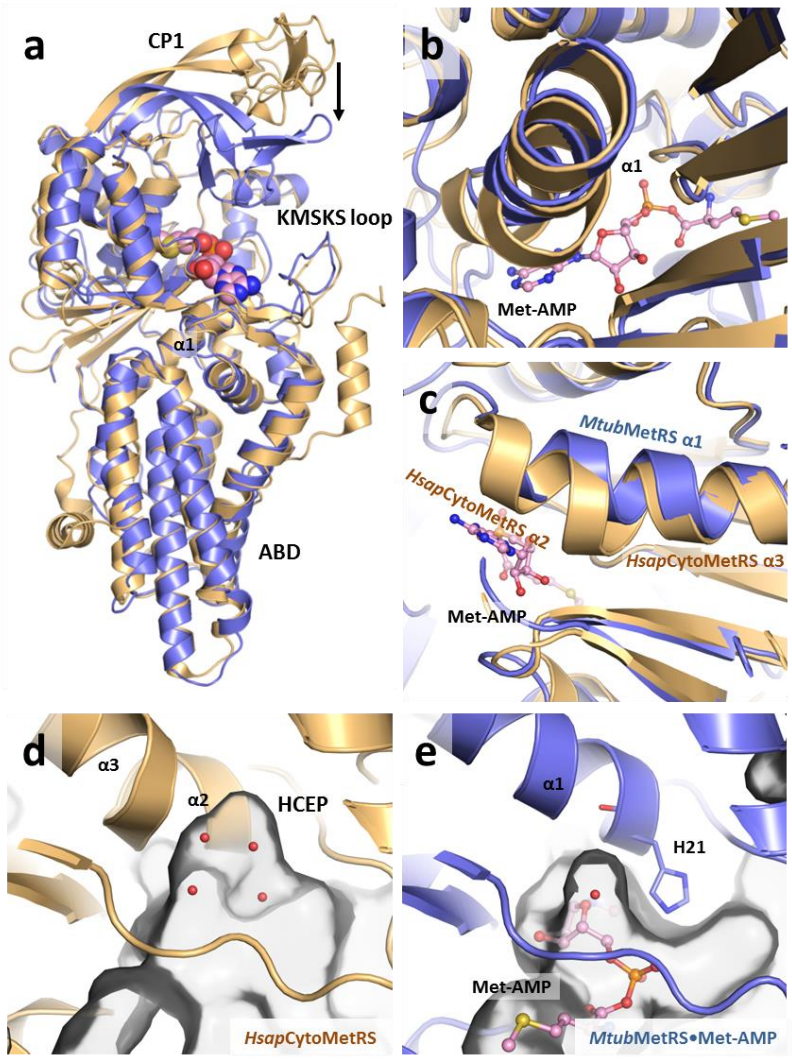


Figure 14. **Comparison of *Mtub*MetRS and *Hsap*CytoMetRS structures.** A) Superposition of *Hsap*CytoMetRS (PDB ID: 5GL7) (light orange) and *Mtub*MetRS (blue) catalytic domains. The mainly β sub-domain CP1 of *Mtub*MetRS is shifted "downwards" towards the active site (black arrow) as compared with the human enzyme. B) In the catalytic domain, a distinct difference between the conformation of *Mtub*MetRS α1 and the same region of *Hsap*CytoMetRS is evident. C) Close to where the Met-AMP ribose ring binds in *Mtub*MetRS, instead of a single helix α1, *Hsap*CytoMetRS has two helices α2 and α3 connected by a hinge. D) A pocket, called the human cytosolic extra pocket (HCEP), harboring four waters (red spheres) is present in *Hsap*CytoMetRS close to the ribose binding region (ribose not shown). E) The presence of residue H21 makes this pocket smaller in *Mtub*MetRS, containing only one water molecule. Met-AMP carbon atoms colored pink, nitrogens blue, oxygens red, phosphorus orange and sulfur yellow.

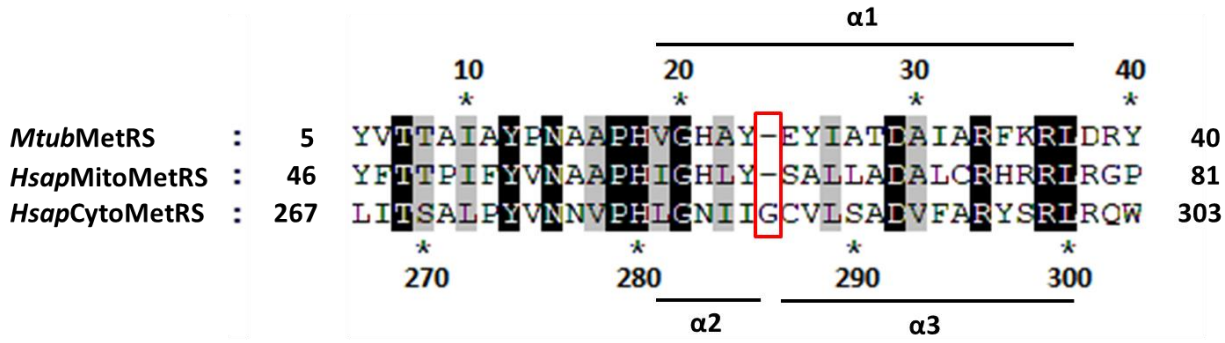


Figure 15. **Sequence alignment of the *Mtub*MetRS α1 helix region with the human MetRS homologs.** In *Hsap*CytoMetRS, the secondary structure of the region corresponds to two α helices, α2 and α3. The insertion, compared to both *Mtub*MetRS and *Hsap*MitoMetRS, of residue G286 in *Hsap*CytoMetRS is indicated by a red box. C*onserved* residues are shown on a black background, conservative changes on a grey background and divergent residues on a white background.

Regarding the design of compounds that have higher affinity for tuberculosis MetRS than for the human homologs it is of interest that in the Met-AMP binding region of the human cytosolic enzyme one out of about every two residues is different from *Mtub*MetRS (Figure 11). Moreover, the cytosolic enzyme is a MetRS2 enzyme and *Mtub*MetRS a MetRS1 enzyme. As reported before for *Saur*MetRS, many compounds inhibiting bacterial MetRS1 are not potent against the human cytosolic enzyme, which belongs to the MetRS2 form [87]. Specifically, REP8839 was reported to have a selectivity of more than 1,000,000-fold for *Saur*MetRS over the 27% identical human cytosolic enzyme [87]. Therefore, differentially targeting *Mtub*MetRS and not the cytosolic enzyme should be readily feasible.

This is somewhat different for the human mitochondrial enzyme since the *M. tuberculosis* and human mitochondrial enzyme both belong to the MetRS1 form. Nevertheless, 22 % of the residues forming the Met-AMP binding in *M. tuberculosis* and human mitochondrial MetRS are

different, and 24 % in the inhibitor binding site (Figure 11). Moreover, compound REP8839, mentioned above, has a 1,000-fold higher affinity for the *S. aureus* enzyme than for human mitochondrial MetRS [87], while the amino acid difference in the inhibitor binding site between the *S. aureus* and the human mitochondrial enzyme is 24%, i.e. the same as in the comparison of *M. tuberculosis* and human mitochondrial MetRS. The fact that considerable selectivity between human mitochondrial and bacterial MetRS enzymes could be achieved in the case of *S. aureus* is promising and indicates that there are opportunities to arrive at inhibitors which have a higher affinity for the *M. tuberculosis* MetRS than for both human homologs.

In conclusion, this work reports the crystal structure of *M. tuberculosis* MetRS bound to its substrate intermediate. Analysis of the structure reveals differences with the human orthologues that will be useful for designing selective inhibitors in the effort to discover new anti-tuberculosis drugs.

Acknowledgments

We thank Robert Steinfeldt for providing support for the computing environment at the Biomolecular Structure Center of the University of Washington. Stephen Nakazawa Hewitt and Cho Yeow Koh contributed to the initial stages of this project. Crystallography performed in support of the work benefitted from remote access to resources at the Stanford Synchrotron Radiation Lightsource supported by the U.S. Department of Energy Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515 and by the National Institutes of Health (P41GM103393). We acknowledge the support of a Fulbright Fellowship to X.B.-A. Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R01AI084004 and R01AI097177.

Chapter 3

Leishmania donovani tyrosyl-tRNA synthetase structure in complex with a tyrosyl adenylate analog and comparisons with human and protozoan counterparts

A modified version of this chapter was previously published in Barros-Álvarez *et al*., 2017 [107].

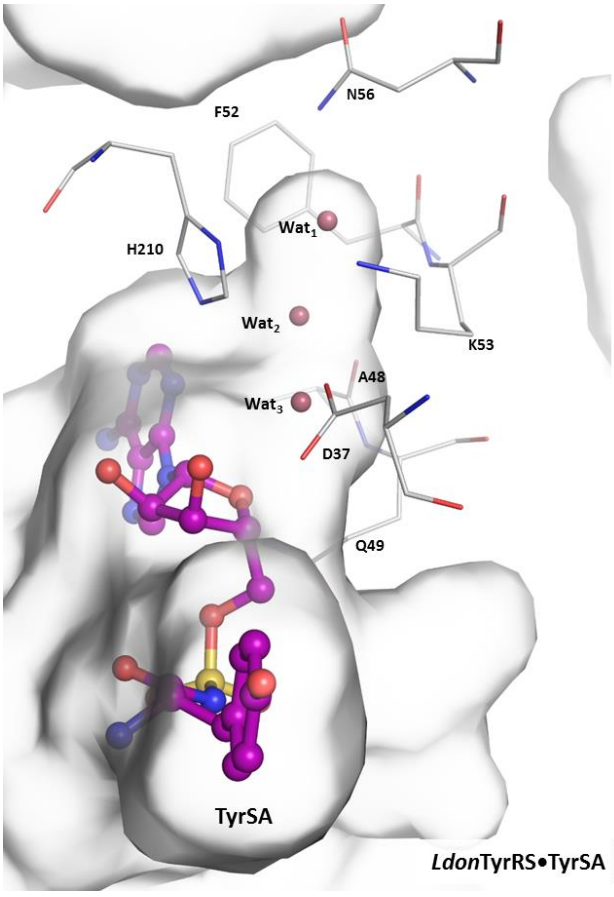
ABSTRACT

The crystal structure of *Leishmania donovani* tyrosyl-tRNA synthetase (*LdonTyrRS*) in complex with a nanobody and the tyrosyl adenylate analog TyrSA was determined at 2.75 Å resolution. Nanobodies are the variable domains of camelid heavy chain-only antibodies. The nanobody makes numerous crystal contacts and in addition reduces the flexibility of a loop of *LdonTyrRS*. TyrSA is engaged in many interactions with active site residues occupying the tyrosine and adenine binding pockets. The *LdonTyrRS* polypeptide chain consists of two pseudo-monomers, each consisting of two domains. Comparing the two independent chains in the asymmetric unit reveals that the two pseudo-monomers of *LdonTyrRS* can bend with respect to each other essentially as rigid bodies. This flexibility might be useful in the positioning of tRNA for catalysis since both pseudo-monomers in the *LdonTyrRS* chain are needed for charging $tRNA^{Tyr}$.

An "extra pocket" (EP) appears to be present near the adenine binding region of *LdonTyrRS*. Since this pocket is absent in the two human homologous enzymes, the EP provides interesting opportunities for obtaining selective drugs for treating infections caused by *L. donovani*, a unicellular parasite causing visceral leishmaniasis, or kala azar, which claims

20,000 to 30,000 deaths per year. Sequence and structural comparisons indicate that the EP is a characteristic which also occurs in the active site of several other important pathogenic protozoa. Therefore the structure of *LdonTyrRS* could inspire the design of compounds useful for treating several different parasitic diseases.

Graphical abstract



3.1. Introduction

The leishmaniases are a variety of diseases caused by more than 20 *Leishmania* species. These protozoans are transmitted through the bites of infected female phlebotomine sandflies. Approximately 350 million people in the tropics and sub-tropics are at risk of infection, with 0.9 to 1.3 million new cases and 20,000 to 30,000 deaths annually [22, 23]. Depending on the species involved, cutaneous, mucocutaneous or visceral leishmaniasis can develop, the latter being the most serious as far as number of deaths is concerned. Caused by *L. donovani* and *L. infantum*, visceral leishmaniasis (VL), or kala-azar, is fatal if left untreated in 95% of the cases. VL is characterized by irregular bouts of fever, weight loss, enlargement of the spleen and liver, and anemia. About 200,000 to 400,000 new cases of VL occur each year and children are the most severely affected group [23, 24]. Available drugs used for the treatment of the leishmaniases are mainly pentavalent antimonial complexes, amphotericin B, the aminoglycoside paromomycin, and the alkylphosphocholine miltefosine, either as single or combination treatments [21, 26]. Most of the available drugs are not oral and, due to their complicated administration regimens, low efficacy for parasite elimination, poor safety profile, occurrence of drug resistance [21] and treatment failure, there is an urgent need to take new oral drug candidates into clinical development [22].

Protein translation is an essential cellular function in Trypanosomatid parasites. The overall fidelity of protein synthesis relies on the accuracy of both codon-anticodon recognition and aminoacyl-tRNA synthesis [41]. Aminoacyl-tRNA synthetases (aaRSs) play a crucial role in accurately pairing each amino acid with its cognate tRNA through a two-step esterification reaction forming aminoacyl-tRNA [111]. The first, ATP dependent, step leads to the formation of an enzyme-bound aminoacyl-adenylate and an inorganic pyrophosphate leaving group. The second

step results in the 3'-esterification of the tRNA with the amino acid moiety and generation of AMP as a leaving group, followed by product release [41].

Human cells have two sets of aaRS, one for cytosolic aminoacyl-tRNA synthesis and a second one, of bacterial evolutionary origin, for aminoacylation of mitochondrial tRNAs. In contrast, Trypanosomatids contain only a single set of aaRS genes, with the exception of three amino acids (aspartate, lysine and tryptophan) for which two aaRS genes exist [50]. Hence, in these parasites 17 aaRSs have to function in both subcellular locations. This is a remarkable difference between human and parasite aaRS. It suggests that inhibiting any of these aaRS in parasites would have the effect of hindering protein synthesis in the cytosol as well as in the mitochondria.

AaRSs have been recognized as validated antimicrobial drug targets [39, 45, 112]. Anti-aaRS compounds can act by blocking the binding site of ATP and/or amino acid, an allosteric site, tRNA recognition, or a secondary editing site [56]. Successes in targeting pathogenic aaRS are encouraging. At least two drugs, mupirocin and tavaborole, are in clinical topical use against, respectively, *Staphylococcus aureus* infections and onychomycosis. Mupirocin binds both the ATP and Ile binding sites of *S. aureus* isoleucyl-tRNA synthetase (IleRS), while tavaborole (Kerydin®) inhibits fungal leucyl-tRNA (LeuRS) by binding to its editing site [39, 58, 60]. Compounds targeting parasitic protozoan aaRS have been shown to interfere with protozoan cell growth [39, 91]. Also, several natural products inhibit aaRS from important tropical parasites. For instance, *Plasmodium falciparum* prolyl-tRNA synthetase, lysyl-tRNA synthetase and threonyl-tRNA synthetase are inhibited by the natural products halofuginone, cladosporin and borrelidin, respectively [56, 62, 139]. These compounds have also inhibitory effects on malaria parasite growth *in vitro* or *in vivo* [140]. Here we focus on tyrosyl-tRNA synthetase (TyrRS) from *L donovani.*

TyrRS belongs to the class I aaRSs, characterized by a Rossmann fold and two hallmark sequence motifs ("HIGH" and "KMSKS"). More specifically, TyrRS belongs to subclass Ic, together with tryptophanyl-tRNA synthetase (TrpRS), and contains an "AIDQ" motif characteristic of the ATP binding site. Progress has been made in inhibiting bacterial TyrRSs both with compounds identified from natural sources, as is the case of SB-219383 [96-98], and with synthetized inhibitors [100-103]. However, there have been no significant advances in experimental work with parasitic TyrRSs as drug targets [39]. A structure guided approach in the design of drugs for the treatment of VL would greatly benefit from the knowledge and analysis of new leishmanial TyrRSs structures, in particular from *L. donovani* and *L. infantum*.

Recent biochemical and parasitological studies on *L. donovani* TyrRS (*LdonTyrRS*) have indicated that the enzyme has a cytoplasmic subcellular location and that, as expected, it is essential for cell proliferation. Interestingly, an intriguing additional function, i.e. an effect on inflammation, has been shown in *L. donovani* for secreted TyrRS. The enzyme appears to be involved in attracting neutrophils and binding to macrophages through an N-terminal ELR motif triggering further cytokine TNF-α and IL-6 release by host macrophages [141]. Regarding inhibition studies of *LdonTyrRS*, the flavonoid fisetin has shown to have anti-leishmanial properties and its effect has been ascribed to its anti-TyrRS activity [141]. This inhibitory effect of fisetin agrees well with the observation that this compound binds to the active site of *L. major* TyrRS (*LmajTyrRS*) [105].

In order to provide a structural platform to assist further anti-leishmanial drug development, we report here the 2.75 Å resolution crystal structure of *LdonTyrRS* in complex with the tyrosyl adenylate analog TyrSA (5'-O-[N-(L-tyrosyl)sulfamoyl]adenosine) and a specific anti-*LdonTyrRS* nanobody (NbA) used as crystallization chaperone [142]. Nanobodies are the variable

domains of camelid heavy chain-only antibodies which can substantially increase the success of protein crystal growth [142]. The structure of *LdonTyrRS* in complex with TyrSA will in particular expand our insight into the binding characteristics of ligands in the neighborhood of the adenine binding pocket, a feature which was not probed in the previous structures of *LmajTyrRS*. Analysis of the structure indicates the presence of promising opportunities to exploit structural differences of the parasite enzyme with both human cytosolic and mitochondrial TyrRS variants. Of particular relevance is the presence of an "extra pocket" (EP) near the adenine binding site in the structure of *LdonTyrRS*. This pocket is absent in the two human homologs and hence provides distinct opportunities to arrive at selective inhibitors. The EP is also present in the structure of *L. major* TyrRS [105], and, in view of the considerable similarities in the active site regions, is highly likely to occur in other *Leishmania* species causing human disease. The same holds for TyrRS from related Trypanosomatids, such as *Trypanosoma brucei,* the causative agent of human African trypanosomiasis, and *T. cruzi,* responsible for Chagas disease. In addition, TyrRS from the malaria parasite *Plasmodium falciparum* [104] appears to contain a close variant of the EP. Amino acid sequence comparisons suggest that the EP occurs in even more parasitic protozoa. Hence the EP may allow the development of selective tyrosyl-tRNA synthetase inhibitors which could become new tools in tackling several major diseases caused by unicellular parasites.

3.2. Materials and methods

3.2.1. *LdonTyrRS* expression and purification

LdonTyrRS was cloned into the AVA0421 vector and expressed in *E. coli* for subsequent purification. A first round of Ni-NTA affinity chromatography was followed by cleavage of the N-terminal His_6-tag using 3C protease (overnight at 4 °C). In a second Ni-NTA step, the cleaved *LdonTyrRS* was purified from the N-terminally His_6-tagged 3C protease. A size-exclusion chromatography on a Superdex 200 column (Amersham Pharmacia Biotech) using SEC buffer (25 mM HEPES at pH 7.25, 500 mM NaCl, 2 mM TCEP, 5% glycerol, 0.025% NaN_3) was performed and a final yield of 6 mg of pure *LdonTyrRS* per liter of *E. coli* culture was obtained and concentrated above 15 mg/mL for co-crystallization with NbA.

3.2.2. NbA production

LdonTyrRS specific nanobodies were generated as previously described [142]. In brief, one llama (*Lama glama*) was immunized six times with in total 0.9 mg of *LdonTyrRS*. Four days after the final boost, blood was taken to isolate peripheral blood lymphocytes. RNA was purified from these lymphocytes and reverse transcribed by PCR to obtain cDNA. The resulting library was cloned into the phage display vector pMESy4 bearing a C-terminal hexa-His tag and a CaptureSelect sequence tag (Glu-Pro-Glu-Ala). Six different families were selected by biopanning. For this, *LdonTyrRS* was solid phase coated directly on plates, *LdonTyrRS* specific phage were recovered by limited trypsinization. After two rounds of selection, periplasmic extracts were made and subjected to ELISA screens [142].

3.2.3. NbA expression and purification

NbA cloned in the pMESy4 vector that carries the *pel*B sequence coding for the secretion signal peptide of PelB was expressed in *E. coli* for subsequent purification from the bacterial periplasm. NbA purification was performed as previously described by Pardon *et al.* [142]. Briefly, after induction, the bacterial pellet was gently resuspended in TES buffer (200 mM Tris at pH 8.0, 0.5 mM EDTA, 500 mM Sucrose) and upon incubation on ice the periplasmic content was recovered by centrifugation. NbA was purified from the periplasmic content by Ni-NTA affinity chromatography followed by size-exclusion chromatography on a Superdex 75 column (Amersham Pharmacia Biotech) using NbSEC buffer (25 mM HEPES at pH 7.25, 300 mM NaCl, 1 mM TCEP, 10% glycerol, 0.025% NaN_3) and concentrated above 10 mg/mL for co-crystallization with *LdonTyrRS*.

3.2.4. Nanobody-*LdonTyrRS* binding studies

Nanobody (Nb) binding studies were systematically carried out by native gel electrophoresis and size-exclusion chromatography for eight nanobodies. The purified Nb was first incubated with *LdonTyrRS* for 30 min at 4 °C and upon a native gel electrophoresis the positions of *LdonTyrRS*, Nb and complex were analyzed. The formation of the complex was also tested by comparing the elution peaks on a Superdex 200 column (Amersham Pharmacia Biotech) where *LdonTyrRS* was run as well as the potential *LdonTyrRS*•Nb complex formed upon 30 min incubation at 4 °C.

3.2.5. *LdonTyrRS* aminoacylation assay

The IC_{50} of TyrSA in the *LdonTyrRS* aminoacylation assay was determined using methods as previously described [88, 92, 117]. Briefly, TyrSA (tested in triplicate) was pre-incubated for 15 minutes

at room temperature with 0.13 nM *LdonTyrRS*, 500 nM [3H]L-tyrosine (40 Ci/mmol), 0.2 mM ATP. 0.1 U/mL pyrophosphatase, 0.2 mM spermine, 0.2 mg/mL bovine serum albumin, 2.5 mM dithiothreitol, 1 mM $MgCl_2$, 25 mM KCl, 50 mM HEPES-KOH pH 7.6, and 2% DMSO. The reactions were started with 200 µg/mL tRNA from brewer's yeast (Roche) and incubated for 30 minutes at room temperature without shaking. The reactions were stopped with cold 10% trichloroacetic acid and processed as previous described [117].

3.2.6. *LdonTyrRS*•NbA•TyrSA complex crystallization

Purified *LdonTyrRS* and NbA proteins were incubated on ice for 30 min at a 1:2 molar ratio followed by buffer exchange to crystallization buffer (25 mM HEPES at pH 7.25, 100 mM NaCl, 1 mM TCEP-HCl, 5% glycerol, 0.025% NaN_3). The complex (5 mg/mL) was then incubated with 200 µM of TyrSA (5'-O-[N-(L-tyrosyl)sulfamoyl]adenosine) on ice for 30 - 60 min prior to setting up the crystallization tray. The crystals were obtained after 5 - 7 days at room temperature by vapor diffusion using sitting drops equilibrated against a reservoir containing 0.1 M sodium cacodylate pH 5.7, 22% PEG 4,000. The drops contained 1 µL of *LdonTyrRS*•NbA•TyrSA complex at 5 mg/mL and 1 µL of reservoir solution. After growth, crystals were flash frozen in liquid nitrogen in cryo-solution (25% glycerol in reservoir solution) and stored until data collection.

3.2.7. Data Collection and Structure Determination

Data was collected under cryogenic conditions at the Stanford Synchrotron Radiation Lightsource (SSRL) using beamline 12-2 at a wavelength of 1 Å. Data processing was carried out with the program HKL2000 [129]. Initial phases were obtained by molecular replacement using Phaser [130] with as models the structure of *L. major* TyrRS•Tyrosinol ([105]; PDB: 3P0J) and a NbA homology

model generated by Phyre2 [143]. This was followed by iterations of manual building and rebuilding using Coot [131] alternated with refinement of the structure with REFMAC5 [132]. Refinement restraints for TyrSA were obtained with the Grade web server [144]. The structure validation server MolProbity [133] was used throughout the process for structure validation. The final data collection and crystallographic refinement statistics are given in Table 4. Pymol [134] was used to create the figures. Coordinates and structure factors of the *LdonTyrRS*•NbA•TyrSA complex have been deposited in the Protein Data Bank under the PDB ID: 5USF.

Table 4. Crystallographic data collection and refinement statistics.

Parameters	*LdonTyrRS*•NbA•TyrSA (PDB: 5USF)
Data collection	
Space group	$P\ 6_5$
Cell dimensions: *a*, *b*, *c* (Å)	96.18, 96.18, 351.83
Resolution (Å)	83.29 – 2.75 (2.85 – 2.75)
R_{merge}	0.210 (1.911)
R_{pim}	0.069 (0.659)
Observed reflections	488525 (43134)
Unique reflections	47547 (4632)
Mean $I / \sigma I$	8.7 (2.0)
Multiplicity	10.3 (9.3)
Completeness (%)	100.0 (100.0)
$CC_{1/2}$	0.995 (0.607)
Refinement	
Resolution (Å)	83.29- 2.75
Reflections used	45156
R_{work} / R_{free}	0.19/0.24
Number of atoms	
Protein	12283
TyrSA	70
Water	353
Number of residues	1600
Average *B*-factors (Å^2)	
Protein	66.0
TyrSA	49.8
Water	55.9
R.m.s. deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.44
Ramachandran plot[#]	
Favored (%)	97
Outlier (%)	0
Ligand (TyrSA)	
Average LLDF[@]	-0.53
Average RSR[^]	0.14

Values in parentheses are for the highest-resolution shell.

[#]Ramachandran Plot statistics as reported by the wwPDB validation report

[@]Local ligand density fit as reported by the wwPDB validation report

[^]Real space R value as reported by the wwPDB validation report

3.3. Results

3.3.1. *Leishmania donovani* tyrosyl-tRNA synthetase structure

The crystal structure of the *LdonTyrRS*•NbA•TyrSA complex was determined at a resolution of 2.75 Å (Table 4), where NbA is a nanobody to be described below, TyrSA a tyrosyl adenylate analog, and the symbol "•" indicates a non-covalent complex. The crystals contain two copies of the *LdonTyrRS*•NbA•TyrSA complex in the asymmetric unit (ASU). Each *LdonTyrRS* molecule is bound to one NbA molecule (Figure 16). Canonical TyrRSs are formed by two identical monomers, each consisting of a catalytic domain (CD) and an anticodon binding domain (ABD). Instead, the 75 kDa *LdonTyrRS* chain is a pseudo-dimer where the structurally similar N- and C-terminal pseudo-monomers share only 23% sequence identity. As described for *LmajTyrRS* [105], the N- and C-terminal pseudo-monomers are connected by a flexible linker between α14 and β9 (Figure 16 A). The N-terminal pseudo-monomer of *LdonTyrRS* contains the three motifs characteristic for the catalytic domain of Class I tRNA synthetases: "HIGH" (^{46}HIAQ^{49} in *LdonTyrRS*), "AIDQ" (^{182}GLDQ^{185} in *LdonTyrRS*) and "KMSKS" (^{222}KMSKS^{226} in *LdonTyrRS*) (Figure 16 B). These motifs have been shown to be involved in the catalytic activity of TyrRS enzymes [105, 145]. The motifs are absent in the C-terminal pseudo-monomer, suggesting that this half of the molecule is not able to perform amino acid activation. This is in agreement with the observations that TyrSA is bound to the active site in the N-terminal CD of *LdonTyrRS*, but not to the C-terminal CD. Hence, the N-terminal CD is called the "functional CD" and the C-terminal CD the "non-functional CD" (Figure 16 A).

The opposite is true for the ABD. Although both ABDs contain the sequence motifs "AC1" (^{244}KIRQAYC^{250} and ^{578}KIKKAYS^{584} in *LdonTyrRS*) and "AC2" (^{313}VSEDALK^{319} and ^{636}LHPADLK^{642} in *LdonTyrRS*) involved in the recognition of the tRNA anticodon arm, as

described for *LmajTyrRS* [105], the loop located between β7 and β8 in the N-terminal ABD is considerably shorter than the corresponding loop between β14 and β15 in the C-terminal pseudo-monomer. This loop is responsible for binding the anticodon base G34 of $tRNA^{Tyr}$ according to the structure of *Saccharomyces cerevisiae* TyrRS [146]. The short loop in the N-terminal ABD homolog is unable to engage with this base (Figure 16 C). Therefore, the N-terminal ABD is called the "non-functional ABD" and the C-terminal ABD the "functional ABD" (Figure 16 A, C).

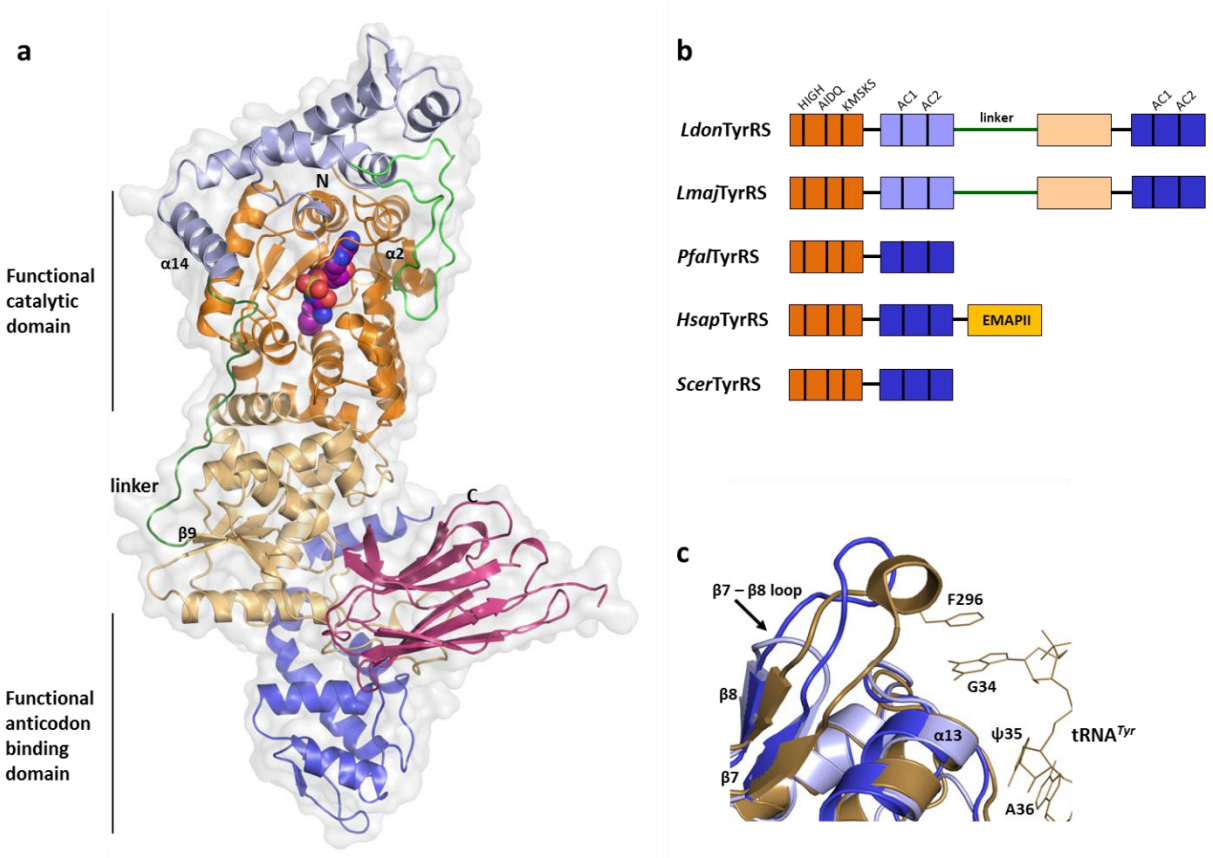


Figure 16. **Domain organization of pseudo-dimeric *LdonTyrRS*.** A) The asymmetric *LdonTyrRS* pseudo-dimer structure in complex with nanobody A (pink) and the tyrosyl adenylate analog TyrSA (CPK model). The N-terminal functional catalytic domain (CD) is shown in dark orange, the C-terminal nonfunctional CD in light orange; the N-terminal non-functional ABD in light blue, the C-terminal functional ABD in dark blue. The linker connecting the N-terminal and C-terminal pseudo-monomers of *LdonTyrRS* is shown in dark green, while the plant/plastid insertion

present in Trypanosomatid TyrRSs, described previously by Larson *et al.* [105], in the N-terminal ABD is shown in bright green. The ELR motif is located in the exposed α2 in the functional CD. B) Domain organization of TyrRS from *L. donovani*, *L. major*, *P. falciparum* (*PfalTyrRS*), human (*HsapTyrRS*) and the yeast *S. cerevisiae* (*ScerTyrRS*). Domain colors correspond to the ones in *LdonTyrRS* structure in part A. The characteristic motifs HIGH, AIDQ and KMSKS in the catalytic domain as well as AC1 and AC2 in the anticodon binding domains are depicted. When processed by an elastase enzyme *HsapTyrRS* gives rise to an N-terminal TyrRS known as Mini TyrRS and a C-terminal EMAPII-like domain (yellow), which has cell signaling activity [104]. (EMAP stands for: "Endothelial Monocyte-Activating Polypeptide II"). C) Anticodon recognition regions of the N-terminal and C-terminal pseudo-monomers of *LdonTyrRS* and *ScerTyrRS*. The *LdonTyrRS* N-terminal (light blue) and C-terminal (dark blue) ABDs are shown superimposed with the *ScerTyrRS*•tRNA^{Tyr} complex (PDB ID: 2dlc) (sand). Residue F296, implied in *ScerTyrRS* tRNA^{Tyr} recognition, and the tRNA^{Tyr} anticodon bases guanine - pseudo-uridine - adenine (^{34}G-ψ-A^{36}) are depicted. The secondary structure elements in the *LdonTyrRS* N-terminal ABD are labeled. A black arrow points to the shortened loop between β7 and β8 in the N-terminal ABD.

The superposition of the two *LdonTyrRS* chains A and B in the *LdonTyrRS*•NbA•TyrSA crystals yields an overall r.m.s.d of 1.57 Å for 678 C^{α} atoms. All four individual domains of the two chains in the asymmetric unit are highly similar, with r.m.s.d values ranging between 0.25 and 0.77 Å. The N- or C- terminal pseudo-monomers of the A and B chains are also similar to each other, with r.m.s.d. values of 0.52 Å and 0.77 Å after superposition. However, after applying the superposition operation of the N-terminal pseudo-monomers to the entire chains, there appears to be a change in orientation of the two C-terminal pseudo-monomers, resulting in a 10.4 Å displacement at the farthest end of the C-terminal pseudo-monomer loop between β14 and β15 (Figure 17). These comparisons of chains A and B indicate that there is intrinsic flexibility within the *LdonTyrRS* molecule, in particular between the N- and C-terminal pseudo-monomers.

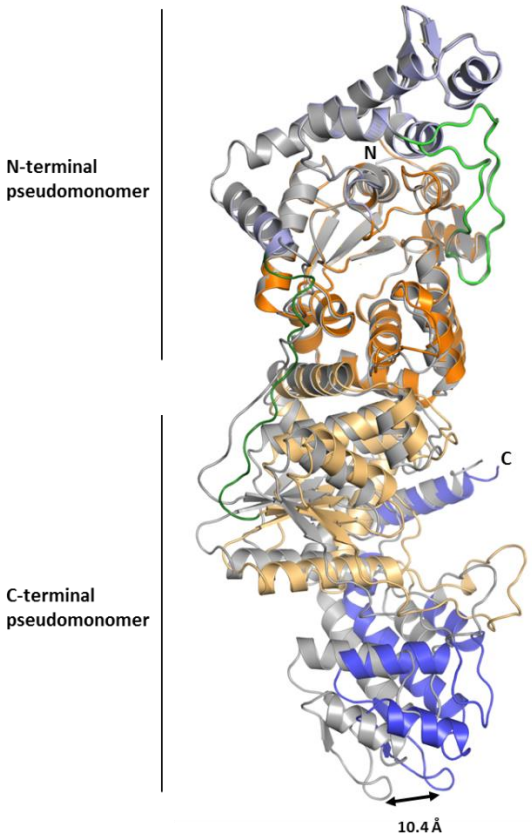


Figure 17. **Analysis of *LdonTyrRS* flexibility through superposition of *LdonTyrRS* chains A and B within the crystal.** Chain A is colored by domains as described in Figure, while chain B is shown in grey. When superimposing the N-terminal pseudo-monomers of both *LdonTyrRS* copies in the crystal, a shift in the C-terminal half of the enzyme is made evident especially in the farthest end of the C-terminal pseudo-monomer where a 10.4 Å displacement occurs (black arrow).

When considering the *LdonTyrRS* active site, all chain A residues are well defined, while in chain B density for two residues (S^{224} and K^{225}) belonging to the ^{221}KMSKS^{226} loop is missing. Therefore, we focus our structural analysis on chain A of *LdonTyrRS*. The KMSKS loop adopts a closed conformation in our *LdonTyrRS* structure with the tyrosyl adenylate analog TyrSA bound (Figure 18). Interestingly, the KMSKS loop was also found in a closed conformation in *LmajTyrRS*, even in the absence of ATP or tyrosyl adenylate analogs [105].

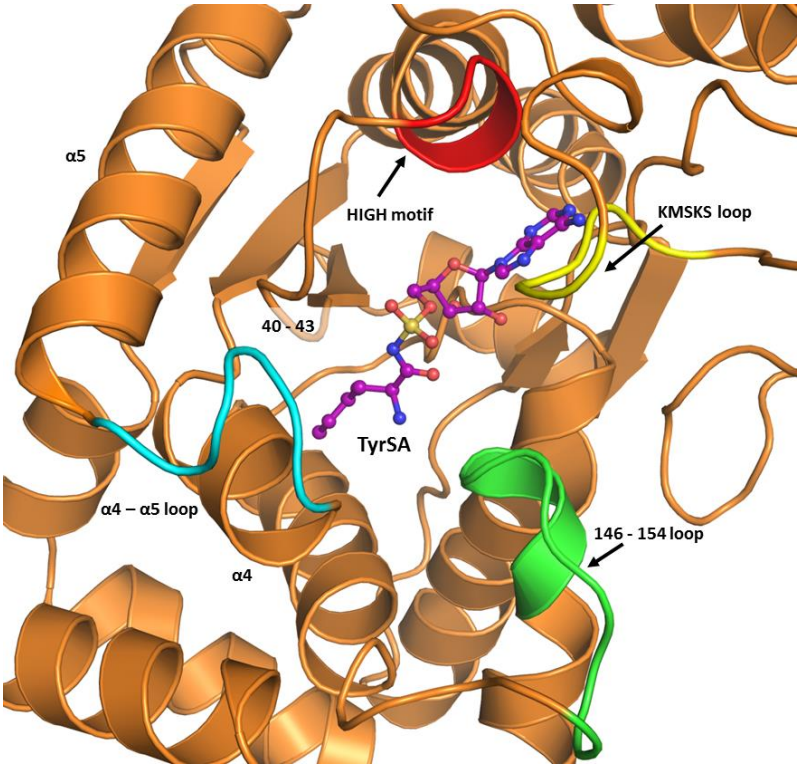


Figure 18. **Closed conformational state of the active site in the *LdonTyrRS*•TyrSA structure**. *LdonTyrRS* in orange while TyrSA in purple. The various loops and motifs that are well ordered in a closed conformation of the enzyme's active site are colored as follows: HIGH motif in red, KMSKS loop yellow, α4 - α5 loop cyan, and 146-154 loop green.

Not only the KMSKS loop can have an open and closed *conformation*, also the entire active site can be in an open or closed *state*. These two distinct conformational states have been described by Larson *et al.* [105] for the active site in the *LmajTyrRS*•Tyrosinol and *LmajTyrRS*•Fisetin crystal structures. In the functional catalytic domains of *LdonTyrRS*•TyrSA both chains adopt the closed state (Figure 18). In this state, the loop containing residues 146-154 is well-ordered and reaches the proximity of the sulfamoyl group of TyrSA, where the phosphate of tyrosyl adenylate or the α and β phosphates of ATP would be during the catalytic reaction by TyrRS. As in the *LmajTyrRS*

closed state, the *LdonTyrRS* α4 and the loop connecting this helix with α5 curl over the active site to interact with residues 40-43 preceding the HIGH motif (^{47}HIAQ^{49} in *LdonTyrRS*) involved in the enzyme's catalytic activity (Figure 18).

3.3.2 NbA structure and its interactions with *LdonTyrRS*

After various failed attempts to crystallize *LdonTyrRS*, we hypothesized that some flexibility in the four domain pseudo-dimer was impairing crystal formation. With the use of nanobodies as crystallization chaperones we aimed to freeze *LdonTyrRS* in a conformation more amenable for crystal growth, while the nanobody might also be able to establish favorable crystal contacts. Nanobodies are small compact single-domain fragments of the original heavy-chain camelid antibodies that retain their full antigen-binding capacity. Typically, three variable loops of β-strands, referred to as the complementarity determining regions (CDRs), are responsible for binding to the antigen. The usually long CDR3 loop of nanobodies can have a special significance since cryptic epitopes located in cavities or clefts of the antigenic protein are sometimes recognized by it [142]. However, in the *LdonTyrRS*•NbA•TyrSA complex the CDR3 loop in anti-*LdonTyrRS* llama nanobody A (NbA) does not interact extensively with *LdonTyrRS* as described below.

We tested 4 anti-*LdonTyrRS* nanobodies as crystallization chaperones before obtaining high quality *LdonTyrRS* crystals with NbA. The presence of the tyrosyl adenylate analog TyrSA was essential as well, and its interaction with the enzyme will be described later. The 1:1 *LdonTyrRS*•NbA complex was purified by size exclusion chromatography (SEC) and the binding of NbA to *LdonTyrRS* was verified through the shift of the 280 nm absorbance elution peak (Figure 19 A), corroborating the formation of a larger molecular weight species. The presence of both proteins *LdonTyrRS* and NbA in the larger molecular weight species was confirmed by SDS-PAGE

(Figure 19 B). In addition, native gel binding assays were performed supporting the interaction between the two proteins (not shown).

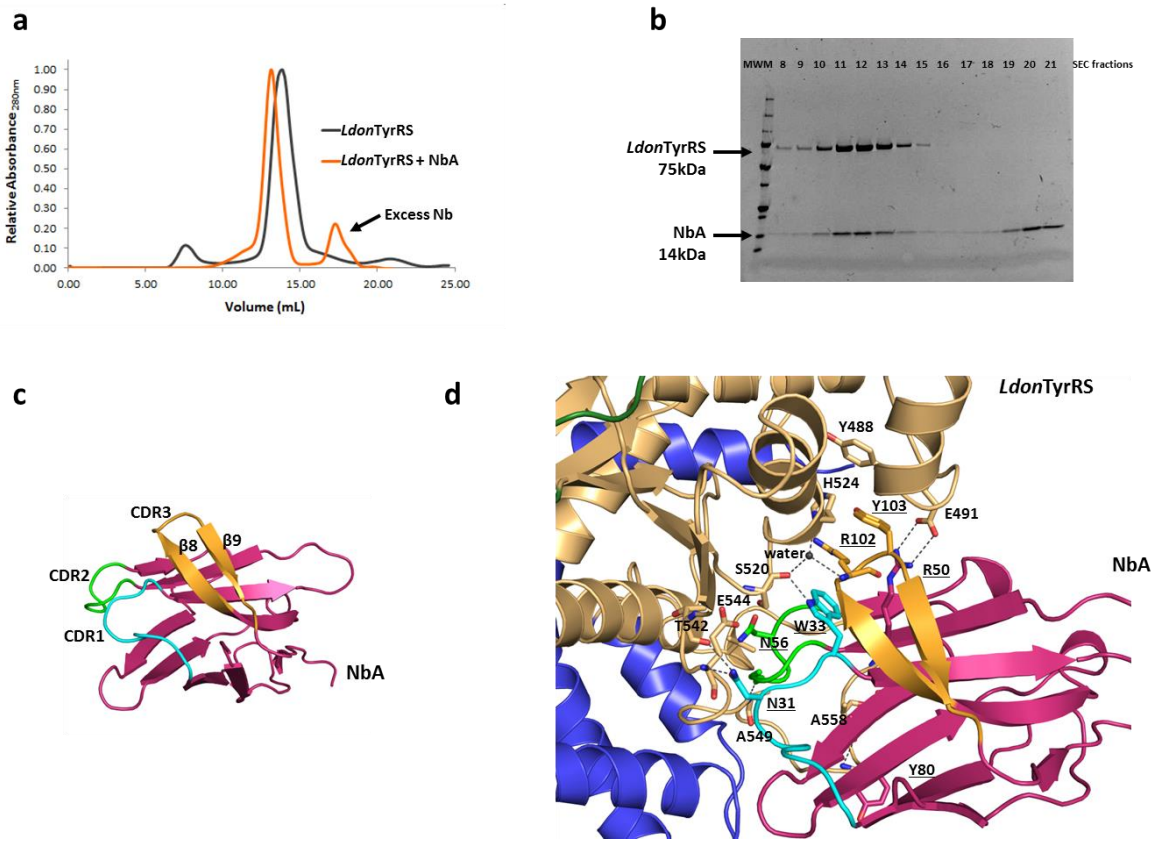


Figure 19. **Interactions between *LdonTyrRS* and NbA.** A) SEC elution profile of purified *LdonTyrRS* (grey) and *LdonTyrRS*•NbA complex (orange). B) SEC fractions were run in SDS-PAGE to corroborate the formation of the 1:1 complex. The low molecular weight present in the SEC chromatogram was proven to be excess NbA present in the mix *LdonTyrRS*/NbA that was loaded into the SEC column. C) NbA structure (pink) showing the three CDR loops: CDR1 (green), CDR2 (cyan) and CDR3 (gold). D) *LdonTyrRS*•NbA complex interactions. The underlined residues correspond to NbA. Most of the interactions are polar (doted lines). The three CDR loops are involved in interactions with the enzyme (in light orange).

NbA is a 14 kDa protein and consists of two β-sheets with a Greek key topology (Figure 19 C) held together by a disulfide bond (C22-C96). Interestingly, the CDR3 loop in NbA adopts an anti-parallel pair of β-strands connected by a short loop of 4 residues. Side chains of the β-strands of CDR3 do not interact with *LdonTyrRS*, but the short connecting loop does. In this case, CDR2 buries a larger solvent accessible surface area (511 $Å^2$) when interacting with *LdonTyrRS* than CDR3 (377 $Å^2$) (calculated by PISA [135]).

With a buried surface area (BSA) of 1947 $Å^2$, one NbA molecule interacts mainly with the non-functional CD of each *LdonTyrRS* chain in the crystal and makes a few additional contacts with the functional ABD. The corresponding estimated free energy of dissociation $\Delta^i G$ is -8 kcal/mol (calculated by PISA). The contacts between the variable region of NbA and *LdonTyrRS* are predominantly hydrophilic (Figure 19 D). Residues N31 and W33 of CDR1, residues R50 and G54 of CDR2, and, residue R102 of CDR3 (through a water molecule), make hydrogen bonds with *LdonTyrRS*. Side chain carbon atoms of residues N56 (CDR2) and Y103 (CDR3) are engaged in hydrophobic interactions with the enzyme. In addition, several frame work residues, not located in CDR loops, establish polar (R19, S69 and S71) and hydrophobic (Y80) interactions with the synthetase.

It is of interest to see if NbA has a stabilizing effect on *LdonTyrRS*, thereby promoting possibly crystal growth. The comparison with the *LmajTyrRS* structure is helpful here. The loop 541-571 connecting the non-functional CD with the functional ABD in *LdonTyrRS* exhibits a well-defined density, but the equivalent loop in the three *LmajTyrRS* crystal structures available is half or completely disordered [105]. Hence, it seems reasonable to conclude that NbA diminishes the flexibility of an exposed loop in *LdonTyrRS*.

A second reason why nanobodies can promote crystal growth is by generating additional crystal contacts. In the *LdonTyrRS*•NbA•TyrSA complex, NbA bound to chain A establishes crystallographic interactions with the other *LdonTyrRS* copy (Chain B) with a BSA of 792 $Å^2$. Moreover, NbA also contacts a symmetry-related NbA copy with a BSA of 969 $Å^2$. The nanobody is clearly engaged in extensive interactions with various protein chains in the crystal. Most likely NbA increases the probability of crystal growth by making numerous crystal contacts and by diminishing the flexibility of an exposed loop.

3.3.3. TyrSA binds to *LdonTyrRS* with its adenine ring near an extra pocket (EP)

The tyrosyl adenylate analog TyrSA (5'-O-[N-(L-tyrosyl)sulfamoyl]adenosine) (Figure 20 A) was necessary for obtaining well diffracting *LdonTyrRS* crystals. As measured by ATP depletion assays, TyrSA binds tightly to *LdonTyrRS* with an IC_{50} of 0.69 nM. The inhibitor was found to occupy the active site of the functional CD (Figure 20 B) in both *LdonTyrRS* chains in the asymmetric unit. The high affinity of the compound for the enzyme is most probably explained by the large number of interactions with *LdonTyrRS* active site residues, limiting thereby the flexibility of this region. This decrease in motility is likely responsible for the fact that TyrSA was crucial for obtaining well diffracting crystals of the *LdonTyrRS*•NbA complex.

The *LdonTyrRS* active site contains two critical pockets: the tyrosine binding pocket (YBP) where the tyrosyl group of TyrSA is situated, and the adenine binding pocket (ABP) where the adenine moiety of TyrSA binds (Figure 20 C). As evidenced through the use of PoseView as part of the Protein*Plus* structure-based modelling software tools [147, 148], the tyrosyl adenylate analog contacts many *LdonTyrRS* active site residues (BSA of 933 $Å^2$), mainly through hydrogen bonds, although a few hydrophobic interactions are present as well. Residues making hydrogen bonds

with the TyrSA tyrosyl group in the YBP are Y36, Y163, Q167, D170 and Q185. Residues G38, A72 and F75 are responsible for the hydrophobic interactions between enzyme and tyrosine moiety in the YBP.

A hydrogen bond is established between an oxygen atom belonging to the sulfamoyl group of TyrSA and the main chain nitrogen atom of residue E40 of *LdonTyrRS*. The residues responsible for hydrogen bond interactions with the ribose ring are D37, G182 and D184. In the ABP, hydrogen bonds with the adenine moiety are made by H210 and L213, while M212 makes hydrophobic contacts with the adenine ring.

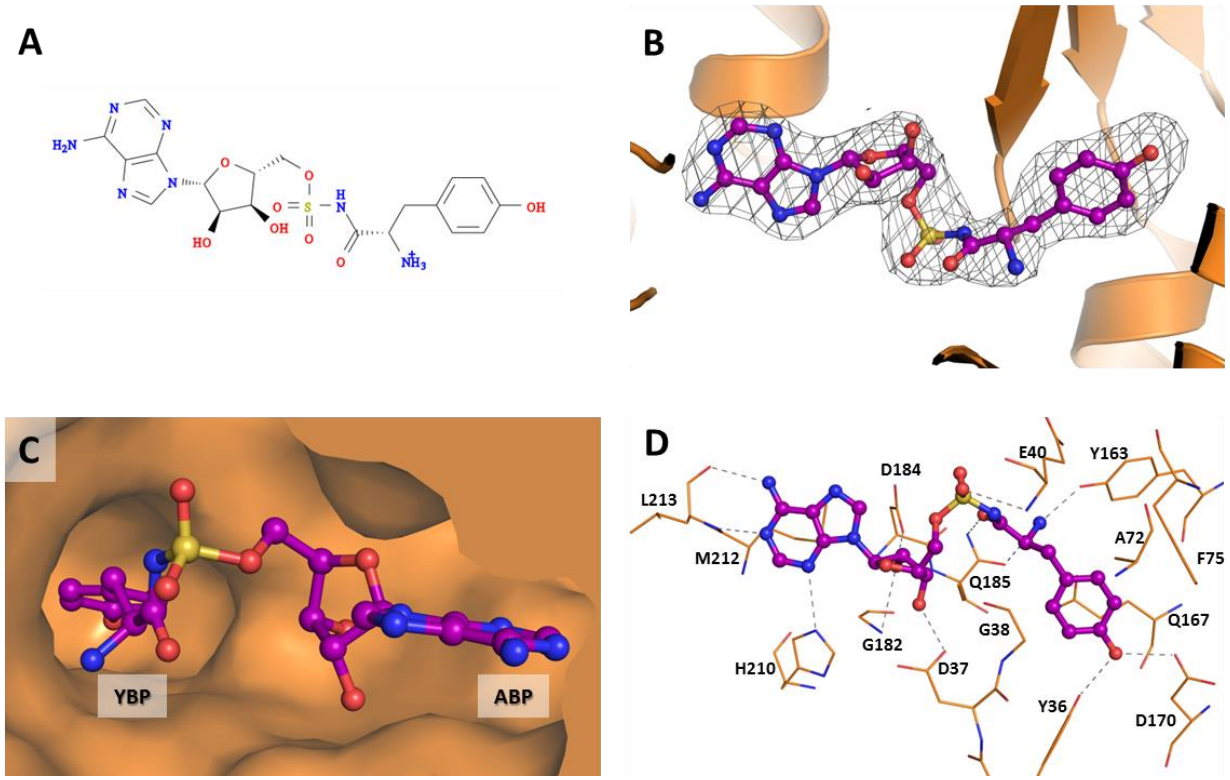


Figure 20. **TyrSA binding to *LdonTyrRS*.** A) Chemical structure of TyrSA. B) *LdonTyrRS*•NbA•TyrSA with difference electron density map calculated by omitting TyrSA, contoured at 3σ (positive density in grey, negative density in red). C) General features of TyrSA binding mode. The protein surface and the two pockets, tyrosine binging pocket (YBP) and adenine binding pocket (ABP), where the compound is bound are shown. D) Extensive hydrogen bond network in the *LdonTyrRS*/TyrSA interaction.

Further analysis of the active site near the ABP showed a previously not reported feature: a pocket in *LdonTyrRS* close to the TyrSA adenine moiety (Figure 21). This pocket, called "extra pocket" (EP), which is also present in the *L. major* TyrRS structure, is mainly lined by residues belonging to helix α3 as well as by the side chain of residue H210. Helix α3 atoms contributing to the pocket (Figure 21) are the carbonyl oxygen of A48, and side chain atoms of Q49, F52, K53 and N56. The EP of *LdonTyrRS* harbors two deeply buried water molecules. The most buried water molecule (Wat1) makes hydrogen bonds with the side chains of residues K53 and N56 and with Wat2. Wat2 engages in hydrogen bonding with the main chain carbonyl oxygen of A48 and with Wat1 and Wat3. The third water molecule (Wat3) interacts with (i) the ether oxygen atom of the TyrSA ribose ring, (ii) Wat2 in the EP, and (iii) the carboxylate group of residue D37 at the pocket entrance. The observation of this extra pocket (EP) in *LdonTyrRS* and *LmajTyrRS* turned out to be particularly interesting when comparing the TyrRS enzymes from the human host and a range of unicellular parasites.

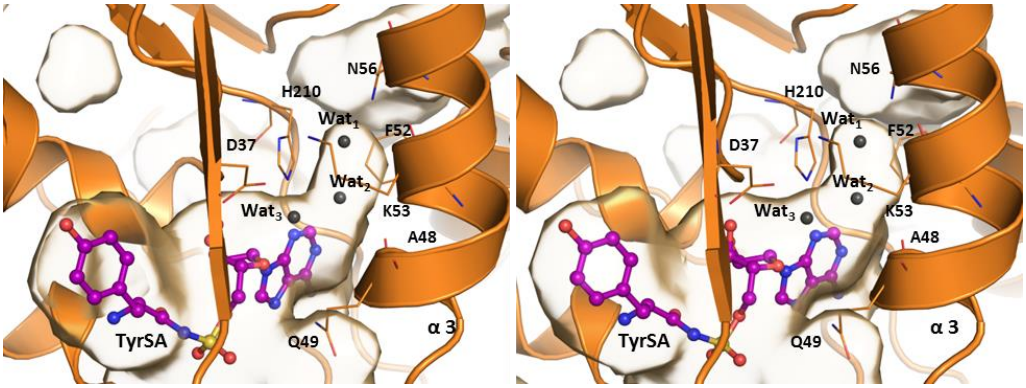


Figure 21. ***LdonTyrRS* residues forming the extra pocket (EP).** Stereo view of the *LdonTyrRS* catalytic site indicating the residues creating the EP. Most of the residues involved in the EP formation are part of helix α3, in addition to residues D37 and H210. TyrSA molecule in purple. Three water molecules, Wat1, Wat2 and Wat3 are located in the pocket and drawn as grey spheres.

3.3.4. The extra pocket (EP) of *LdonTyrRS* is absent in the human TyrRS enzymes

In order to discover new opportunities for drug design, the active sites of *LdonTyrRS* and the two *HsapTyrRS* enzymes were compared (Figure 22; for a list of TyrRS structures compared with *LdonTyrRS* in this paper see Table 5). While the YBPs are highly similar among the compared TyrRS structures, the analysis of the ABP reveals interesting amino acid differences between human and parasite enzymes as also described by Larson *et al*. [105] on the basis of the comparison of their *L. major* TyrS structure with the two human TyrRS enzymes. The basic conclusions of these authors are confirmed by our current *LdonTyrRS* structure.

Interestingly, the EP offers additional exploitable differences between the enzymes from parasite and host. When superimposing the *L. donovani* and human TyrRS structures, a substantial difference in position of α3 in the human enzymes *vs.* parasite enzyme is evident (Figures 22 and 23). Specifically, in cytosolic *HsapTyrRS*, the helix corresponding to the *LdonTyrRS* α3 helix is moved 3.0 Å towards the EP of *LdonTyrRS* (Figure 23 B), while in the mitochondrial *HsapTyrRS* structure this helix is moved by 4.6 Å towards the EP of *LdonTyrRS* (Figure 23 C, D). As result of this shift, the EP is absent in the human enzymes.

It is worthwhile to try to discern why the EP is absent in the human enzymes. It appears that helix α3 side chains forming the EP in *LdonTyrRS* are all different in the human enzymes. For instance, Q49 of *LdonTyrRS* is equivalent to Y52 and H91 in the human cytosolic and mitochondrial enzymes, F52 of *LdonTyrRS* to V54 and L93 in these two human homologs, K53 of *LdonTyrRS* to P55 and A94, and N56 of *LdonTyrRS* to K58 and G97, respectively. Actually, the *L. donovani* TyrRS residues from C35 to K59, comprising helix α3, and the *L. donovani* TyrRS residues V206 to L213, surrounding *Ldon*TyrRS H210, are essentially all different when comparing *LdonTyrRS* to the two human TyrRS enzymes (Figure 24). As a result of all these

changes, helix α3 shifts in the human enzymes compared to the *L. donovani* enzyme, and the EP is absent in the human homologs (Figure 22).

Table 5. Compared structures of tyrosyl tRNA synthetases.

PDB ID	Organism	Ligand(s)	Crystallized TyrRS	Reference
5usf	*Leishmania donovani*	TyrSA and nanobody	Full length	This publication
3p0h	*Leishmania major*	Fisetin	Full length	*105*
3p0j	*Leishmania major*	Tyrosinol	Full length	*105*
3vgj	*Plasmodium falciparum*	TyrAMP	Full length	*104*
4qbt	*Homo sapiens* Cytosolic	L-Tyrosine	Truncated (1-341)	*149*
2pid	*Homo sapiens* Mitochondrial	TyrSA	Full length	*150*
2dlc	*Saccharomyces cerevisiae*	Tyr-AMP analogue and $tRNA^{Tyr}$	Truncated (1-364)	*146*
1jii	*Staphylococcus aureus*	SB-219383	Full length	*97*
2jan	*Mycobacterium tuberculosis*	None	Full length	*151*
1vbm	*Escherichia coli*	TyrSA	Truncated (1-322)	*145*
1j1u	*Methanocaldococcus jannaschii*	L-Tyrosine and $tRNA^{Tyr}$	Full length	*152*

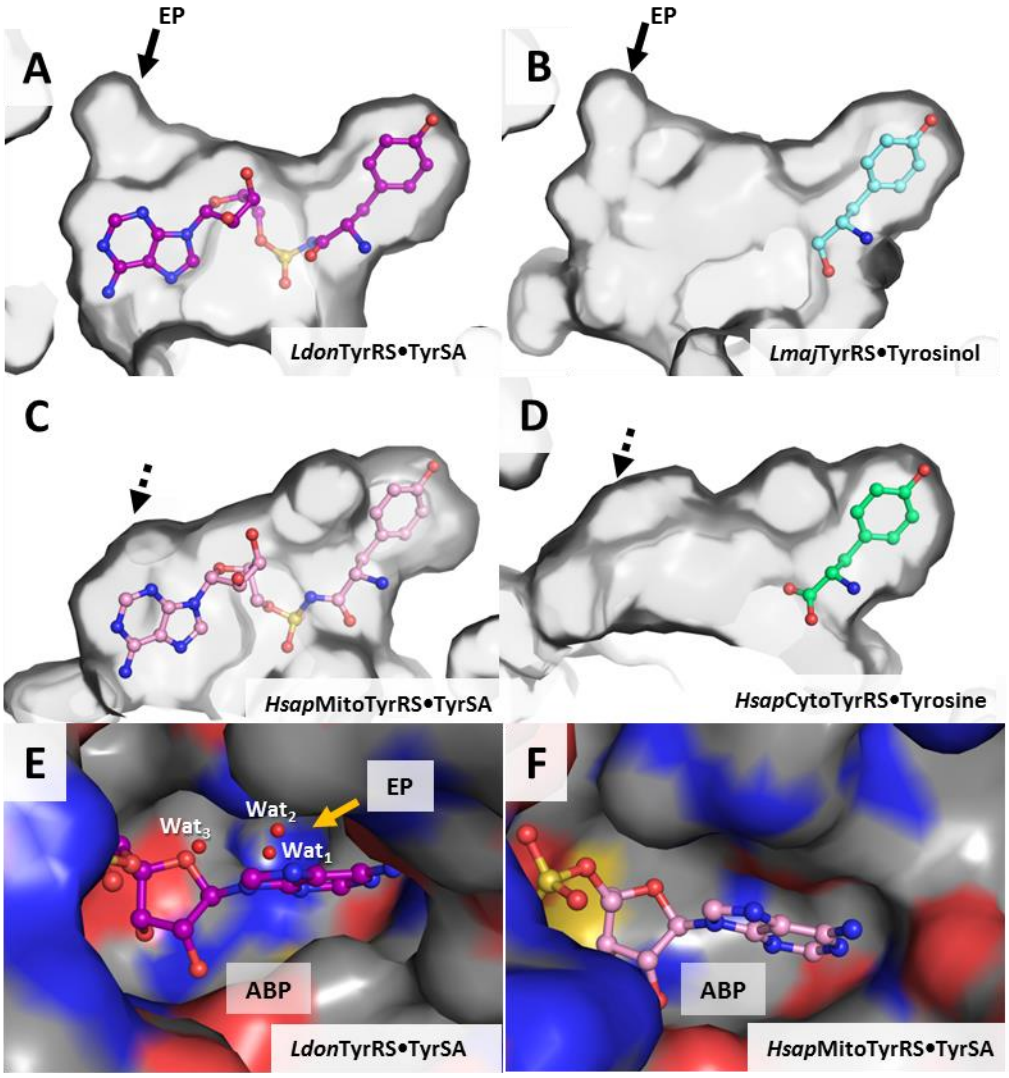


Figure 22. ***LdonTyrRS* catalytic pocket surface representation and comparison to human TyrRSs.** In panels A, B and E, the extra pocket (EP) is labeled. A) The *LdonTyrRS*•TyrSA structure in complex with TyrSA with purple carbon atoms. B) The *LmajTyrRS*•Tyrosinol structure with the tyrosinol molecule with cyan carbon atoms (PDB: 3POJ). The EP in the *L. major* enzyme is very similar to that in *L. donovani* TyrRS. C) *Hsap*Mito*TyrRS*•TyrSA structure with TyrSA molecule in pink (PDB: 2PID). The absence of the EP is indicated with a dashed arrow. D) *Hsap*Cyto*TyrRS*•Tyrosine structure with tyrosine molecule with green carbon atoms (PDB: 4QBT). The lack of the EP is indicated with a dashed arrow. E) Surface representation of *LdonTyrRS* binding pocket with bound TyrSA molecule with purple carbon atoms. Protein carbon atoms are colored grey, nitrogens blue, oxygens red, and sulfur atoms yellow. The adenine binding pocket (ABP) and EP are indicated. The three water molecules occupying the EP are shown as red spheres. Two of those (Wat1 and Wat2) are deeply positioned in the EP. F) Surface representation

of *Hsap*Mito*TyrRS* binding pocket in with bound TyrSA molecule with pink carbon atoms (PDB: 2PID) in the same orientation as the *LdonTyrRS* structure in E. Like in the human cytosolic enzyme binding pocket (not shown in surface representation), no EP is present.

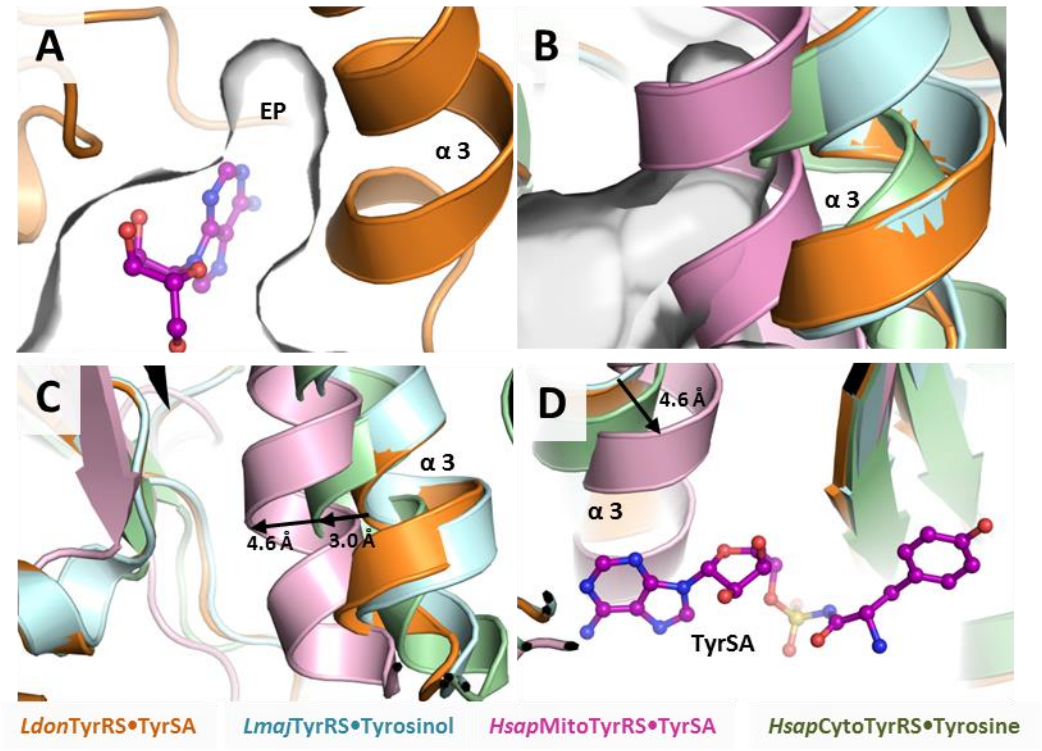


Figure 23. **Comparison of α3 in human and *Leishmania* TyrRS structures.** The superposition of structures was done to understand the absence of the EP in the human TyrRSs. A) *LdonTyrRS*•TyrSA structure in orange with TyrSA molecule in purple showing the EP surface in grey. B) Superposition of *LdonTyrRS*•TyrSA with *Hsap*Mito*TyrRS*•TyrSA (PDB: 2PID) in pink, *Hsap*Cyto*TyrRS*•Tyrosine (PDB: 4QBT) in light green and *LmajTyrRS*•Tyrosinol (PDB: 3POJ) in cyan; *LdonTyrRS* EP surface in grey. C) A 4.6 Å and 3.0 Å shift in α3 prevents the formation of the EP in *Hsap*Mito*TyrRS* and *HsapCytoTyrRS*, respectively. D) Another view of α3 superposition and its shift in *HsapTyrRS* structures.

3.3.5. A sequence fingerprint for the extra pocket (EP)

The seven residues forming the EP in *LdonTyrRS*: D37, A48, Q49, F52, K53, N56 and H210 can be used as an "EP fingerprint": D-x_{10}-AQ-x_2-FK-x_2-N-z-H, where x stands for any amino acid and

z for a variable large number of residues. This EP fingerprint is present in the TyrRS sequences of all *Leishmania* species analyzed (Figure 24). Hence, it is likely that the EP is present in all *Leishmania* TyrRS enzymes, as confirmed by a comparison of the *L. major* and *L. donovani* TyrRS structures (Figure 22 A, B). The EP fingerprint also occurs in other Trypanosomatids. The latter include the important human pathogens *Trypanosoma brucei* and *T. cruzi*, as well as *T. vivax*, causing nagana in cattle in sub-Saharan Africa. Given the absence of the EP in the two human homologs, this difference indicates interesting opportunities for arriving at compounds with higher affinity for Trypanosomatid than for the human tyrosyl-tRNA synthetases.

3.4. Discussion

The structure of the 75 kDa *LdonTyrRS* in complex with TyrSA has increased our insights in the architecture of this unusual member of the tRNA synthetase family. The enzyme consists of two pseudo-monomers, each containing two domains. The N-terminal pseudo-monomer is formed by a functional catalytic domain and a non-functional anticodon-binding domain, while the C-terminal pseudo-monomer is formed by a non-functional catalytic domain and a functional anticodon binding domain. A similar architecture had been observed for *L. major* TyrRS [105]. Comparison of the two crystallographically independent chains in our structure showed that the two pseudo-monomers can flex with respect to each other. This results in a considerable difference in position of the functional anti-codon binding domain at the C-terminus of the second pseudo-monomer when the N-terminal pseudo-monomers are superimposed (Figure 17). This flexibility might be essential in positioning $tRNA^{Tyr}$ properly with respect to the catalytic domain during the second catalytic step of the reaction catalyzed by leishmania TyrRS.

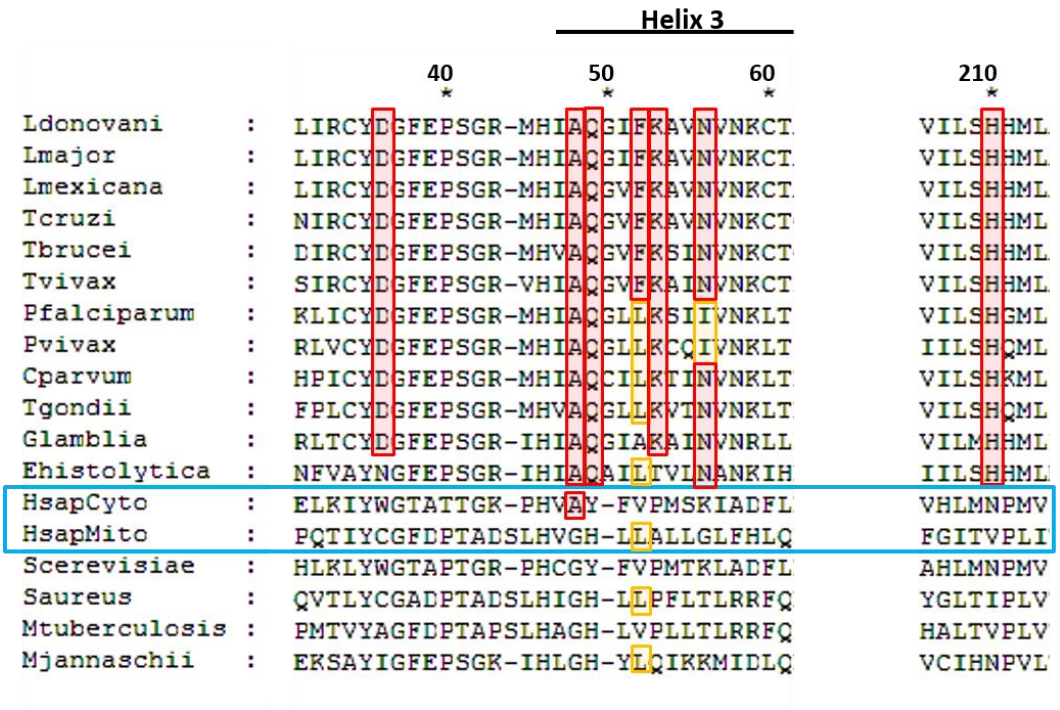


Figure 24. **Amino acid sequence alignment of selected species and the signature fingerprint of the extra pocket (EP).** The residues involved in the EP formation in *L. donovani* (top line) are indicated with red boxes. When the corresponding residues in other species are identical to those of the EP fingerprint in *LdonTyrRS* then these residues are also enclosed in a red box. Residues L and I at positions 4 and 6 of the *LdonTyrRS* EP fingerprint are indicated with orange boxes. The EP was not found in the available structures of human, yeast, bacteria or archaeal (*Methanocaldococcus jannaschii* [152]) TyrRSs which agrees with the lack of conservation of the EP fingerprint residues. The amino acid sequence alignment of *LdonTyrRS* with enzymes from selected pathogenic bacterial species shows that these bacterial TyrRS do not contain an EP, and comparisons of TyrRS crystal structures from (*S. aureus* and *M. tuberculosis* [97, 151]) with *LdonTyrRS* confirm that this indeed the case (not shown). The full species names are: Ldonovani = *Leishmania donovani*; Lmajor = *Leishmania major*; Lmexicana = *Leishmania mexicana*; Tcruzi = *Trypanosoma cruzi*; Tbrucei = *Trypanosoma brucei*; Pfalciparum = *Plasmodium falciparum*; Pvivax = *Plasmodium vivax*; Cparvum = Cryptosporidium parvum; Tgondii = *Toxoplasma gondii*; Glamblia = *Giardia lamblia*; Ehistolytica = *Entamoeba hystolitica*; HsapCyto = cytosolic *Homo sapiens*; HsapMito = mitochondrial *Homo sapiens*; Scerevisiae = *Saccharomyces cerevisiae*; Saureus = *Staphylococcus aureus*; Mtuberculosis = *Mycobacterial tuberculosis*; Mjannaschii = *Methanocaldococcus jannaschii*.

The binding mode of TyrSA to *LdonTyrRS* provides additional insight as to how the tyrosyl-adenylate adduct binds to the active site of this group of TyrRS enzymes. A new feature revealed by our analysis of the *LdonTyrRS* structure is the presence of an "extra pocket" (EP), close to the adenine binding pocket and filled by three mutually interacting water molecules. This EP is present as well in the *L. major* TyrRS structure [105] and, based on amino acid sequence comparisons, also in other Trypanosomatids. Several of the latter, such a *T. brucei* and *T. cruzi*, are major global human pathogens causing death and disease in particular in low-income countries [23, 24]. Since the EP is absent in the two human TyrRS enzymes, this difference between enzymes from parasite and human host indicates opportunities to arrive at inhibitors with high affinity and selectivity.

It is of interest to evaluate if the EP is present in the TyrRS from other medically important parasites. The enzyme from the protozoan *Plasmodium falciparum,* the major causative agent of malaria, has been the subject of several studies, including the determination of the crystal structure of cytosolic *P. falciparum* TyrRS (*PfalTyrRS*) [104]. The EP fingerprint is present in *PfalTyrRS* with two changes: the F at the fourth position in the fingerprint is in *PfalTyrRS* an L, which conserves the hydrophobic side chain characteristic at this position, and the N at the sixth position is an I in *PfalTyrRS* (Figure 24). A comparison of the structure of *LdonTyrRS* with the crystal structure of *P. falciparum* TyrRS [104], reveals that in the latter enzyme the EP is indeed present (Figure 25). In this case, the EP harbors only one deep water molecule, the equivalent of Wat2 in the *LdonTyrRS* EP. The absence of Wat1 is due to the fact that residue N56, in the back of the EP of *LdonTyrRS* (Figure 22 B), is substituted by I80 in the *P. falciparum* enzyme, providing a more hydrophobic environment within the pocket and no hydrogen-bonding partner for Wat1. Based on the amino acid sequence alignment (Figure 24), this same feature is present in *P. vivax*, another important

malaria parasite. These comparisons suggest that there are opportunities to arrive at high affinity *PfalTyrRS* inhibitors with good selectivity. This is extra interesting since Khan [140] has recently emphasized the opportunities of TyrRS as a drug target for malaria.

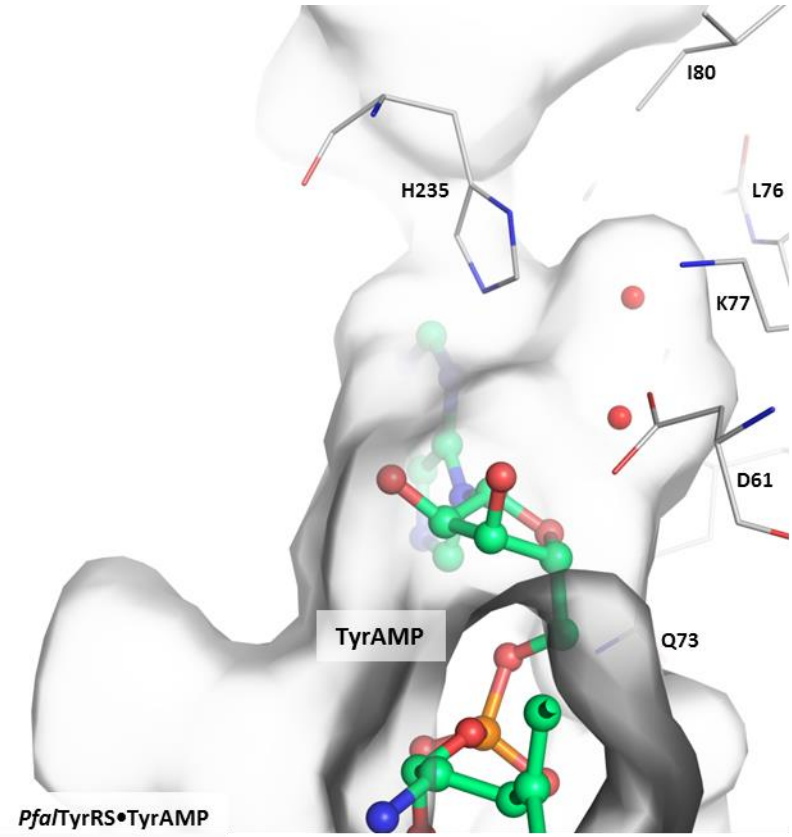


Figure 25. ***P. falciparum*** **TyrRS residues forming the extra pocket (EP).** Surface representation of the EP in the *PfalTyrRS*•TyrAMP crystal structure [104] (PDB: 3VGJ). Protein carbon atoms are colored grey, nitrogens blue, oxygens red and phosphorous orange. All but two (L76 and I80) of the residues involved in the EP formation are shared with trypanosomatids. *PfalTyrRS* harbors one deep water in the EP (equivalent to Wat2 in *LdonTyrRS*) and a water molecule at the EP entrance (equivalent to Wat3 in *LdonTyrRS*), shown as red spheres. TyrAMP carbon atoms are shown in green.

Turning to other protozoa causing considerable human suffering and death across the globe, the comparison of EP fingerprints in *Cryptosporidium parvum*, *Toxoplasma gondii* and *Giardia lamblia* reveals that in these pathogens there is only a single residue difference in the

fingerprint: the F at the fourth position is an L, I or A in these three parasites, respectively. Hence the EP in these protozoa is even closer to that of *LdonTyrRS* than the EP in the two *P. falciparum* species. The N at the sixth position of the fingerprint is maintained so that it is very likely that in these three species the EP is present with Wat1 deep in the pocket. Consequently, the EP might enable the development of effective and selective TyrRS inhibitors for treating cryptosporidiosis, toxoplasmosis and giardiasis. Although in *Entamoeba histolytica* the EP fingerprint is less perfectly conserved, with differences in the first (D to N), fourth (F to L) and fifth (K to T) positions with respect to *LdonTyrRS*, two of these three side chain changes are very conservative. Hence, an exploitable EP might also exist in the TyrRS of this pathogen.

In conclusion, the structure of the *LdonTyrRS*•NbA•TyrSA complex elucidated revealed different mutual orientations of its two pseudo-monomers. The comparison of the *Ldon*TyrRS structure to both human homologs showed major differences compared in the active site, in particular the presence of an "extra pocket" in the parasite enzyme. The latter feature is also found in the TyrRSs from other unicellular parasites and could be a key element in the development of novel compounds for treating diseases caused by a wide range of important pathogenic protozoa.

Acknowledgments

We like to thank Ethan Merritt for stimulating discussions. Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award number R01AI084004 (to WGJH) and R01AI097177 (to FSB and EF). We thank Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI), and the Research Foundation Flanders (FWO) for their support to the Nanobody discovery. We are grateful to Nele Buys for the technical assistance during Nanobody discovery. We acknowledge the support of a Fulbright Fellowship to X.B.-A. We thank Robert Steinfeldt for providing support for computing environment at the Biomolecular Structure Center of the University of Washington. Crystallography performed in support of the work benefitted from remote access to resources at the Stanford Synchrotron Radiation Lightsource supported by the U.S. Department of Energy Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515 and by the National Institutes of Health (P41GM103393). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Chapter 4

Trypanosoma brucei methionyl-tRNA synthetase structures with inhibitors: structural biology assisting the drug design process against human African trypanosomiasis

A compilation of the structural biology work done as part of collaborative SGDD projects published in Devine *et al.* (2017) [108], Huang *et al.* (2016) [109] and Huang *et al.* (2017) [110].

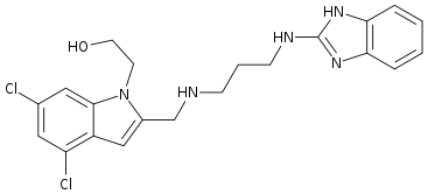
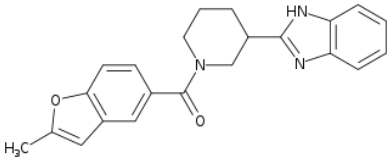
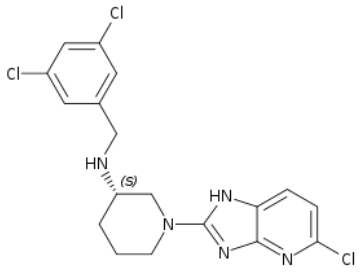
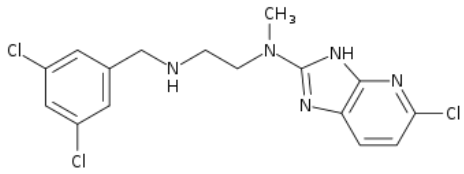
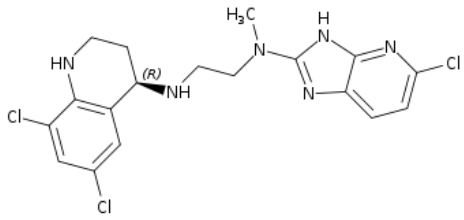
4.1. Introduction

The initial crystal structures of *Trypanosoma brucei* methionyl-tRNA synthetase (*Tbru*MetRS) in complex with methionine, methionyl adenylate (Met-AMP) and inhibitors revealed that *Tbru*MetRS undergoes extensive conformational changes upon inhibitor binding by conformational selection [90]. Obtaining structures of *Tbru*MetRS with inhibitor bound was not straightforward, however, and required that first high-quality crystals with the substrate methionine was obtained, to be followed by rapid soaking of the *Tbru*MetRS•Met crystals in a solution of the inhibitor plus DMSO. Koh *et al.* (2012) [90], described the remarkable result that the methionine bound to *Tbru*MetRS in only one of the two chains (chain B) in the asymmetric unit of the *Tbru*MetRS•Met crystal could be replaced through soaking with *Tbru*MetRS inhibitors (representative compounds can be found in Chapter 1, Table 1). Upon inhibitor binding, methionine remained bound to chain A maintaining the *Tbru*MetRS•Met complex while chain B adopted a new conformation while forming a *Tbru*MetRS•Inhibitor complex. Interestingly, upon

movement of several hydrophobic side chains that delineated the methionine pocket in the *Tbru*MetRS•Met complex, the methionine binding pocket became enlarged in what was called the "enlarged methionine pocket" (EMP) in the *Tbru*MetRS•Inhibitor structure. At the same time, a new pocket, non-existent in the *Tbru*MetRS•Met complex, named the auxiliary pocket (AP), was formed in *Tbru*MetRS•Inhibitor structures [90] . *Tbru*MetRS inhibitors had a common structure of two aromatic ring systems linked by a linear linker (Table 1). Each of the EMP and AP pockets in the *Tbru*MetRS•Inhibitor structure was able to accommodate one of the ring systems [90, 91]. Some of these inhibitors showed to be very potent *in vitro* against *Tbru*MetRS and also detrimental for the *T. brucei* parasites [88, 91, 92]. However, several pharmacological proprieties had shortcomings and needed to be optimized and therefore other scaffolds were explored in order to find better compounds.

This chapter is the compilation of three research publications in which structural studies supported drug development projects targeting MetRS from *T. brucei* parasites in collaboration with computational biology, medicinal chemistry and parasitology groups inside and outside the University of Washington. The *Tbru*MetRS•Inhibitor structures solved as part in these publications are listed in Table 6.

Table 6. *Tbru*MetRS•Inhibitor structures solved and published as part of this dissertation.

PDB code	Compound identifier	Compound structure	Publication
5TQU	NEU-1053 (compound **1** in [108])		Devine *et al*. (2017) [108]
5J5A	70786556 (compound **1** in [109])		Huang *et al*. (2016) [109]
5J58	1856 (compound **16** in [109])		Huang *et al*. (2016) [109]
5J59	1893 (compound **31** in [109])		Huang *et al*. (2016) [109]
5V49	1917 (compound **13** in [110])		Huang *et al*. (2017) [110]

4.1.1. Structural studies revealed surprising interactions of a new inhibitor

NEU-1053 (Table 6) was identified in a kinase-targeted *high-throughput* screening (HTS) [153] effort against *T. brucei* from the GlaxoSmithKline screening collection. NEU-1053 was able to cure a murine bloodstream infection of *T. brucei rhodesiense*. The compound had an unknown target but its structure resembled that of proven *Tbru*MetRS inhibitors [88]. In Devine *et al*. (2017) [108] (pages 95-106 of this dissertation), it was demonstrated that *Tbru*MetRS was indeed the target for NEU-1053 and showed an IC_{50} against *Tbru*MetRS that was lower than the limit of detection of the assay (<25 nM). The compound was very potent against the parasites as well, with an EC_{50} of 0.96 nM. A structure-activity relationship (SAR) of a series of compounds derived from NEU-1053 (SB-443342) was performed. To further support the target identification and understand the binding mode of NEU-1053 to *Tbru*MetRS, the *Tbru*MetRS•NEU-1053 complex crystal structure was elucidated. Rapid soaking of *Tbru*MetRS•Met crystals in a solution of NEU-1053 was again succesful.[108]. The compound appeared to bind to *Tbru*MetRS essentially in the same way other *Tbru*MetRS inhibitors did, with one of the two ring systems each occupying the EMP and AP pockets. Interestingly, the compound established a novel interaction through the OH group in its hydroxyethyl substituent with the catalytic aspartate residue (D287) in the *Tbru*MetRS catalytic site. NEU-1053 has since become a lead compound for further optimization to improve its central nervous system (CNS) permeability (crucial to treat late-stage HAT) and pharmacokinetics, as well as a new scaffold to inspire further drug development against HAT.

4.1.2. The structure of a weak inhibitor in complex with *Tbru*MetRS provided structural insights for the discovery of two very potent inhibitor series with improved pharmacological properties

Compound **1** (nomenclature as used in Huang *et al*. 2016 [109], Table 6), which had been identified previously through HTS of the NIH Molecular Libraries Small Molecule Repository [117] as a micromolar *Tbru*MetRS inhibitor, was optimized through SGDD in Huang *et al*. 2016 [109] (pages 107-118 of this dissertation). The screening hit had shown 78 % and 20 % inhibition in a *Tbru*MetRS aminoacylation assay at 10 and 1 µM respectively. In trying to discover new scaffolds and compounds with good pharmacological proprieties to feed our collaborative drug development initiative against HAT, characteristics of compound **1**, were combined with other features of a previously reported potent *Tbru*MetRS inhibitor, compound **2** (nomenclature as used in Huang *et al*. 2016 [109]). The optimization was based on the analysis of the crystal structures of *Tbru*MetRS•**1** and *Tbru*MetRS•**2** complexes. Similar to earlier *Tbru*MetRS inhibitors (Table 1), both compounds consist of two ring systems, connected by a linker. The benzofuran moiety of compound **1** binds to the EMP and the benzimidazole moiety occupies the AP, this time connected by a cyclic linker (Table 6). The *Tbru*MetRS•**1** complex structure revealed a void in the EMP that was occupied by one of the chlorines of the dichloro-benzene moiety in the *Tbru*MetRS•**2** complex structure. At the same time, the longer linear linker in compound **2** allowed the benzimidazole moiety to bind deeper into the AP and suggested the possibility to exploit an extra interaction with residue L456 at the bottom of the AP. In addition, the ring in the linker of compound **1** establishes extensive hydrophobic interactions with *Tbru*MetRS residue Y250. Combining the dichloro-benzene moiety of compound **2** with the cyclic linker of compound **1** and a chlorine-substituted benzimidazole moiety (reaching residue L456) gave rise to a new scaffold that was further optimized resulting in

two novel *Tbru*MetRS series of potent inhibitors discovered: a cyclic linker and a linear linker (formed after opening the ring in the linker) series. The new series showed good potency, high selectivity, promising PK properties as well as brain permeability. Compounds **16** and **31** (Table 6) are the best compound in each series, with IC_{50} values against the *Tbru*MetRS enzyme that were lower than the limit of detection of the assay and were both able to kill the parasites with EC_{50}s of 39 and 22 nM, respectively. The new lead compounds showed low toxicity to mammalian cells and have also promising PK properties after oral dosing in mice. Additionally, compound **31** had moderately good brain permeability. The crystal structures of the *Tbru*MetRS•**16** and *Tbru*MetRS•**31** complexes show similar binding modes in the EMP and AP, in addition these compounds take advantage of novel interaction with *Tbru*MetRS residue Y250.

4.1.3. Structural evidence of a better fragment bound to the *Tbru*MetRS EMP

Further optimization of the lead compounds **16** and **31** discovered in Huang *et al*. 2016 [109] was carried out by Huang *et al*. 2017 [110], (pages 119-124 of this dissertation). In a SAR study, the inhibitory activities of new *Tbru*MetRS inhibitors exploiting the EMP were described [110]. The superior potency of bacterial MetRS [154, 155] and *Tbru*MetRS inhibitors [93] that contained a tetrahydroquinoline group occupying the methionine binding site, inspired the substitution of the dichlorophenyl moiety in the most potent compounds of the linear and cyclic series described above [109] by a tetrahydroquinoline group. The resulting compounds **11** and **13** (nomenclature as used in Huang *et al*. 2017 [110], Table 6) exhibited an EC_{50} of 57 nM and 4 nM, respectively. The crystal structure of the *Tbru*MetRS•**13** complex was determined which showed that the 6,8-dichloro-tetrahydroquinoline ring occupied the EMP, as expected. The structural analysis showed that both compound **13** and its predecessor containing a 3,5-dichloro-benzyl group (compound **2**

[110]) bind similarly to *Tbru*MetRS. The *Tbru*MetRS•**13** crystal structure is now available to explain the SAR data generated as part of the optimization. Additionally, compound **13** showed low toxicity to mammalian cell lines and moderately good brain penetration in mice.

In conclusion, Huang *et al*. 2017 [110] identified the 6,8-dichloro-tetrahydroquinoline ring to be potentially a better fragment for binding the EMP than the 3,5-dichloro benzyl group in the cyclic and linear linker series of *Tbru*MetRS inhibitors and the crystal structure of *Tbru*MetRS•13 complex will help guide future inhibitor design.

ACS Infectious Diseases

Article

pubs.acs.org/journal/aidcbc

From Cells to Mice to Target: Characterization of NEU-1053 (SB-443342) and Its Analogues for Treatment of Human African Trypanosomiasis

William G. Devine,[†] Rosario Diaz-Gonzalez,[§] Gloria Ceballos-Perez,[§] Domingo Rojas,[§] Takashi Satoh,[†] Westley Tear,[†] Ranae M. Ranade,[#] Ximena Barros-Álvarez,[⊥,Δ] Wim G. J. Hol,[⊥] Frederick S. Buckner,[#] Miguel Navarro,[§] and Michael P. Pollastri*[,†]

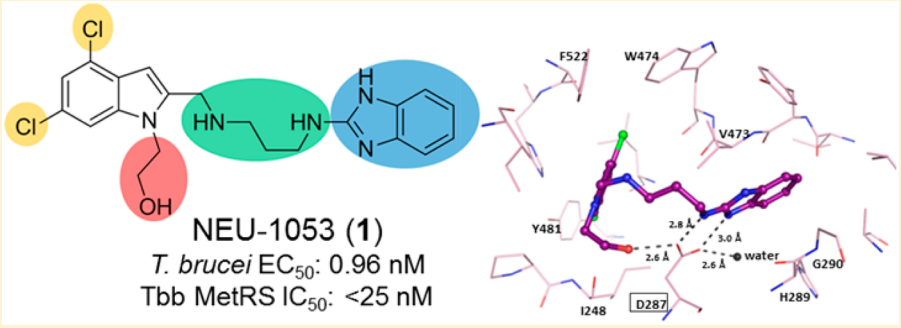
[†]Department of Chemistry & Chemical Biology, Northeastern University 360 Huntington Avenue, Boston, Massachusetts, United States

[§]Instituto de Parasitologia y Biomedicina "Lopez-Neyra", Consejo Superior de Investigaciones Cientificas (CSIC), Granada 18100, Spain

[#]Departments of Medicine and [⊥]Biochemistry, University of Washington, Seattle, Washington 98195, United States

[Δ]Laboratorio de Enzimología de Parásitos, Facultad de Ciencias, Universidad de los Andes, Mérida, Venezuela

Supporting Information



ABSTRACT: Human African trypanosomiasis is a neglected tropical disease that is lethal if left untreated. Existing therapeutics have limited efficacy and severe associated toxicities. 2-(2-(((3-((1*H*-Benzo[*d*]imidazol-2-yl)amino)propyl)amino)methyl)-4,6-dichloro-1*H*-indol-1-yl)ethan-1-ol (NEU-1053) has recently been identified from a high-throughput screen of >42,000 compounds as a highly potent and fast-acting trypanocidal agent capable of curing a bloodstream infection of *Trypanosoma brucei* in mice. We have designed a library of analogues to probe the structure−activity relationship and improve the predicted central nervous system (CNS) exposure of NEU-1053. We report the activity of these inhibitors of *T. brucei*, the efficacy of NEU-1053 in a murine CNS model of infection, and identification of the target of NEU-1053 via X-ray crystallography.

KEYWORDS: *Trypanosoma brucei, methionyl-tRNA synthetase, medicinal chemistry*

Insect-borne trypanosomal diseases are a menace to human health. Human African trypanosomiasis (HAT, or sleeping sickness) is a neglected tropical disease caused by two subspecies of *Trypanosoma brucei* for which current therapeutics are toxic and inconvenient. Although there are two compounds in clinical trials, SCYX-7158[1] and fexinidazole,[2] given the clinical failure rate for infectious diseases,[3] it is prudent to continue the search for new drugs.

To that end, we recently described a high-throughput screening (HTS) campaign performed as part of an industry−academic partnership between GlaxoSmithKline, Spanish National Research Council (CSIC), and Northeastern University, in which we uncovered 798 inhibitors of *T. brucei* cellular proliferation. We also reported in vitro drug metabolism, physicochemical properties, and pharmacokinetics data, plus kinase selectivity data for key analogues.[4] Included in that paper was NEU-1053 (SB-443342, **1**), a singleton compound identified in the screen that showed rapid and irreversible proliferation inhibition of *T. brucei*, showed good plasma exposure, and cured a bloodstream infection in a *T. brucei rhodesiense* mouse model of HAT. As a singleton hit with no other analogues included in the screening campaign, there was no structure−activity relationship (SAR) information apparent from the HTS. Besides looking to better understand the SAR of this series, we wished to explore the various structural regions of the compound to identify effective analogues with more attractive physicochemical properties. We report those efforts here, definitively demonstrate the

Received: November 30, 2016
Published: January 22, 2017

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

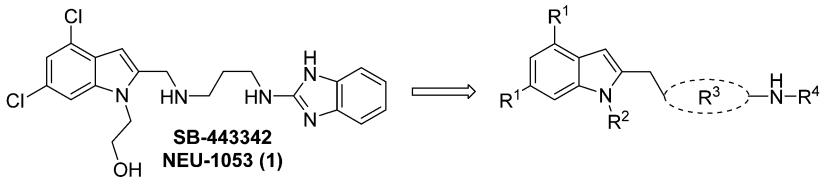
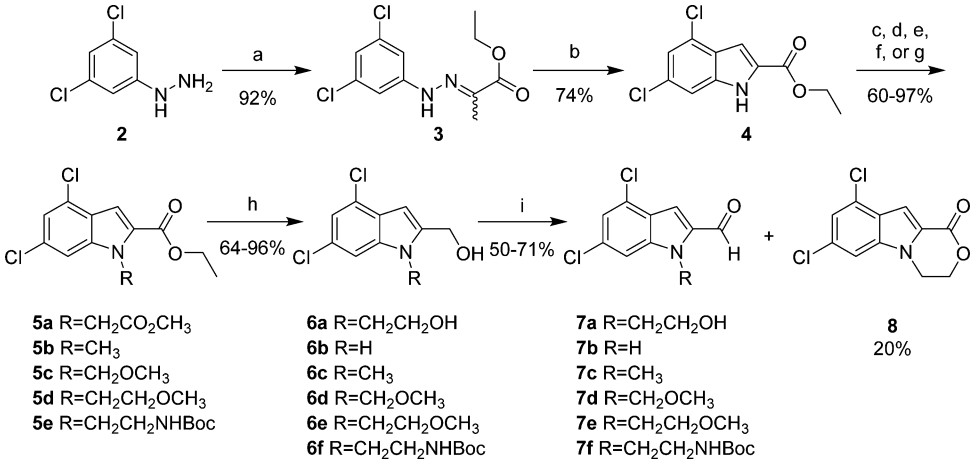


Figure 1. 1 SAR regions of interest.

Scheme 1. Synthesis of the 3,5-Dichloroindole Western End of 1 and Its Analogues[a]



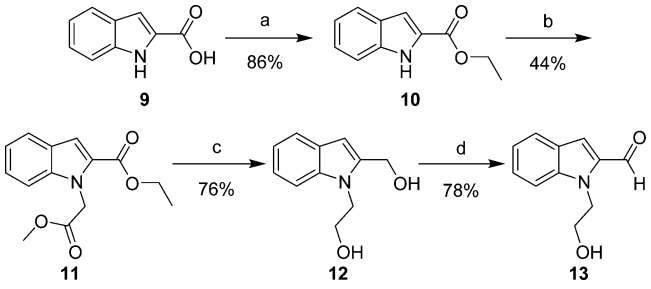
[a]Reagents and conditions: (a) ethyl pyruvate, EtOH, reflux, 1 h; (b) PPA, 130 °C, 10 min; (c) methyl bromoacetate, K_2CO_3, DMF, rt, overnight; (d) NaH, DMF, 0 °C, 30 min; then CH_3I, 0 °C → rt, 3 h; (e) NaH, DMF, 0 °C, 30 min; then CH_3OCH_3Cl, 0 °C → rt, 3 h; (f) NaH, DMF, 0 °C, 30 min; then $CH_3OCH_2CH_2Br$, 0 °C → rt, 18 h; (g) K_2CO_3, DMF, $BocNH(CH_2)_2OTs$, rt, overnight; (h) DIBAL, THF, rt, 1 h; (i) MnO_2, THF, rt, 2 h.

mechanism of action for **1**, and describe the results of an in vivo efficacy experiment for this compound in a murine model of stage 2 HAT.

RESULTS

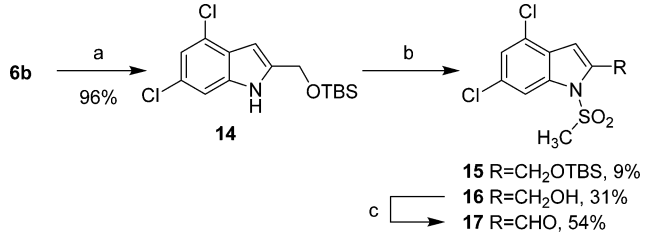
The overall strategy for the SAR exploration of **1** is shown in Figure 1. First, noting that the chlorine atoms on the indole provide substantial contribution to molecular weight and lipophilicity, we were interested to understand their importance. Second, the hydroxyethyl substituent on the indole nitrogen seemed to be a potential metabolic liability, and exploration of this region was needed. Third, the linker between the two aromatic systems in the molecule needed to be probed in terms of length, vector, and rigidity. Lastly, looking to reduce the size of the molecule, we wished to better understand the requirements of the 2-aminobenzimidazole eastern end by replacing the benzimidazole functionality. We first describe the synthesis of these analogues and will then discuss the impact on potency and properties for this series.

Scheme 2. Synthesis of the Western End of Dechloro-1 Analogues[a]



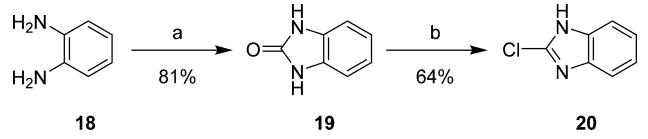
[a]Reagents and conditions: (a) H_2SO_4, EtOH, reflux, overnight; (b) methyl bromoacetate, K_2CO_3, DMF, rt, 3 h; (c) $LiAlH_4$, THF, 0 °C, 30 min, rt, 1.5 h; (d) MnO_2, CH_2Cl_2, rt, overnight.

Scheme 3. TBS-Protected *N*-Substituted Indole Synthesis[a]



[a]Reagents and conditions: (a) TBSCl, imidazole, CH_2Cl_2, rt, 1 h; (b) NaH, DMF, rt, 30 min; CH_3SO_2Cl, rt, 1 h; (c) MnO_2, CH_2Cl_2, rt, 2 h.

Scheme 4. Chloro-1*H*-benzo[*d*]imidazole Synthesis[a]

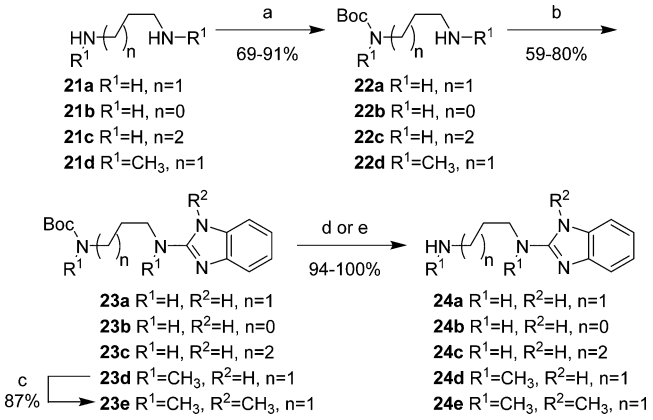


[a]Reagents and conditions: (a) urea, $(CH_2OH)_2$, 140 °C, 1 h, 170 °C, 7 h; (b) $POCl_3$, reflux, 2 h.

The synthesis of **1** and its 3,5-dichloroindole analogues commenced with the condensation of hydrazine **2** and ethyl pyruvate to yield a mixture of *E* and *Z* isomers of **3** (Scheme 1). Cyclization of **3** generated 3,5-dichloroindole **4**, which could be

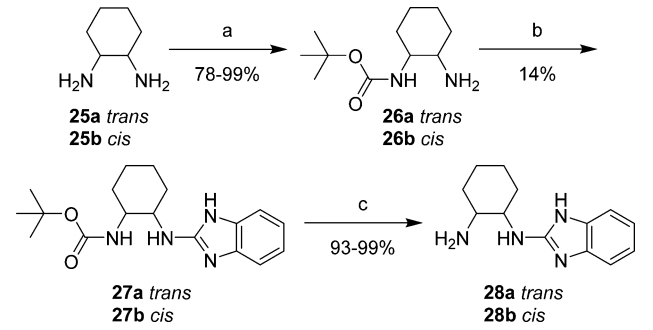
DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

Scheme 5. Acyclic Eastern Half Synthesis[a]



[a]Reagents and conditions: (a) Boc_2O, $CHCl_3$, 0 °C → rt, overnight; (b) **20**, toluene, μW, 150 °C, 6 h; (c) NaH, DMF, CH_3I, rt, 1 h; (d) TFA, CH_2Cl_2, rt, overnight; (e) H_2O, μW, 150 °C, 1 h.

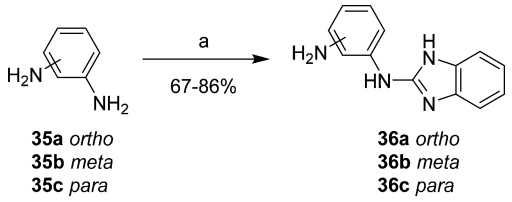
Scheme 6. 1 Cyclohexyl Linker Replacement Synthesis[a]



[a]Reagents and conditions: (a) Boc_2O, CH_2Cl_2, 0 °C → rt, overnight; (b) **20**, toluene, μW, 150 °C, 6 h; (c) TFA, CH_2Cl_2, rt, overnight.

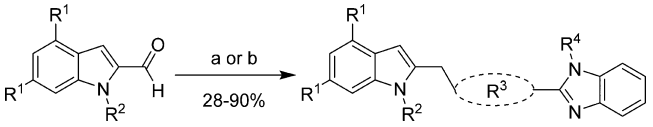
recrystallized from H_2O/ethanol. Alkylation with the appropriate alkyl halides or tosylates[5,6] and subsequent reduction with DIBAL gave corresponding *N*-substituted 2-(hydroxymethyl)indoles **6a–f** in good to high yields. Oxidation of the benzylic alcohol with freshly prepared MnO_2 following the Attenburrow procedure[7] produced aldehydes **7a–f** in 50–71% yield. The oxidation of diol **6a** also generated lactone **8** in 20% yield, likely stemming from overoxidation of the intermediate lactol. This material could be reverted to **6a** via reduction or used for further analogue synthesis (vide infra). The nonhalogenated **13** was synthesized in a similar manner from the commercially available indole **9** (Scheme 2).

Scheme 8. Phenylenediamine Linker Synthesis[a]



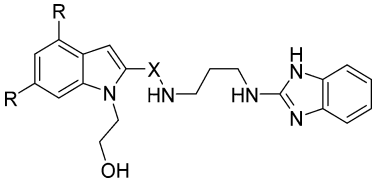
[a]Reagents and conditions: (a) **20**, xylenes, μW, 200 °C, 15 min.

Scheme 9. Synthesis of 1 Analogues via Reductive Amination[a]



[a]Reagents and conditions: (a) amine, KOAc, $NaBH_3CN$, CH_3OH; (b) amine, AcOH, $NaBH_3CN$, CH_3OH.

Table 1. Activity of Modified Indole Analogues of 1 against *T. brucei*



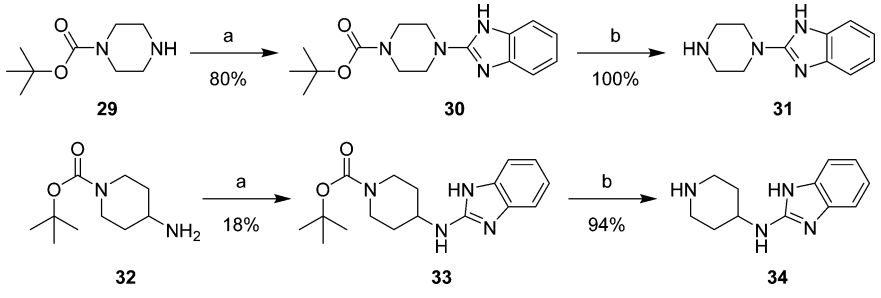
compound	R	X	*T. brucei* EC_{50}[a] (nM)	MRC5-SV2 TC_{50}[b] (μM)	selectivity (TC_{50}/EC_{50})
1	Cl	CH_2	0.96	12	
39	Cl	CO	500	18	36
46	H	CH_2	75	27[c]	360

[a]SD < 0.20. [b]SD < 1.5 unless otherwise noted. [c]SD = 9.3.

The alcohol moiety of compound **6b** was protected as the TBS ether, and the sulfonylation of this compound was performed using the same protocol for **5c–e**; this reaction also effected desilylation, providing **16** in 31% yield (Scheme 3). The free alcohol was oxidized to the corresponding aldehyde with MnO_2 in modest yield.

Aminoalkyl-2-aminobenzimidazole intermediates were prepared as shown in Schemes 4–7. Briefly, **20** was formed via heating of **18** in the presence of urea followed by deoxychlorination with $POCl_3$. Mono-Boc protected diamines **22a–d**, **26a,b**, **29**, and **32**, prepared according to known

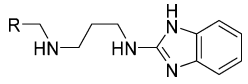
Scheme 7. 1 Cyclic Linker Analogue Synthesis[a]

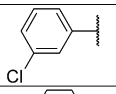
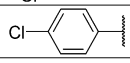
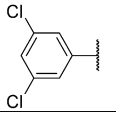


[a]Reagents and conditions: (a) **20**, toluene, μW, 150 °C, 6 h; (b) TFA, CH_2Cl_2, rt, overnight.

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225–236

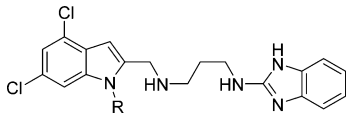
Table 2. Activity of Chlorophenyl Analogues of 1 against *T. brucei*

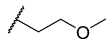
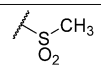
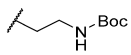
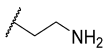


Compound	R	*T. brucei* EC_{50} (nM)[a]	MRC5-SV2 TC_{50} (μM)[b]	Selectivity (TC_{50}/EC_{50})
38a		150	22	147
38b		960	22	23
38c		16	22	1375

[a]SD < 0.20. [b]SD < 3.0.

Table 3. Activity of Indole *N*-Substituted Analogues of 1 against *T. brucei*



Compound	R	*T. brucei* EC_{50} (nM)[a]	MRC5-SV2 TC_{50} (μM)[b]	Selectivity (TC_{50}/EC_{50})
47a	H	9.0	8.8	978
47b	CH_3	5.0	9.8	1960
47c		8.0	7.4	925
47d		620	10	16
47e		340	1.4	4.1
47f		790	8.1	10
47g		200	8.4	42

[a]SD < 0.20. [b]SD < 3.0.

procedures[8−13] or commercially available, were coupled to **20** using microwave irradiation at 150 °C. Removal of the Boc group was effected either under acidic conditions to yield the bis-TFA salt or via microwave-promoted thermolysis to give the free base. Attempts to couple **20** with **21a** directly were successful in generating product, although the purification of **24a** was complicated due to its high water solubility. This was overcome by the introduction a Boc group, allowing purification via extraction followed by chromatography. The isolation of **24** was accomplished by simple concentration of the reaction medium.

2-(Aminophenyl)aminobenzimidazoles **36a−c** were prepared and purified without Boc protection under microwave heating (Scheme 8). Reaction times were reduced to 15 min by increasing the temperature from 150 to 200 °C.

Reductive amination of 2-indolecarboxaldehydes **7a−f**, **13**, and **17** with **24a** was carried out with $NaBH_3CN$ under either mildly basic conditions with the bis-TFA salt of **24a** or mildly acidic conditions with the free base (Scheme 9, product structures in Tables 1 and 3). The same reaction conditions were used to couple **7a** with amines **24**, **28**, **31**, **34**, and **36** (Table 4).

The synthesis of analogues **38**, substituting the dichloroindole scaffold with chlorophenyl rings, is shown in Scheme 10, via reductive amination from commercially available chlorobenzaldehydes **37**.

The side product **8** (Scheme 1) was used to produce **39** via ester aminolysis with the free base of **24a** and microwave irradiation (Scheme 11).

To synthesize analogues varying the benzimidazole of **1**, aldehyde **7a** was first reductively aminated with **22a** (Scheme 12). Exchange of the Boc protecting group for the acid-stable *N*-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl) (Dde) protecting group provided **41** in high yield. The secondary amine was Boc protected and the enamine removed upon heating in the presence of hydrazine monohydrate. The free primary amine of **43** was then reacted with phenylisocyanate to provide **44a**, 2-chlorobenzoxazole (**44b**), or 4-chloropyrimidine (**44c**). Deprotection of the Boc-protected secondary amines under acidic conditions gave the final analogues in modest yields.

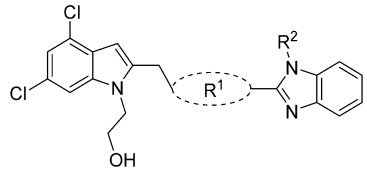
In Vitro Assessment against *T. brucei*. *Indole Halogens.* Removal of both chlorides on the indole ring resulted in a 78 times loss in potency (**46**, EC_{50} = 75 nM, Table 1). A significant reduction in MW (432 → 364) and clogP (3.60 → 2.39) was effected, however, resulting in the highest CNS multiparameter optimization (MPO) score[14] among these analogues. Central nervous system (CNS) exposure is essential to effectively treat stage 2 HAT when the parasites have invaded the CNS and crossed the blood−brain barrier (BBB). Noting that a CNS MPO score ≥4 is suggestive of CNS penetration, the score for **46** was calculated to be 4.3. This compound also retained a high lipophilic ligand efficiency (LLE, pEC_{50}-cLogP) of 4.71 (the desirable range for LLE is ≥4).[15] Replacement of the aminomethyl group of **1** with an amide (**39**) resulted in a 520-fold loss in activity. This correlates well with the conclusions from the linker SAR in that the basicity of the dialkylamine is essential for high potency.

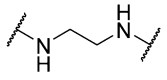
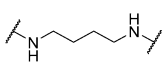
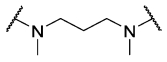
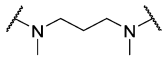
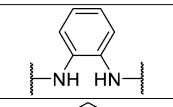
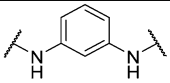
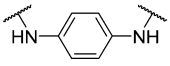
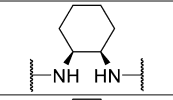
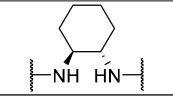
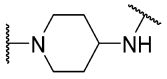
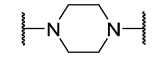
Table 2 shows the results of replacing the dichloroindole scaffold with simplified chlorophenyl rings. Compound **38c**, retaining the 3,5-dichloro motif, showed only a slight reduction in potency (EC_{50} = 16 nM). Removal of one of the chlorine atoms (**38a**) resulted in an 8-fold loss in activity (EC_{50} = 0.15 μM) compared to **38c**, and migration of the chloride to the *para*-position (**38b**) resulted in a further 6-fold loss in activity (EC_{50} = 0.96 μM). Our observation that the indole moiety in this class of compounds can be replaced with a simple dichlorophenyl group may provide a useful path forward in the event that toxicity due to indole metabolism is observed.

Indole N-Substituents. Modifications to the 2-hydroxyethyl group of **1** displayed mixed results (Table 3). Truncations (**47a**, **47b**) or a shift of the oxygen atom to an internal position (**47c**) resulted in a <10-fold loss in activity. Larger groups (**47d**, **47e**, **47f**) comparatively caused a far larger loss in activity (>300 times). Replacement of the alcohol with an amine (**47g**) led to around a 200-fold loss in activity. MPO scoring for these compounds was similar (≤Δ 0.3) to that for **1** except for the large increase resulting from the addition of the methanesulfonamide group (**47e**, MPO = 4.1) and a large decrease apparent in the Boc-protected **47f** (MPO = 1.5). Generally, the smaller indole *N*-substituents remained highly active, whereas those

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

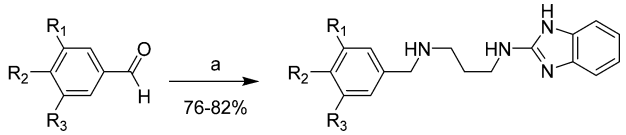
Table 4. Activity of Linker Analogues of 1 against *T. brucei*



Compound	R^1	R^2	*T. brucei* EC_{50} (nM)[a]	MRC5-SV2 TC_{50} (μM)[b]	Selectivity (TC_{50}/EC_{50})
48a		H	5.0	26[c]	5200
48b		H	5.0	11	2200
48c		H	980	22	22
48d		CH_3	2200	10	4.5
48e		H	420	25	60
48f		H	960	21	22
48g		H	5500	23	4.2
48h		H	680	14	21
48i		H	340	17	50
48j		H	400	12	30
48k		H	>5000	50	--

[a]SD < 0.55. [b]SD < 3.0 unless otherwise noted. [c]SD = 8.8.

Scheme 10. Synthesis of Chlorophenyl 1 Analogues[a]

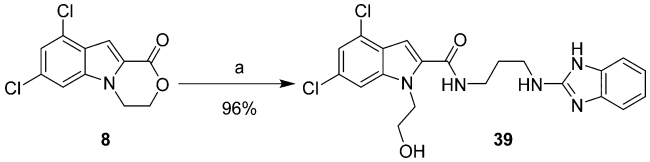


37a R^1=R^2=H, R^3=Cl
37b R^1=R^3=H, R^2=Cl
37c R^1=R^3=Cl, R^2=H

38a R^1=R^2=H, R^3=Cl
38b R^1=R^3=H, R^2=Cl
38c R^1=R^3=Cl, R^2=H

[a]Reagents and conditions: (a) **24a**, AcOH, $NaBH_3CN$, CH_3OH.

Scheme 11. Synthesis of 38 via Ester−Amide Exchange[a]



8 **39**

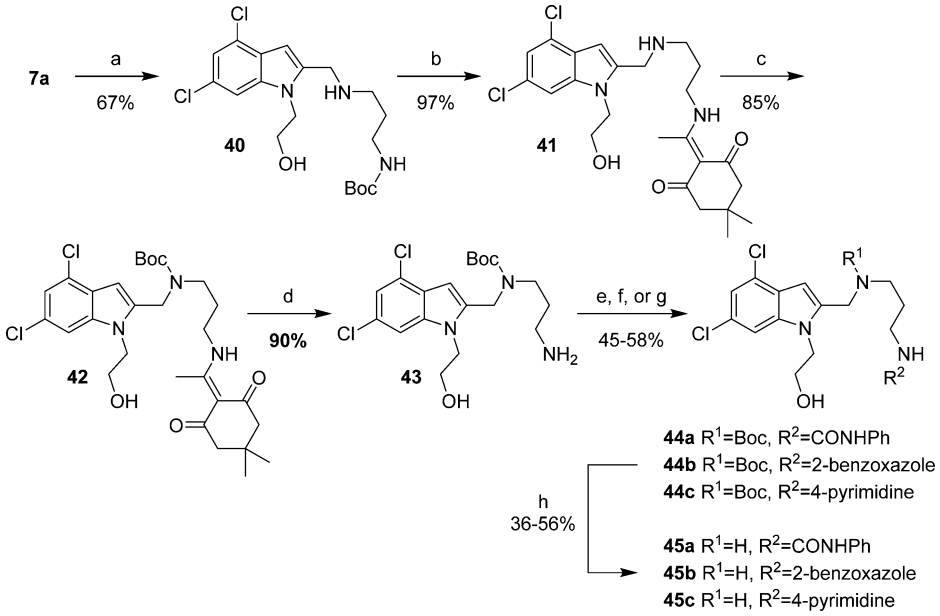
[a]Reagents and conditions: (a) **24a**, DMF, μW, 150 °C, 1 h.

that increased steric bulk resulted in a drop in potency by 2−3 orders of magnitude.

Linker. Contracting or extending the diamino linker by a methylene unit (**48a** and **48b**) did not lead to a loss in potency (EC_{50} = 5.0 nM each, Table 4) compared to **1**. Compound **48a** did, however, possess a slightly improved MPO score (3.8). Capping both amines of the linker with methyl groups was not tolerated (**48c**, EC_{50} = 0.98 μM). The trimethylated compound (**48d**) showed approximately an additional 3-fold loss in activity (EC_{50} = 2.2 μM).

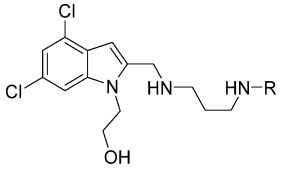
DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

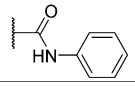
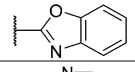
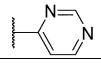
Scheme 12. Heterocyclic Replacements on the Eastern Half of 1[a]



[a]Reagents and conditions: (a) **22a**, AcOH, $NaBH_3CN$, CH_3OH, rt, 12 h; (b) 3 M aq HCl, THF, rt, 1 h, 80 °C, 1 h; then DDE-OH, $(i\text{-}Pr)_2NEt$, CH_2Cl_2, CH_3OH, rt, 12 h; (c) Boc_2O, K_2CO_3, CH_3CN, H_2O, rt, 12 h; (d) $H_2NNH_2{\cdot}H_2O$, CH_3CN, 45 °C, 3 h; (e) PhCNO, CH_3CN, rt, 12 h; (f) 2-chlorobenzoxazole, $(i\text{-}Pr)_2NEt$, CH_3CN, 80 °C, 12 h; (g) 4-chloropyrimidine, $(i\text{-}Pr)_2NEt$, CH_3CN, μW, 150 °C, 2 h; (h) 4 M HCl in 1,4-dioxane, rt, 6 h.

Table 5. Activity of Chlorophenyl Analogues of 1 against *T. brucei*



Compound	R	*T. brucei* EC_{50} (nM)[a]	MRC5-SV2 TC_{50} (μM)[b]	Selectivity (TC_{50}/EC_{50})
45a		2.0	14	7000
45b		790	26	33
45c		270	>50	--

[a]SD < 0.20. [b]SD < 3.0.

The tolerance for a chain of two to four carbons between amines suggests adequate space within the biological target(s) of action for conformational variations of the linker. The *p*-, *m*-, and *o*- phenylenediamine linkers, designed to restrict the conformation of the linker, all showed a dramatic loss in activity. The ortho-linked **48e** was the most active (EC_{50} = 0.42 μM), followed closely by the meta-linked **48f** (EC_{50} = 0.96 μM). The para-linked compound, **48g**, was an additional order of magnitude less active (EC_{50} = 5.5 μM). These data suggested either that (1) a bent orientation of the linker is closest to the active conformation of **1** or (2) the basicity of these nitrogens is important.

Compounds **48h**−**k** further probed the tolerance for reduced linker flexibility and diamine orientation. Three of the four saturated cyclic linked analogues, **48h**−**j**, remained sub-micromolar in activity. Compounds **48h** and **48i** were within 2-fold of one another, showing no preference for *cis* or *trans* substitution of the cyclohexyl ring, and were essentially equipotent to the aromatic analogue **48e**. Compound **48k**, bearing a piperazinyl linker, was >5 μM, compared to the 4-aminopiperadinyl linked **48j** (EC_{50} = 0.40 μM). This difference in activity is likely not a result of linker reduction as **48a** retained essentially all activity compared to **1**. Rather, taken together with the poor activity of **48c** and **48d**, the reduced activity of **48k** illustrates the importance of (1) the hydrogen bonding donor motif and (2) a high degree of flexibility in the linker for high potency.

Benzimidazole Replacements. Any replacements for the benzimidazole ring that lacked a hydrogen bond donor were not tolerated (Table 5, **45b** and **45c**). Specifically, changing to an isosteric benzoxazole (**45b**) led to around an 800-fold loss in activity compared to **1**. Compound **45c**, also devoid of a hydrogen bond donating 4-pyrimidine ring, possessed activity approximating that of **45b** (0.79 vs 0.27 μM). Reintroduction of a hydrogen bond donor with a urea N−H (**45a**), which could recapitulate that of the benzimidazole, regained the high potency observed for **1**.

All analogues of **1** were screened against MRC5-SV2 human lung cells to assess toxicity. Most compounds possessed a selectivity index in the 10−50 times range. Compounds in the single-digit nanomolar range against *T. brucei* displayed a higher selectivity index in the range of 900−7000 times.

Physicochemical and ADME Properties. The calculated physicochemical properties of compounds **1**, **38a**−**c**, **39**, **45a**−**c**, **46**, **47a**−**g**, and **48a**−**k** are presented in Table 6, color coded in terms of desirability. Compound **47e** is noteworthy among analogues replacing the 2-hydroxyethyl group for its high MPO score above 4.0, due to its reduced clogP and pK_a (calculated using JChem for Excel, Chemaxon, Inc.) Analogues **48e**−**i** bearing phenylenediamine and 1,2-diaminocyclohexyl linkers

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

Table 6. Physicochemical Properties of 1 and Its Analogues

Cmpd	MW (g/mol)[a]	clogP[a,c]	clogD[a,d]	H-bond donors[a]	pKa[b,c]	TPSA (Å²)[b,c]	MPO Score[e]	LE	LLE
1	432.4	3.60	1.59	4	9.29	82.48	3.5	0.31	5.57
38a	314.8	3.44	1.57	3	9.15	57.32	4.4	0.31	3.38
38b	314.8	3.44	1.53	3	9.19	57.32	4.4	0.27	2.49
38c	349.3	4.04	2.33	3	8.98	57.32	4.0	0.33	3.65
39	446.3	3.11	2.98	4	6.98	94.97	3.7	0.21	3.19
45a	435.4	3.28	1.36	4	9.34	82.90	3.7	0.30	5.42
45b	433.3	3.66	1.78	3	9.29	79.83	3.7	0.21	2.45
45c	394.3	2.30	0.42	3	9.29	79.58	4.3	0.25	4.27
46	363.5	2.39	0.36	4	9.31	82.48	4.3	0.26	4.71
47a	388.3	4.06	2.06	4	9.28	73.11	3.6	0.31	4.04
47b	402.3	4.29	2.28	3	9.29	62.25	3.4	0.31	4.11
47c	432.4	4.35	2.36	3	9.27	71.48	3.2	0.28	3.75
47d	446.4	4.24	2.23	3	9.29	71.48	3.2	0.21	1.96
47e	466.4	2.88	2.08	3	8.01	96.39	4.1	0.22	3.59
47f	531.5	4.98	2.97	4	9.29	100.58	1.5	0.17	1.12
47g	431.4	3.49	-0.73	4	9.83	89.89	3.3	0.23	3.21
48a	418.3	3.54	1.88	4	8.97	82.48	3.8	0.30	4.76
48b	446.4	4.11	1.82	4	9.60	82.48	3.0	0.28	4.39
48c	460.4	4.62	3.43	2	8.54	61.52	3.0	0.20	1.48
48d	474.4	4.84	3.66	1	8.54	50.66	3.0	0.18	0.81
48e	466.4	5.34	5.07	4	7.38	79.15	2.2	0.20	1.06
48f	466.4	5.34	5.06	4	7.39	79.15	2.2	0.19	0.76
48g	466.4	5.34	5.04	4	7.42	79.15	2.2	0.16	-0.14
48h	472.4	4.95	2.94	4	9.35	82.48	2.1	0.20	1.51
48i	472.4	4.95	2.94	4	9.35	82.48	2.1	0.19	1.22
48j	458.4	3.93	3.61	3	7.38	70.31	3.2	0.21	2.47
48k	444.4	4.38	4.30	2	6.73	60.32	3.2	0.18	0.92

[a]Values that are within the MPO more desired ranges are green and outside the less desired range are red. [b]Calculated using ChemAxon JChem for Excel. [c]Calculated at pH 7.4 using ChemAxon JChem for Excel. [d]Calculated using MPO calculator previously reported.[14]

Table 7. ADME Measurements of 1 and Analogues

compound	aqueous solubility (uM)	human PPB (%)	HLM Cl_{int} (μL/min/mg)	HLM $t_{1/2}$ (min)	rat hepatocyte Cl_{int} (μL/min/10^6)	rat hepatocyte $t_{1/2}$ (min)	logD 7.4
1	6.33	>99.9	24	*[a]	24.9	*	4.0
46	298	98.61	22.6	*	81.4	*	2.0
47a	4.02	99.8	15.5	*	17.2	*	4.3
47b	12.2	99.8	12.9	*	29.3	*	4.5
47c	<1	99.9	34.2	*	33.6	*	4.0
47d	3.77	>99.9	43.8	*	49.7	*	4.1
47e	19	*	80.9	8.6	98.8	7	*
47f	32	*	142	4.9	33.3	20.8	*
47g	*	*	*	*	7.96	87.1	*
48a	21	99.8	36	*	30.7	*	4.3
48b	4.26	99.9	22.5	*	17	*	3.7
48c	11.7	>99.9	135	*	61.7	*	4.0
48d	*	*	*	*	>300	<2.3	*
48e	*	>99.9	51.6	*	100	*	4.9
48f	<1	>99.9	33.9	*	22.5	*	4.3
48g	1.39	>99.9	27.3	*	18.7	*	4.7
48h	11	>100	64.2	*	115	*	4.0
48j	0.3	100	*	*	38.8	*	4.4
48k	<1	99.0	21.1	*	12.8	*	4.9

[a]An asterisk (*) indicates the value was not obtained.

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

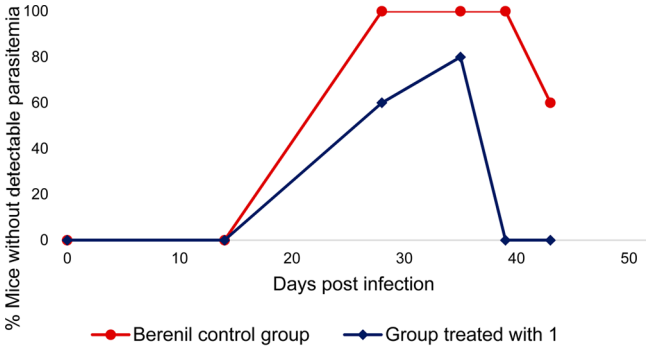


Figure 2. Mice without detectable parasitemia in a mouse model of CNS infection with *T. b. brucei*. Limit of detection for parasitemia = 125,000 parasite/mL blood.

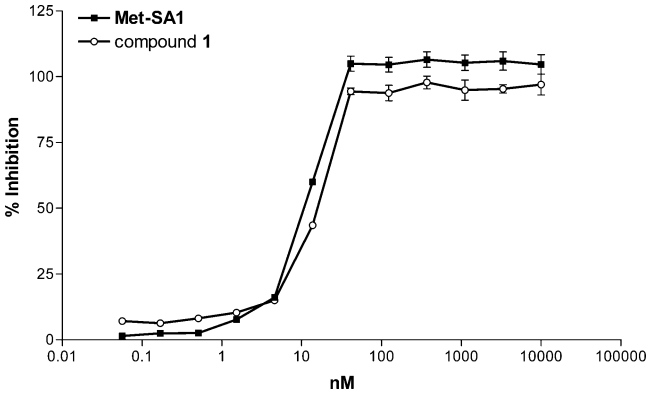


Figure 4. *T. brucei* MetRS inhibition assay. IC_{50}values of both the positive control, Met-SA1, and compound **1** are <25 nM (the lower detection limit of the assay as described under Biochemical Assessment in the Target ID section). Error bars represent standard error of the mean of three replicates. Pentamidine, a negative control, had no *T. brucei* MetRS inhibition up to 10 μM (results not shown).

possess poor MPO scores as a result of high molecular weights, clogP, and clogD values.

An excellent MPO score (4.3) is calculated for **46**, which resulted from reduced molecular weight compared to **1** (363.5, Δ68.9), clogP (2.39, Δ1.21), and clogD (0.36, Δ1.23); this results simply from the removal of the chlorine atoms from the indole. In addition to the improved MPO scores of many compounds in this series, several compounds retained LE values ≥0.30 and LLE values ≥4.0, indicating a good balance of size and lipophilicity.

In vitro absorption, distribution, metabolism, and excretion (ADME) properties were collected for several compounds and are provided in Table 7. Few compounds improved the aqueous solubility to an appreciable extent, with **46** as the only significant exception at 298 μM. Compound **46** was also the only analogue to show reduced human plasma protein binding (PPB) to <99% or to reduce the $logD_{7.4}$ below 3.0 among those tested. Of the 16 compounds assessed for human liver microsome clearance (HLM Cl_{int}), only 4 analogues possessed lower clearance than **1**. Compounds **47a** and **47b** displayed the lowest human microsomal clearance, implying the 2-hydroxyethyl group may indeed be a metabolic handle as we expected. Others with high microsomal clearance are **47d**− **f**, **48c**, **48e**, and **48h**; the Boc protected analogue **47f** and the dimethylamino analogue **48c** both showed rapid metabolism (142 and 135 μL/min/mg respectively). Similarly, few compounds possessed good rat hepatocyte clearance, with only **47g** (Cl_{int} = 7.96 μL/min/10^6) and **48k** (Cl_{int} = 12.8 μL/min/10^6) showing any significant improvement over **1**.

Pharmacokinetics sStudy. Although we previously reported plasma pharmacokinetics, we did not determine CNS drug levels. Thus, female BALB/c mice were injected with a single, 10 mg/kg intraperitoneal dose of **1**, and plasma and brain levels were measured, showing a brain-to-plasma ratio of 0.39, with drug levels of 31.6 ng/g of brain tissue measured at the 24 h time point. (The PK parameters are summarized in the Supporting Information, Tables S1−S3 and Figure S1.) With this evidence in hand, we proceeded to assess this compound in a mouse model of stage 2 HAT.

Murine CNS Infection Model. We advanced **1** to a mouse model of a CNS infection with *T. b. brucei* (GVR 35 strain). On day 0, two groups of five mice each were infected with *T. b. brucei* and checked for parasitemia on day 14. One group, taken as the control, was treated with a single dose of 40 mg/kg Berenil on day 21. From days 22 to 25 the control group was treated with the vehicle DMSO in PBS. The other group of mice was treated with **1** at a dose of 20 mg/kg/day on days 21−25. Following a 2-day hiatus, treatment with **1** was resumed on days 28−32. Parasitemia was checked from both groups via examination of tail blood on days 14, 28, 35, 39, and 43. The results are reported in Figure 2 and Table S-4.

On day 28, 2 days after the first regimen of treatment was completed, three of five mice treated with **1** showed no detectable parasitemia. Following the second round of treatment, this number increased to four of five mice showing no detectable parasitemia on day 35. When checked on day 39, however, all mice treated with **1** showed a full relapse of parasites in the blood. The control group comparatively showed no detectable levels of parasitemia when checked on days 28, 35, and 39 in all five mice, and only two mice showed a relapse

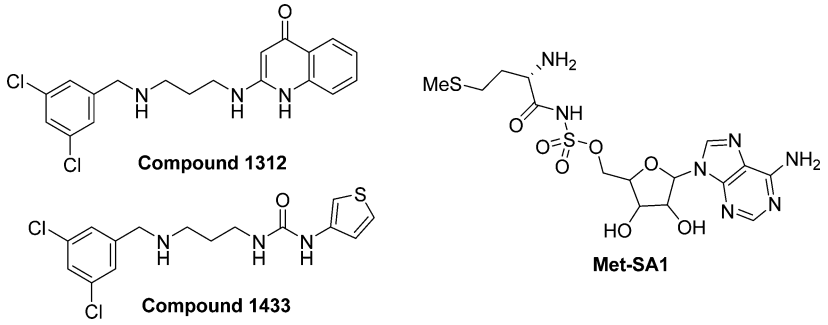


Figure 3. Structures of previously published *T. brucei* MetRS inhibitors (**1312**,[16] **1433**,[17] and **Met-SA1**[19]).

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

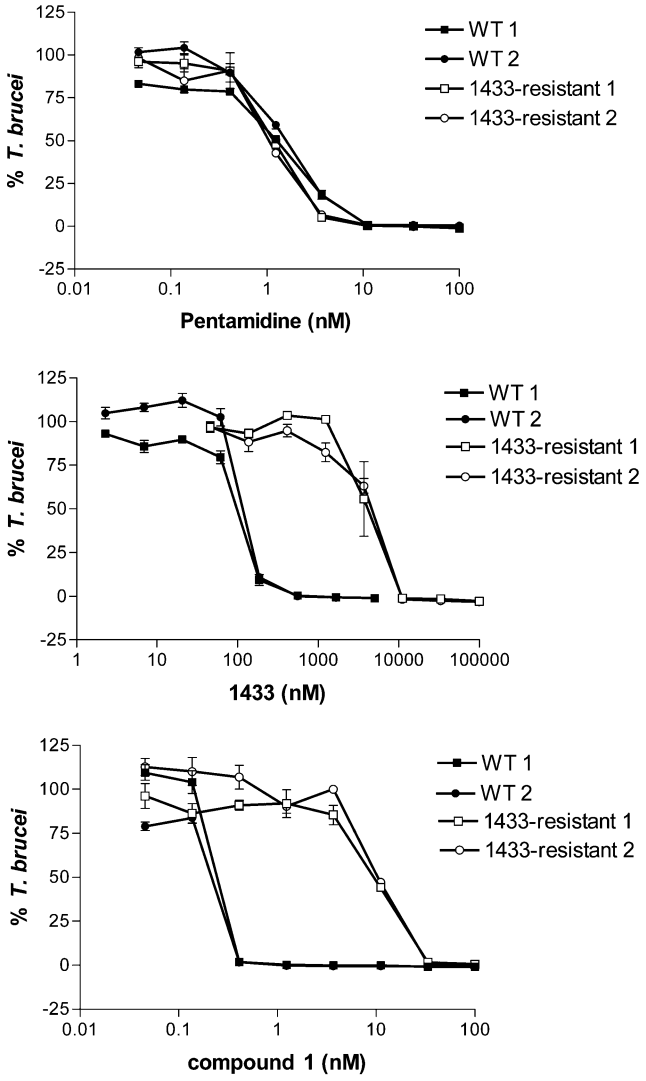


Figure 5. Wild-type and **1433**-resistant *T. brucei* growth inhibition assays. Results from two independent assays, as indicated by "1" and "2". Error bars represent standard error of the mean of three replicates.

Table 8. Wild-Type and 1433-Resistant *T. brucei* Growth Inhibition Assays[a]

	strain		
compound	WT EC_{50} (nM)	**1433**-resistant EC_{50} (nM)	**1433**-resistant EC_{50}/WT EC_{50}
pentamidine	1.5 ± 0.2	1.1 ± 0.1	0.73
1433	98 ± 1	4700 ± 230	48
1	0.13 ± 0.01	14 ± 1	108

[a] EC_{50} values represent averages and standard error of the mean of two independent assays.

on day 43. Because recrudescence was observed even in animals in which the bloodstream infection was cleared by **1**, we interpret this result to suggest that compound **1** does not have sufficient antiparasitic activity in the brain to cure stage 2 HAT. This may be attributed in part to the high PPB (>99.9%) and rapid clearance (rat hepatocyte Cl_{int} = 24.9 μL/min/10^6 cells) of **1**.

Target Identification. Although the potent activity of **1** was uncovered from a set of compounds selected on the basis of their likely kinase inhibition activity, we noted that the structure of **1** was quite similar to a class of previously reported *T. brucei* methionyl-tRNA synthetase (MetRS) inhibitors (such as compound **1312**, Figure 3).[16] *T. brucei* has a single MetRS and is necessary for cell growth as previously reported using RNA interference.[16] We wished to determine whether **1** inhibited *T. brucei* MetRS.

Biochemical Assessment. A functional assay for the *T. brucei* MetRS enzyme based on ATP depletion was previously described.[20] Compound **1** was determined to be a highly potent inhibitor, with an IC_{50} below the lower detection limit of the assay (<25 nM) (Figure 4). The positive control compound, Met-SA1[19] Figure 3, is also a potent inhibitor with an IC_{50} of <25 nM.

Potency Shift in Resistant Parasites. To further test if compound **1** is targeting the *T. brucei* MetRS, a *T. brucei* growth inhibition cell assay was conducted with a strain that is resistant to MetRS inhibitors. The resistant strain was generated by serial in vitro passage of wild-type *T. brucei* with the MetRS inhibitor **1433** (Figure 3) and was shown to have ∼35-fold up-regulated expression of MetRS mRNA.[18] The EC_{50} values of **1** and **1433** are ∼108- and ∼48-fold less potent against the **1433**-resistant strain compared to the wild-type (Figure 5 and Table 8), consistent with the hypothesis that **1** acts by inhibiting the MetRS enzyme. Pentamidine does not inhibit *T. brucei* MetRS (data not shown) and does not show a difference in EC_{50} values between the **1433**-resistant and wild-type strains.

Structural Biology. Even though crystals of TbruMetRS·Met would disintegrate after ∼20 s in contact with a solution of **1**, it was possible to replace the bound methionine with **1** in the binding site using short soaking times.[21] The structure in complex with **1** was determined at 2.6 Å resolution and deposited in the Protein Data Bank with PDB identifier 5TQU. Data collection and refinement statistics are shown in Table S5 (Supporting Information), and the electron density map for the compound is shown in Figure 6A. Compound **1** binds to TbruMetRS in a manner similar to that previously reported for compound-bound structures.[21,22] Characteristically, two pockets are involved in binding of the inhibitor (Figure 6B): the enlarged methionine pocket (EMP), formed mainly by hydrophobic residues including the ones engaged in methionine binding, and the "auxiliary pocket" (AP), not present in the methionine-bound state.[21] The EMP, formed mainly as a result of the movement of TbruMetRS residues Val473, Trp474, and Phe522, is occupied by the 4,6-dichloroindolyl ring moiety, which also interacts with residues Ile248 and Tyr481. The AP is occupied by the benzimidazole moiety, which interacts with His289, Gly290, and Val473, while establishing a hydrogen bond with Asp287 through one of the nitrogens of the benzimidazole ring with a distance of 3.0 Å. The flexible linker of **1** forms hydrophobic interactions with the aromatic group of residue Tyr250 and a hydrogen bond through the exocyclic amine of the benzimidazole moiety with the catalytic Asp287 with a distance of 2.8 Å (Figure 6C). In a novel interaction pattern for TbruMetRS inhibitors, the OH group in the hydroxyethyl substituent of the 4,6-dichloroindolyl ring moiety makes a hydrogen bond with residue Asp287 with a distance of 2.6 Å. As a result, Asp287 plays a central role in establishing polar contacts with the three different regions of **1**: the linker, the benzimidazole moiety, and the OH group in the hydroxyethyl moiety. Furthermore, at least one water molecule is part of what seems to be a crucial hydrogen bonding network involving Asp287 and **1** (Figure 6C).

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

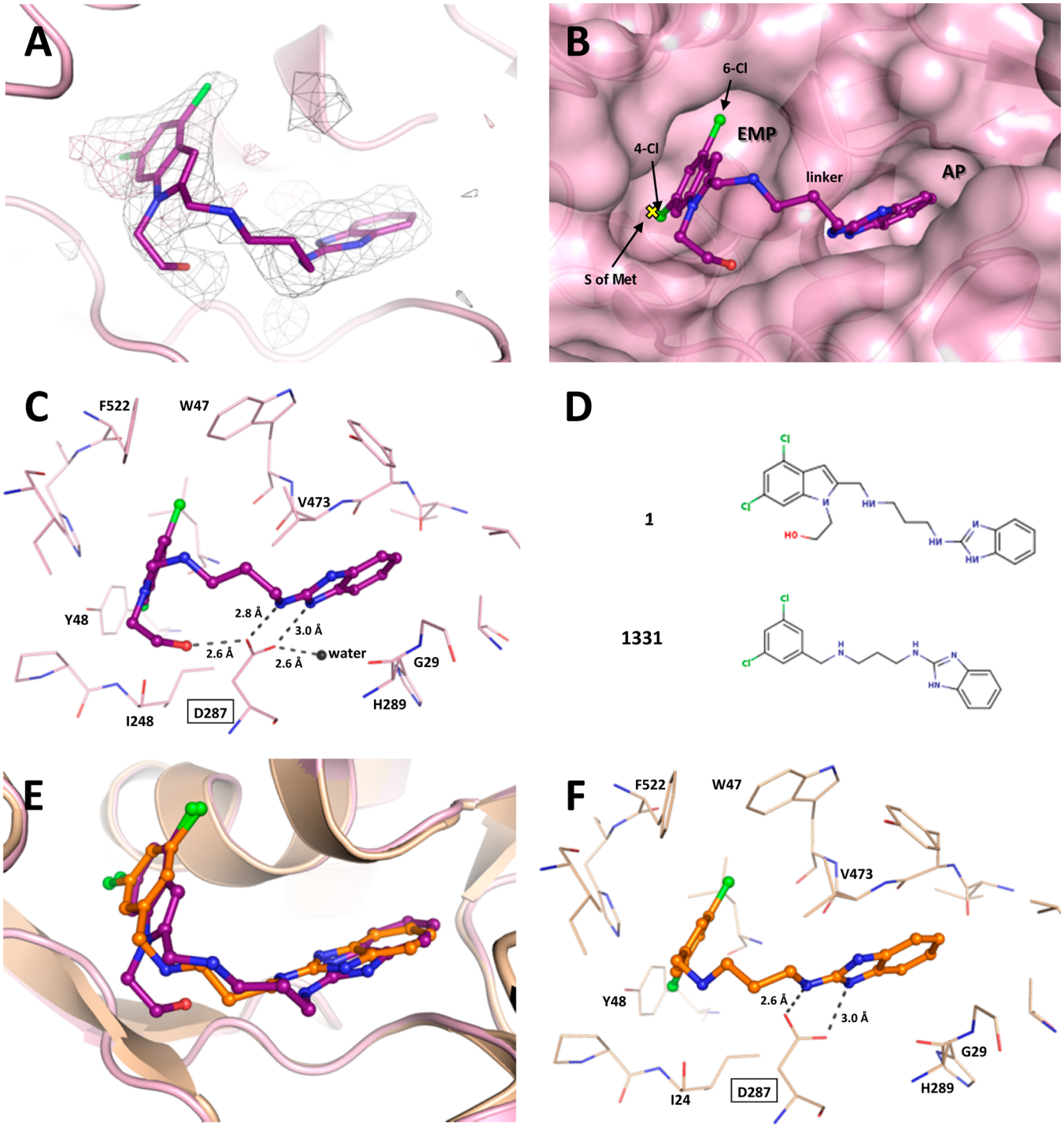


Figure 6. Binding of **1** to TbruMetRS. (A) TbruMetRS·**1** structure (PDB: 5TQU) with the difference electron density map calculated by omitting the inhibitor, contoured at the 3σ level (positive density in gray, negative density in red). (B) General features of the **1** binding mode. The position of the methionine sulfur atom in the TbruMetRS·methionine structure (PDB: 4EG1)[21] is shown with a yellow cross. The protein surface and the two pockets, EMP and AP, where the inhibitor is bound are shown. (C) Hydrogen bond network in the TbruMetRS·**1** structure. The label for residue Asp287 is highlighted with a rectangle. The hydroxyethyl substituent of the 4,6-dichloroindolyl-ring moiety establishes an extra hydrogen bond with residue Asp287. (D) Chemical structures of compound **1** and the previously reported TbruMetRS inhibitor **1331**. (E) Superposition of compounds **1** and **1331** (PDB: 4EG7)[21] bound to TbruMetRS. (F) Hydrogen bond interactions between inhibitor and enzyme in the TbruMetRS·**1331** structure.

When the new structure is compared to previously reported TbruMetRS complexes with inhibitors bound, the overall binding mode of **1** is quite similar. Taking the TbruMetRS·**1331** structure (Figure 6D) as a representative of previously determined inhibitor-bound structures, it appears that the two ring systems of each inhibitor occupy similar positions. For instance, the two chlorine atoms are only 0.8 and 0.3 Å (Cl-4 and Cl-6, respectively) apart in the two structures (Figure 6E) and the linkers follow a similar path. The key difference is the extra interaction of the hydroxyl group of **1** with Asp287

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

(compare Figure 6C,F), representing one additional hydrogen bond between the inhibitor and the target enzyme.

CONCLUSIONS

In summary, we report the structure−activity relationships for the various regions of **1** and describe its activity in a mouse model of CNS infection of *T. brucei*. In addition, the compound's mechanism of action via *T. brucei* MetRS was conclusively elucidated by a combination of biochemical, cellular, and crystallographic techniques. Despite the excellent in vitro antiparasitic potency of **1**, the lack of CNS activity in the late-stage model of HAT necessitates further optimization of CNS permeability and pharmacokinetics; this work is ongoing.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsinfecdis.6b00202.

Experimental details for the pharmacokinetics experiments, chemical syntheses, X-ray crystallography, and in vitro and in vivo biological methods; annotation of compounds with NEU registry numbers and SMILES strings (PDF)

AUTHOR INFORMATION

Corresponding Author

*(M.P.P.) E-mail: m.pollastri@neu.edu.

ORCID

Michael P. Pollastri: 0000-0001-9943-7197

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under Awards R56AI099476 (to M.P.P.), R01AI114685 (to M.P.P. and M.N.), R01AI084004 (to W.G.J.H.), and R01AI097177 (to F.S.B. and R.M.R.). We acknowledge the support of a Fulbright Fellowship to X.B.-A and funding from Northeastern University. We thank Erkang Fan (University of Washington) for providing compounds **1433** and Met-SA1. We thank Christophe L. M. Verlinde for his contributions to the structure validation process. We thank Stewart Turley and Robert Steinfeldt for providing support for the X-ray data collection and computing environment at the Biomolecular Structure Center of the University of Washington. Crystallography performed in support of the work benefitted from remote access to resources at the Stanford Synchrotron Radiation Lightsource supported by the U.S. Department of Energy Office of Basic Energy Sciences under Contract DE-AC02-76SF00515 and by the National Institutes of Health (P41GM103393).

REFERENCES

(1) Jacobs, R. T., Nare, B., Wring, S. A., Orr, M. D., Chen, D., Sligar, J. M., Jenks, M. X., Noe, R. A., Bowling, T. S., Mercer, L. T., Rewerts, C., Gaukel, E., Owens, J., Parham, R., Randolph, R., Beaudet, B., Bacchi, C. J., Yarlett, N., Plattner, J. J., Freund, Y., Ding, C., Akama, T., Zhang, Y. K., Brun, R., Kaiser, M., Scandale, I., and Don, R. (2011) SCYX-7158, an Orally-Active Benzoxaborole for the Treatment of Stage 2 Human African Trypanosomiasis. *PLoS Neglected Trop. Dis. 5*, e1151.

(2) Torreele, E., Bourdin Trunz, B., Tweats, D., Kaiser, M., Brun, R., Mazué, G., Bray, M. A., and Pécoul, B. (2010) Fexinidazole − A New Oral Nitroimidazole Drug Candidate Entering Clinical Development for the Treatment of Sleeping Sickness. *PLoS Neglected Trop. Dis. 4*, e923.

(3) Hay, M., Thomas, D. W., Craighead, J. L., Economides, C., and Rosenthal, J. (2014) Clinical development success rates for investigational drugs. *Nat. Biotechnol. 32*, 40−51.

(4) Diaz, R., Luengo-Arratta, S. A., Seixas, J. D., Amata, E., Devine, W., Cordon-Obras, C., Rojas-Barros, D. I., Jimenez, E., Ortega, F., Crouch, S., Colmenarejo, G., Fiandor, J. M., Martin, J. J., Berlanga, M., Gonzalez, S., Manzano, P., Navarro, M., and Pollastri, M. P. (2014) Identification and characterization of hundreds of potent and selective inhibitors of *Trypanosoma brucei* growth from a kinase-targeted library screening campaign. *PLoS Neglected Trop. Dis. 8*, e3253.

(5) Mansueto, M., Frey, W., and Laschat, S. (2013) Ionic Liquid Crystals Derived from Amino Acids. *Chem.−Eur. J. 19*, 16058−16065.

(6) Loison, S., Cottet, M., Orcel, H., Adihou, H., Rahmeh, R., Lamarque, L., Trinquet, E., Kellenberger, E., Hibert, M., Durroux, T., Mouillac, B., and Bonnet, D. (2012) Selective Fluorescent Nonpeptidic Antagonists For Vasopressin V2 GPCR: Application To Ligand Screening and Oligomerization Assays. *J. Med. Chem. 55*, 8588−8602.

(7) Attenburrow, J., Cameron, A. F. B., Chapman, J. H., Evans, R. M., Hems, B. A., Jansen, A. B. A., and Walker, T. (1952) 194. A synthesis of vitamin A from cyclohexanone. *J. Chem. Soc.*, 1094−1111.

(8) Montero, A., Goya, P., Jagerovic, N., Callado, L. F., Meana, J. J., Girón, R. o., Goicoechea, C., and Martín, M. I. (2002) Guanidinium and aminoimidazolinium derivatives of *N*-(4-piperidyl)propanamides as potential ligands for μ opioid and I2-imidazoline receptors: synthesis and pharmacological screening. *Bioorg. Med. Chem. 10*, 1009−1018.

(9) New, O. M., and Dolphin, D. (2009) Design and Synthesis of Novel Phenothiazinium Photosensitiser Derivatives. *Eur. J. Org. Chem. 2009*, 2675−2686.

(10) Schmuck, C., and Dudaczek, J. (2007) Ion Pairing Between the Chain Ends Induces Folding of a Flexible Zwitterion in Methanol. *Eur. J. Org. Chem. 2007*, 3326−3330.

(11) Marugan, J. J., Zheng, W., Motabar, O., Southall, N., Goldin, E., Westbroek, W., Stubblefield, B. K., Sidransky, E., Aungst, R. A., Lea, W. A., Simeonov, A., Leister, W., and Austin, C. P. (2011) Evaluation of Quinazoline Analogues as Glucocerebrosidase Inhibitors with Chaperone Activity. *J. Med. Chem. 54*, 1033−1058.

(12) Cortez, N. A., Aguirre, G., Parra-Hake, M., and Somanathan, R. (2009) New heterogenized C2-symmetric bis(sulfonamide)-cyclohexane-1,2-diamine-RhIIICp* complexes and their application in the asymmetric transfer hydrogenation (ATH) of ketones in water. *Tetrahedron Lett. 50*, 2228−2231.

(13) Beena, Joshi, S., Kumar, N., Kidwai, S., Singh, R., and Rawat, D. S. (2012) Synthesis and Antitubercular Activity Evaluation of Novel Unsymmetrical Cyclohexane-1,2-diamine Derivatives. *Arch. Pharm. 345*, 896−901.

(14) Wager, T. T., Hou, X., Verhoest, P. R., and Villalobos, A. (2010) Moving beyond rules: The development of a central nervous system multiparameter optimization (CNS MPO) approach to enable alignment of druglike properties. *ACS Chem. Neurosci. 1*, 435−449.

(15) Shultz, M. D., Cheung, A. K., Kirby, C. A., Firestone, B., Fan, J., Chen, C. H., Chen, Z., Chin, D. N., Dipietro, L., Fazal, A., Feng, Y., Fortin, P. D., Gould, T., Lagu, B., Lei, H., Lenoir, F., Majumdar, D., Ochala, E., Palermo, M. G., Pham, L., Pu, M., Smith, T., Stams, T., Tomlinson, R. C., Toure, B. B., Visser, M., Wang, R. M., Waters, N. J., and Shao, W. (2013) Identification of NVP-TNKS656: The Use of Structure-Efficiency Relationships To Generate a Highly Potent, Selective, and Orally Active Tankyrase Inhibitor. *J. Med. Chem. 56*, 6495−6511.

(16) Shibata, S., Gillespie, J. R., Kelley, A. M., Napuli, A. J., Zhang, Z., Kovzun, K. V., Pefley, R. M., Lam, J., Zucker, F. H., Van Voorhis, W.

C., Merritt, E. A., Hol, W. G., Verlinde, C. L., Fan, E., and Buckner, F. S. (2011) Selective inhibitors of methionyl-tRNA synthetase have potent activity against *Trypanosoma brucei* Infection in Mice. *Antimicrob. Agents Chemother. 55*, 1982−1989.

(17) Shibata, S., Gillespie, J. R., Ranade, R. M., Koh, C. Y., Kim, J. E., Laydbak, J. U., Zucker, F. H., Hol, W. G. J., Verlinde, C. L. M. J., Buckner, F. S., and Fan, E. (2012) Urea-Based Inhibitors of *Trypanosoma brucei* Methionyl-tRNA Synthetase: Selectivity and in Vivo Characterization. *J. Med. Chem. 55*, 6342−6351.

(18) Ranade, R. M., Gillespie, J. R., Shibata, S., Verlinde, C. L., Fan, E., Hol, W. G., and Buckner, F. S. (2013) Induced resistance to methionyl-tRNA synthetase inhibitors in Trypanosoma brucei is due to overexpression of the target. *Antimicrob. Agents Chemother. 57*, 3021.

(19) Ranade, R. M., Zhang, Z., Gillespie, J. R., Shibata, S., Verlinde, C. L. M. J., Hol, W. G. J., Fan, E., and Buckner, F. S. (2015) Inhibitors of Methionyl-tRNA Synthetase Have Potent Activity against *Giardia intestinalis* Trophozoites. *Antimicrob. Agents Chemother. 59*, 7128−7131.

(20) Pedro-Rosa, L., Buckner, F. S., Ranade, R. M., Eberhart, C., Madoux, F., Gillespie, J. R., Koh, C. Y., Brown, S., Lohse, J., Verlinde, C. L., Fan, E., Bannister, T., Scampavia, L., Hol, W. G., Spicer, T., and Hodder, P. (2015) Identification of potent inhibitors of the *Trypanosoma brucei* methionyl-tRNA synthetase via high-throughput orthogonal screening. *J. Biomol. Screening 20*, 122−130.

(21) Koh, C. Y., Kim, J. E., Shibata, S., Ranade, R. M., Yu, M., Liu, J., Gillespie, J. R., Buckner, F. S., Verlinde, C. L. M. J., Fan, E., and Hol, W. G. J. (2012) Distinct States of Methionyl-tRNA Synthetase Indicate Inhibitor Binding by Conformational Selection. *Structure 20*, 1681−1691.

(22) Koh, C. Y., Kim, J. E., Wetzel, A. B., de van der Schueren, W. J., Shibata, S., Ranade, R. M., Liu, J., Zhang, Z., Gillespie, J. R., Buckner, F. S., Verlinde, C. L. M. J., Fan, E., and Hol, W. G. J. (2014) Structures of *Trypanosoma brucei* Methionyl-tRNA Synthetase with Urea-Based Inhibitors Provide Guidance for Drug Design against Sleeping Sickness. *PLoS Neglected Trop. Dis. 8*, e2775.

DOI: 10.1021/acsinfecdis.6b00202
ACS Infect. Dis. 2017, 3, 225−236

European Journal of Medicinal Chemistry 124 (2016) 1081–1092

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

journal homepage: http://www.elsevier.com/locate/ejmech

Research paper

Structure-guided design of novel *Trypanosoma brucei* Methionyl-tRNA synthetase inhibitors

Wenlin Huang [a], Zhongsheng Zhang [a], Ximena Barros-Álvarez [a, b], Cho Yeow Koh [a, 1], Ranae M. Ranade [c], J. Robert Gillespie [c], Sharon A. Creason [c], Sayaka Shibata [a], Christophe L.M.J. Verlinde [a], Wim G.J. Hol [a], Frederick S. Buckner [c, *], Erkang Fan [a, **]

[a] Department of Biochemistry, University of Washington, Seattle, WA 98195, United States
[b] Laboratorio de Enzimología de Parásitos, Facultad de Ciencias, Universidad de los Andes, Mérida, Venezuela
[c] Department of Medicine, Division of Allergy and Infectious Diseases, and the Center for Emerging and Re-emerging Infectious Diseases (CERID), University of Washington, Seattle, WA 98109, United States

ARTICLE INFO

Article history:
Received 30 July 2016
Received in revised form
29 September 2016
Accepted 13 October 2016
Available online 14 October 2016

Keywords:
Human African trypanosomiasis
Methionyl-tRNA synthetase
Structure-guided design

ABSTRACT

A screening hit **1** against *Trypanosoma brucei* methionyl-tRNA synthetase was optimized using a structure-guided approach. The optimization led to the identification of two novel series of potent inhibitors, the cyclic linker and linear linker series. Compounds of both series were potent in a *T. brucei* growth inhibition assay while showing low toxicity to mammalian cells. The best compound of each series, **16** and **31**, exhibited EC_{50}s of 39 and 22 nM, respectively. Compounds **16** and **31** also exhibited promising PK properties after oral dosing in mice. Moreover, compound **31** had moderately good brain permeability, with a brain/plasma ratio of 0.27 at 60 min after IP injection. This study provides new lead compounds for arriving at new treatments of human African trypanosomiasis (HAT).

1. Introduction

Human African trypanosomiasis (HAT) is a parasitic disease caused by the protozoan *Trypanosoma brucei* [1,2]. The disease is endemic in regions of sub-Saharan Africa, causing infection risk to 70 million people [3,4]. Without treatment, the disease is invariably fatal. Current treatment for HAT includes suramin, pentamidine, melarsoprol, eflornithine, or a combination of nifurtimox and eflornithine [2,5]. These drugs have many shortcomings, including high toxicity and/or require administration by injection [6]. Thus, there is urgent need for the development of new therapeutics that are effective, safe, easy to administer, and affordable.

Methionyl-tRNA synthetase (MetRS) of *T. brucei* (*Tb*MetRS) plays an essential role in protein translation, providing the methionine charged tRNAs needed for biosynthesis of protein peptide chains [7]. It has been shown that *Tb*MetRS is an attractive drug target for the development of a new HAT treatment [8,9]. However, some previously reported inhibitors have drawbacks. For example, the aminoquinolone-based analogues of inhibitors that target bacterial MetRS [10,11] have poor PK profiles and poor membrane permeability despite potent *in vitro* activity against *T. brucei* parasites [8]. Urea-based inhibitors have improved pharmacokinetic characteristics and membrane permeability, but their potency against the parasites is suboptimal [9]. As part of our continued effort to discover novel MetRS inhibitors, a high-throughput screen of the NIH Molecular Libraries Small Molecule Repository was performed with *Tb*MetRS [12], leading to the discovery of compound **1**. In this paper we report the structure-guided optimization of **1** that resulted in novel *Tb*MetRS inhibitors with good potency, high selectivity, promising PK properties as well as brain permeability.

2. Results and discussion

2.1. Molecular design

Compound **1** (Fig. 1E) was identified as a weak inhibitor of *Tb*MetRS in a high-throughput orthogonal screening [12], showing

Abbreviations: AP, auxiliary pocket; CNS, central nervous system; EMP, enlarged methionine pocket; HAT, human African trypanosomiasis; MetRS, methionyl-tRNA synthetase; SAR, structure-activity relationship.

* Corresponding author.
** Corresponding author.
E-mail addresses: fbuckner@uw.edu (F.S. Buckner), erkang@uw.edu (E. Fan).
[1] Present address: Department of Biological Sciences, National University of Singapore, 117543, Singapore.

http://dx.doi.org/10.1016/j.ejmech.2016.10.024
0223-5234/

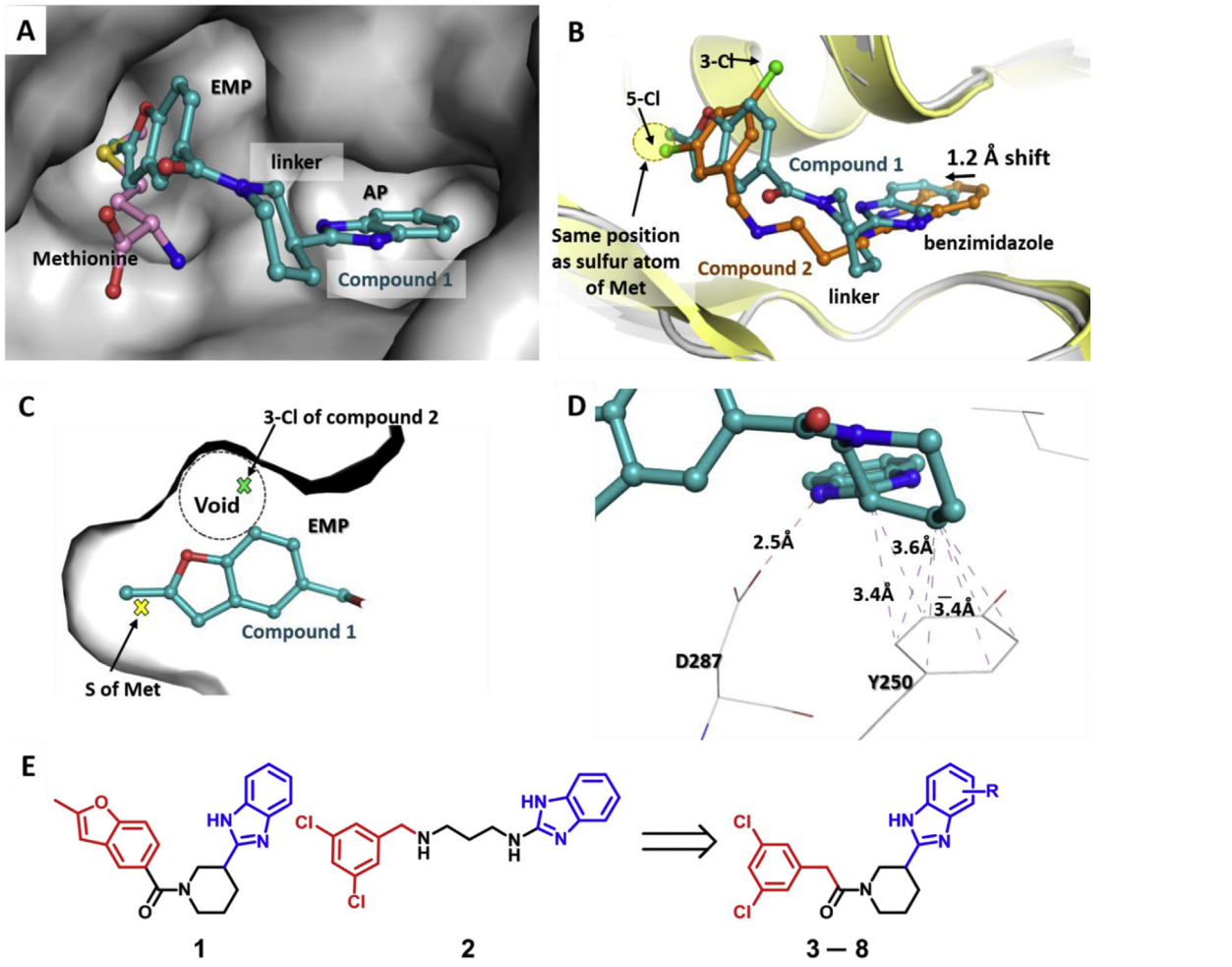


Fig. 1. Crystal structures of compounds **1** and **2** in complex with *Tb*MetRS and the design strategy of new compounds. (A) Compound **1** (cyan carbons) in the *Tb*MetRS active site, bound methionine from a previously determined structure (PDB: 4EG1) is shown as a reference (pink carbons) [13]; (B) Superimposed structures of compounds **1** (cyan) and **2** (orange) in complex with *Tb*MetRS; (C) A void existing in the EMP when the benzofuran group of compound **1** binds; (D) The piperidine ring in the linker of compound **1** interacts with Tyr250; (E) Inhibitor design strategy of combining compounds **1** and **2**. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

78% and 20% inhibition in a *Tb*MetRS aminoacylation assay at 10 and 1 μM respectively. Compound **1** can be chemically parsed into three parts: a benzimidazole moiety, a benzofuran moiety, and a linker connecting these two aryl rings. Similarly, compound **2** (Fig. 1E), a potent *Tb*MetRS inhibitor reported previously [12–15], also contains a benzimidazole moiety connected by a linker to an aryl ring (3,5-dichlorophenyl group). The benzimidazole moiety has already been identified as an effective fragment that binds the "auxiliary pocket" (AP) of *Tb*MetRS [14]. The new attachment mode of the benzimidazole group to the other aryl moiety (benzofuran) seen in compound **1** provided opportunities to arrive at *Tb*MetRS inhibitors with alternative linkers.

To aid the design of new inhibitors, a crystal structure of *Tb*MetRS with compound **1** (Fig. 1A) was obtained at 2.7 Å resolution (Table S1). The binding mode of compound **1** was compared to compound **2** bound to *Tb*MetRS and analyzed in detail (Fig. 1B). The benzimidazole group of compound **1** binds similarly to the "auxiliary pocket" (AP) compared to compound **2**. Both compounds interact with the AP by means of hydrophobic interactions with residues His289, Gly290 and Val473. A previously observed hydrogen bond interaction between the catalytic residue Asp287 and the benzimidazole moiety in compound **2** is also formed by compound **1** (Fig. 1D). However, there is a shift of 1.2 Å away from the binding pocket of the benzimidazole moiety of compound **1** with respect to **2** (Fig. 1B). At the other end of compound **1**, the benzofuran moiety occupies the "enlarged methionine pocket" (EMP) (Fig. 1C), which is formed mainly by hydrophobic residues (including the ones engaged in methionine binding) and interacts favorably with residue Phe522. However, compared to compound **2**, the benzofuran binding of compound **1** leaves a void that is filled by the 3-chlorine of the 3,5-dichloro-benzene moiety in compound **2** (Fig. 1C) providing in the latter case favorable hydrophobic interactions especially with residue Trp474. In the linker region, the piperidine ring in compound **1** restricts the conformational freedom of the linker. The piperidine ring is also engaged in hydrophobic interactions with residue Tyr250 (Fig. 1D), adding to the previously described *Tb*MetRS·inhibitor stacking interactions [13].

From comparing the binding modes of compounds **1** and **2**, we hypothesized that hybridizing elements of both compounds would lead to a novel chemotype exploiting the best features of both parental compounds (Fig. 1E). Considering that the 3,5-dichlorophenyl moiety in compound **2** was established as a preferred fragment for the EMP [9,14] and it binds more efficiently (no void) than the benzofuran moiety in compound **1**, we decided to fix it as the fragment to fill the EMP in all the designed compounds reported here. We focused our investigation on varying the

portion that occupies the AP as well as the linker. First, we made variants of the fragment occupying the AP. The shift in binding position of the benzimidazole as described in Fig. 1B suggests that adding substituents to the benzene ring of the benzimidazole moiety might improve contacts with the AP. Thus, compounds **3**–**9** with different substitutions on the benzimidazole (or imidazopyridine) moiety (Table 1) were designed, synthesized and evaluated, which led to the discovery of 5-chloro-benzimidazole (or 5-chloro-imidazopyridine) moiety as the favored fragment for binding the AP. Then, we continued our investigation of the linker part while maintaining the 3,5-dichlorophenyl and the 5-chloro-benzimidazole (or 5-chloro-imidazopyridine) moieties. A series of compounds **10**–**31** with varied linkers were designed, synthesized and evaluated as summarized in Table 2.

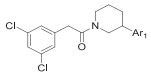
2.2. Compound synthesis

Several synthetic strategies were used to generate the target compounds. The synthetic routes used to generate inhibitors are shown in Schemes 1–5. Compounds **3**–**13** and **26** were prepared using the procedure shown in Scheme 1. Substituted carboxylic acids **32** were reacted with an appropriate aryl diamine **33** to generate the amide intermediates **34**. The amide intermediates were cyclized in acetic acid to generate the target compounds.

The synthetic route shown in Scheme 2 was used to synthesize compounds **14**–**25** and **31**. Linkers were introduced by nucleophilic substitution of 2-bromo-5-chloro-1H-imidazo[4,5-b]pyridine with appropriate amines, followed by Boc deprotection with TFA when necessary. Reductive amination of **37** with 3,5-dichlorobenzaldehyde produced the target compounds.

Linear linker compound **27** was synthesized following procedures shown in Scheme 3. The 1H-imidazo[4,5-b]pyridine ring was generated through reacting 6-chloropyridine-2,3-diamine with the

Table 1
Binding and inhibitory activity of compounds designed to investigate the benzimidazole substitution.



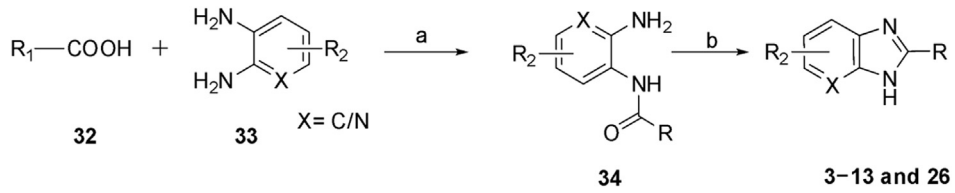
Compound number	Ar_1	ΔTm (°C)[a]	IC_{50} (nM)[b]	EC_{50} (nM)[c]
3		7.8	288	>5000
4		7.4	926	8810
5		8.0	184	3260
6		7.2	60	1397
7		–	<50	1058
8		7.0	138	3334
9		7.0	812	4603

[a] Values are the averages of two or more experiments.
[b] Control for *Tb*MetRS IC_{50} assay average ± SEM: **Met-SA1** [15] (9.9 nM ± 1.0 nM; n = 11 assays).
[c] Control for *Tb*EC_{50} assay average ± SEM: Pentamidine (0.90 nM ± 0.17 nM; n = 11).

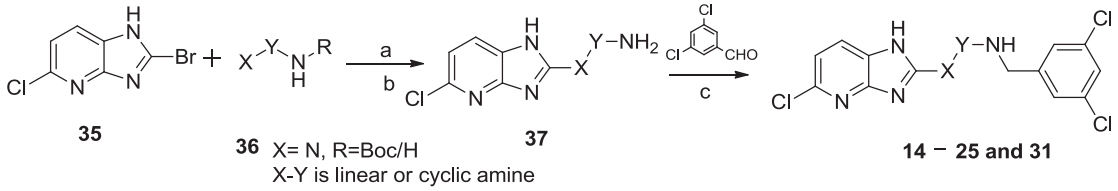
Table 2
Binding and inhibitory activity of compounds designed to investigate the linkers.

Compound number	Linker	X	ΔTm (°C)[a]	IC_{50} (nM)[b]	EC_{50} (nM)[c]
10		N	6.0	>10000	3763
11		C	12.6	580	1952
12		N	14.2	124	359
13		N	13.1	106	580
14		N	18.4	<50	82
15		N	13.1	192	999
16		N	19.0	<50	39
17		N	7.5	196	8802
18		N	14.4	<50	343
19		N	2.6	>10000	>20000
20		N	0	>10000	>20000
21		N	5.8	4970	13311
22		N	6.4	2166	15263
23		N	7.2	696	9013
24		N	11.4	<50	1539
25		N	6.8	682	1813
26		C	10.8	243	2154
27		N	12.8	<50	307
28		N	–	1054	1399
29		N	14.5	<50	61
30		C	14.4	85	539
31		N	18.4	<50	22[a]

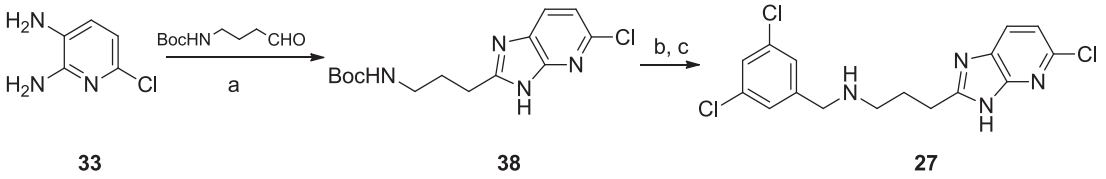
[a] Values are the average of two or more experiments.
[b] Control for *Tb*MetRS IC_{50} assay average ± SEM: **Met-SA1** [15] (9.9 nM ± 1.0 nM; n = 11 assays).
[c] Control for *Tb*EC_{50} assay average ± SEM: Pentamidine (0.90 nM ± 0.17 nM; n = 11).



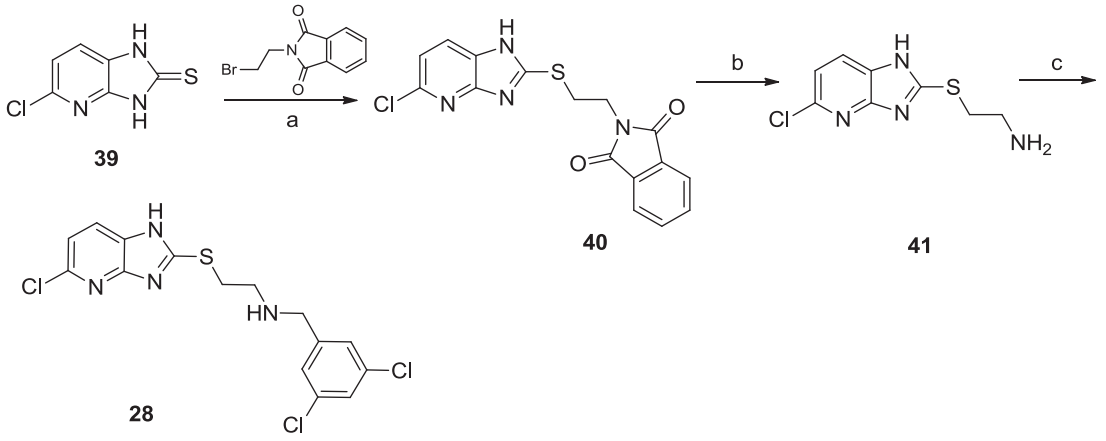
Scheme 1. Reagents and conditions: (a) EDC, pyridine, r.t. overnight; (b) X = C: AcOH, 60 °C, 3 h; X = N: AcOH, $POCl_3$, Microwave 150 °C, 1 h.



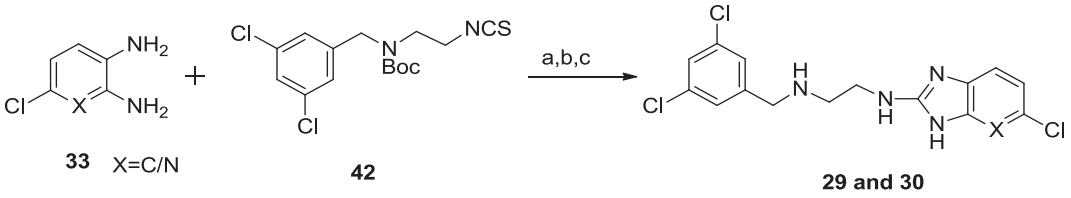
Scheme 2. Reagents and conditions: (a) Pyridine, MW, 100 °C, 30 min; (b) R = Boc, TFA, DCM, r.t., overnight; (c) DIPEA, $NaBH_3CN$, AcOH, CH_3OH, r.t., overnight.



Scheme 3. Reagents and conditions: (a) $NaHSO_3$, DMF, 100 °C; (b) TFA, CH_2Cl_2; (c) 3,5-dichloro benzaldehyde, MeOH, $NaBH_3CN$.



Scheme 4. Reagents and conditions: (a) K_2CO_3, CH_3CN, r.t. overnight; (b) H_2NNH_2, CH_3OH, MW, 70 °C, 20 min, (c) 3,5-dichlorobenzaldehyde, MeOH, $NaBH_3CN$.



Scheme 5. Reagents and conditions: (a) CH_3CN, reflux, overnight; (b) DIC, CH_3CN, reflux, overnight; (c) HCl, dioxane.

desired aldehyde. The Boc protected intermediate **38** was then deprotected with TFA. After Boc deprotection, reductive amination was performed to give compound **27**.

The synthetic route shown in Scheme 4 was used to obtain linear linker compound **28**. The linker was introduced by alkylating 5-chloro-1H-imidazo[4,5-b]pyridine-2(3H)-thione with 2-(2-bromoethyl)isoindoline-1,3-dione. The alkylated intermediate **40** was then deprotected with hydrazine to produce the free amine intermediate **41**. Reductive amination of **41** with 3,5-dichlorobenzaldehyde generated compound **28**.

Linear linker inhibitors **29** and **30** were prepared using the procedure shown in Scheme 5. The 1*H*-benzo[*d*]imidazole or 1*H*-imidazo[4,5-b]pyridine ring were generated by first reacting aryl diamines with isocyanate **42**, followed by treatment with DIC. Target compounds **29** and **30** were obtained by removing the Boc protecting group with HCl in dioxane.

2.3. Affinity of inhibitors for TbMetRS and potency against T. brucei parasites

All the compounds reported here were first assessed for binding to *Tb*MetRS by a thermal shift assay, followed by testing in a *T. brucei* growth inhibition assay. A good correlation was observed between ΔTm and EC_{50}, which is consistent with previous observations [8,14]. The higher the affinity the compound for the enzyme (higher ΔTm), the more potent the compound inhibits *T. brucei* parasite growth. These results support the hypothesis that the compounds act on target and their cellular activity is directly related to their affinity to the target. To evaluate the potency of the inhibitors, an enzymatic ATP depletion assay was performed as described previously [12]. For compounds with an IC_{50} below 50 nM (the enzyme concentration) the thermal shift magnitude should be used for potency ranking.

As shown in Table 1, all the compounds designed to investigate the effect of substitution on the benzimidazole ring (or imidazopyridine) were more potent than compound **1**. It was also noted that the substitution pattern on the benzimidazole ring has a significant impact on activity. Compound **3** without substitution on benzimidazole ring showed moderate enzyme inhibition with an IC_{50} of 288 nM against *Tb*MetRS. Introduction of a substitution to the C-4 of the benzimidazole ring (compound **4**) decreased potency compared to compound **3**. However, the C-5 substituted compounds (compounds **5**–**7**) showed improved activity compared to **3**. Especially, both of the 5-chloro substituted compounds **6** and **7** exhibited approximately 5-fold cellular activity improvement compared to **3**. Substitution at C-6 (compounds **8** and **9**) decreased potency compared to **6**. This investigation led to the discovery of 5-chloro-benzimidazole (or 5-chloro-imidazopyridine) moiety as a preferred fragment for the AP.

We then focused our investigation on the linker part while maintaining the 3,5-dichlorophenyl (preferred fragment for the EMP) and the 5-chloro-benzimidazole (or 5-chloro-imidazopyridine) (preferred fragments for the AP). The compounds with different linkers are summarized in Table 2. Changing the amide of the linker to an amine gave compound **10**, which decreased potency around 3-fold compared to **6**. As many of the previously reported *Tb*MetRS inhibitors contain a benzyl amine group, which is believed to be important for the binding between the ligand and the target [9,14], the N in the linker of **10** was shifted to a position equivalent to the benzyl position of the 3,5-dichlorophenyl ring, which resulted in compounds **11** and **12**. Compared to compound **6**, compound **11** did not show improvement in cellular potency though the affinity for the enzyme improved (5.4 °C higher ΔTm). In contrast, compound **12** improved both cellular potency (4-fold lower EC_{50}) and enzyme affinity (7.0 °C higher ΔTm) compared to **6**. As compound **12** contains two chiral centers, thus a mixture of 4 stereoisomers, it is reasonable to predict that the actual isomer responsible for the activity would be more potent. Presently, we were only able to obtain one single stereoisomer **13** which showed slightly worse activity compared to compound **12**. To reduce the number of chiral centers, the cyclohexane of **12** was replaced by a piperidine in compound **14**, with only one chiral center. The racemic **14** exhibited improved potency (~4-fold lower EC_{50}) and affinity (4.2 °C higher ΔTm) compared to compound **12**. After synthesis of the stereoisomers of **14**, the *S*-isomer **16** was shown to be markedly more potent (>25 fold lower EC_{50}) with higher affinity (5.9 °C higher ΔTm) compared to the *R*-isomer **15**, indicating that the *S*-isomer was responsible for the activity of **14**.

Based on compound **16**, alternative linkers between the 3,5-dichlorophenyl moiety and the piperidine ring were investigated. Replacement of the amine by an ether resulted in compound **17**. Changing the amine into two different versions of an amide furnished compounds **19** and **20**. Unfortunately, these three compounds exhibited poorer potency. Shifting the N in the linker to the benzyl position gave compound **18**, which showed worse affinity and potency compared to compound **14**. Shifting the 3,5-dichlorophenyl moiety from the meta-position to the para-position of the piperidine ring gave compound **21**. This modification led to loss of potency, indicating the meta-substitution pattern is important for activity. In addition, the ring size of the cyclic linker was explored. A 4-member ring linker (compound **22)** was detrimental. 5-member ring linker compounds were much weaker compared to compound **16**, though the *S*-isomer **24** was ~6-fold more potent than the *R*-isomer **23**. The 7-member ring compound **25** did not shown improvement either. Thus, the 6-member ring is the preferred cyclic linker so far.

Modifying the rigid cyclic linker into a more flexible linear linker while keeping the hydrophobic interactions established with *Tb*MetRS residue Tyr250 would remove the chiral center and be predicted to maintain the binding affinity. The linear linker compounds **26** and **27** exhibited comparable potency compared to their equivalent cyclic linker compounds **11** and **12**. Linear linker compound **29** was comparable to its equivalent cyclic linker compound **16** as well, demonstrating that the linear linker is well tolerated as a replacement of the cyclic linker. Encouragingly, introducing a methyl group to the imidazopyridine side NH group of compound **29** resulted in compound **31**, which improved cellular potency (~3-fold lower EC_{50}) and affinity (3.9 °C higher ΔTm) compared to compound **29**. Compound **31** was the most potent compound observed of this series.

2.4. Crystal structures of TbMetRS in complex with compounds *16* and *31*

To check for unexpected changes in the binding mode, we obtained crystal structures of *Tb*MetRS in complex with compounds **16** and **31**. The structures were solved at 2.8 and 2.4 Å resolution, respectively (Table S1). Compounds **16** and **31** bind similarly to the target protein (Fig. 2). The almost perfect superposition of the linkers in compounds **16** and **31** shows that the linear linker maintained the desired conformation for compound **31**, allowing the compound to interact hydrophobically with *Tb*MetRS through the two ring systems (3,5-dichlorophenyl moiety and the 5-chloro-imidazopyridine) as well as through the methylated linker. Additionally, the distance between residue Tyr250 and the methyl substituent in compound **31** (~3.6 Å) is more favorable than the distance between the corresponding carbon atom in the cyclic ring of compound **16** (~4.0 Å) (Fig. 2A). The void in the EMP shown in the structure of *Tb*MetRS bound to compound **1** is now filled by compounds **16** and **31** (Fig. 2B). The imidazopyridine moieties of compounds **16** and **31** are in a similar position as the corresponding group in compound **1**. However, the presence of the 5-chloro substituent fills the AP better than compound **1**. This Cl adds an interaction with *Tb*MetRS residue Leu456 (Fig. 2C).

The availability of inhibitor-bound crystal structures should also help to understand the structure-activity relationship (SAR) data generated in Tables 1 and 2 in order to guide future inhibitor design. Note that in the SAR tables, some enzyme inhibition data are shown as <50 nM because the enzyme concentration used for the activity assay was 50 nM (see Experimental Procedures section). Therefore,

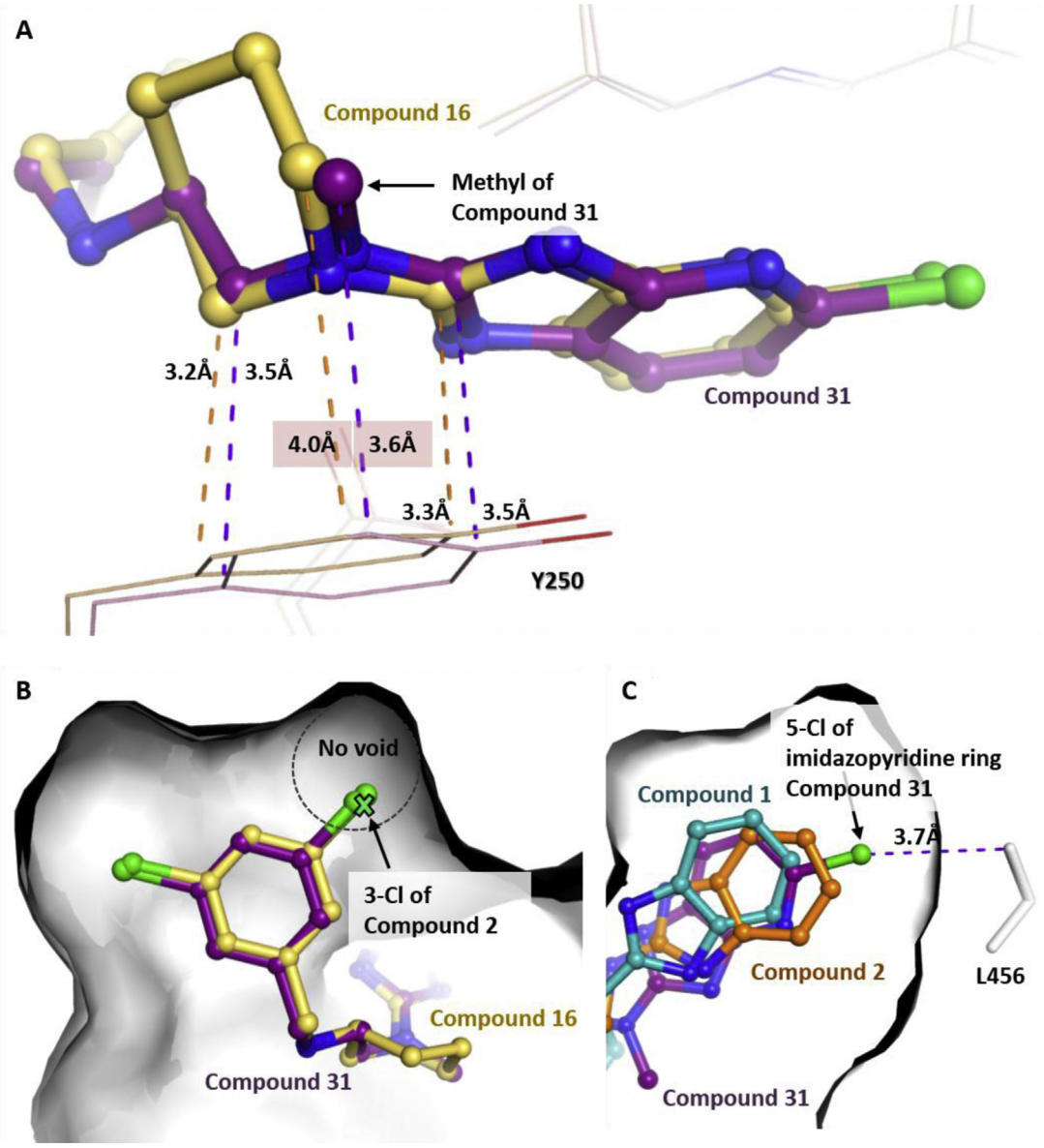


Fig. 2. Crystal structures of compounds **16** and **31** in complex with *Tb*MetRS. (A) Interactions between the linker of compounds **16** (yellow) and **31** (magenta) and the protein residue Tyr250; (B) 3,5-dichlorophenyl ring of compounds **16** and **31** bound to the EMP of the superimposed *Tb*MetRS·inhibitor structures; (C) Benzimidazole or imidazopyridine ring of compounds **1** (cyan), **2** (gold) and **31** (magenta) bound to the AP of the superimposed *Tb*MetRS·inhibitor structures. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for compounds with IC_{50} <50 nM, enzyme inhibition data only confirms on-target inhibitory potency but is not useful for quantitative SAR analysis. Fortunately, the thermal shift data as an affinity gauge is available for most compounds. In fact, as observed previously [8,14], there is an excellent linear correlation between ΔTm and pEC_{50} ($-\log EC_{50}$) for data in Tables 1 and 2 ($R^2 = 0.84$ for linear regression). Given that compounds **16** and **31**, the most potent inhibitors in the two chemical series under study, have essentially the same binding modes toward *Tb*MetRS, one may assume that these two compounds present the optimal binding modes in occupying the EMP and the AP pockets while having the least strain in the ring- or linear-linker regions. Therefore, the binding modes of **16** and **31** can be used as the starting points for comparison to molecular docking studies of other inhibitors described in this work to check for deviations in binding poses inside the EMP and the AP pockets, or in energetic strains of the linker, and the correlation to ΔTm and EC_{50} data. New designs of *Tb*MetRS inhibitors that can reduce linker strains and preserve optimal interactions at the EMP and the AP pockets might lead to further improvements in affinity.

2.5. Selectivity and toxicity of select inhibitors

With novel potent inhibitors in hand, we selected a group of compounds to examine their selectivity and toxicity. The selected compounds were tested for host cell toxicity using a human lymphoblast cell line (CRL-8155) and a hepatocellular carcinoma cell line (Hep G2) (Table 3). These compounds exhibited low toxicities to mammalian cell lines with EC_{50}s ranging from 9.9 to >50 μM. The most potent compounds on *T. brucei* (**16** and **31**) exhibited high selectivity indices of 751 and 1027, respectively.

2.6. Pharmacokinetics and permeability of compounds *16* and *31*

Compounds **16** and **31** were also subjected to PK studies in mice

Table 3
Host cell toxicity data of select inhibitors.

Compound number	CRL-8155[a] EC_{50} (μM)	Hep G2[b] EC_{50} (μM)	Compound number	CRL-8155[a] EC_{50} (μM)	Hep G2[b] EC_{50} (μM)
6	22.8	35.4	**27**	45.7	45.6
12	33.1	32.3	**29**	33.8	45.9
14	46.2	>50.0	**30**	14.0	9.9
16	29.3	49.0	**31**	22.6	39.2
18	15.4	29.2			

[a] Control for CRL-8155 EC_{50} assay average ± SEM: Quinacrine (2.12 μM ± 0.3 μM; n = 5).
[b] Control for HepG2 EC_{50} assay average ± SEM: Quinacrine (11.9 μM ± 0.6 μM; n = 4).

following oral administration at 50 mg/kg. Compound **16** exhibited outstanding PK properties, showing a Cmax at 37.6 μM, and a AUC of 6223 min•μmol/L. Compound **31** exhibited promising PK properties with Cmax at 9.7 μM and AUC of 952 min•μmol/L.

The potential permeability of compounds to the CNS is important for treating late stage HAT. Hence, compounds **16** and **31** were tested for brain permeability in mice at a dose of 5 mg/kg IP. Comparing plasma with brain tissue concentrations at a single time-point (60 min), the cyclic linker compound **16** was observed to have poor permeability, with nearly undetectable brain levels at 60 min after IP injection. This permeability result suggests that the cyclic-linker compound is unlikely to pass through the blood-brain barrier, thus makes it a poor candidate for treating HAT. In contrast, the linear linker compound **31** exhibited good permeability, with brain/plasma ratio of 0.27 at 60 min after IP injection. This feature makes compound **31** a promising lead compound for developing therapies for treating late stage HAT.

3. Conclusions

In summary, starting with a weak screening hit **1**, two new series of potent *T. brucei* methionyl tRNA synthetase inhibitors were obtained through structure-guided design. The best compounds **16** in the cyclic-linker series and **31** in the linear-linker series were potent in a *T. brucei* growth inhibition assay, with EC_{50}s of 39 and 22 nM, respectively. These compounds also showed low toxicity to the mammalian cells, resulting in a high selectivity index. Compound **16** exhibited outstanding PK properties but poor brain permeability, therefore further investigations are ongoing with the aim to improve its permeability. Compound **31** exhibited good PK properties and, importantly, it showed moderately good brain penetration in mice. These studies provide novel lead compounds for developing drugs for treating HAT.

4. Experimental procedures

4.1. General chemistry

Unless otherwise stated, all chemicals were purchased from commercial suppliers and used without further purification. Microwave irradiation was performed on a CEM Discover System. Reaction progress was monitored by thin-layer chromatograph on silica gel containing an inert binder and a fluorescent indicator (activated at 254 nm) coated flexible sheet (J. T. Baker). Chromatography was performed using an automated flash chromatography system, eluting on pre-packed silica gel columns with CH_2Cl_2/MeOH or cyclohexane/Ethyl acetate gradient solvent system. The purification by preparative RP-HPLC was performed on Waters Xterra Prep RP18 OBD 5 μM (19 mm × 50 mm), eluting with a CH_3CN/H_2O solvent system with 0.1% TFA. The purity of all final compounds was determined by analytical LCMS using an Onyx Monolithic C18 column (4.6 mm × 100 mm) (Phenomenex, Torrance, CA) and eluting with CH_3CN/H_2O solvent system (+0.1% TFA). The products were detected by UV at 220 nm. All compounds were determined to be >95% pure by this method. The mass spectra were recorded with an Ion Trap Mass Spectrometer (Agilent, Santa Clara, CA). NMR spectra were recorded on Bruker 300 or 500 MHz spectrometers at ambient temperature. Chemical shifts are reported in parts per million (δ) referenced to the internal standards (7.26 ppm for $CDCl_3$, 3.34 ppm for CD_3OD and 2.50 ppm for $(CD_3)_2SO$) and coupling constants in Hz.

4.2. General procedures for the synthesis of compounds *3*–*13* and *26* (Scheme 1)

(a) Substituted carboxylic acid **32** (1 eq) and aromatic diamine **33** (1 eq) in pyridine was added EDC (1.5 eq). The mixture was stirred at r.t. overnight, and pyridine was then removed under reduced pressure. After addition of saturated aqueous sodium bicarbonate to the residue, the mixture was extracted with EA. The organic layer was dried over anhydrous sodium sulfate, and concentrated in vacuum. Purification through flash chromatography on silica gel eluted with MeOH-DCM (0.5% ammonia hydroxide) gave the amide intermediate **34.**

(b) When X = C: The amide intermediate **34** (1 eq) was dissolved in glacial acetic acid and heated at 60 °C for 3 h. The reaction was concentrated in vacuum and the residue partitioned between saturated sodium bicarbonate and DCM. The organic extract was dried over anhydrous sodium sulfate, and concentrated in vacuum. Purification through flash chromatography on silica gel eluted with MeOH-DCM (0.5% ammonia hydroxide) gave target compound.

When X = N: The amide intermediate **34** was dissolved in glacial acetic acid, $POCl_3$ (3 eq) was added, the mixture was microwave irradiated at 150 °C for 1 h. The reaction was concentrated in vacuum and the residue partitioned between saturated sodium bicarbonate and DCM. The organic extract was dried over anhydrous sodium sulfate, and concentrated in vacuum. The residue was purified through flash chromatography on silica gel eluted with MeOH-DCM (0.5% ammonia hydroxide) to give the target compound.

4.2.1. Compound *3*

1-(3-(1H-benzo[d]imidazol-2-yl)piperidin-1-yl)-2-(3,5-dichlorophenyl) ethanone was synthesized using 1-(2-(3,5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and benzene-1,2-diamine. ^{1}H NMR (500 MHz, MeOD) δ 7.85–7.71 (m, 2H), 7.67–7.54 (m, 2H), 7.31 (s, 1H), 7.27–7.21 (m, 2H), 4.74 (d, J = 12.0 Hz, 1H), 4.03 (d, J = 13.7 Hz, 1H), 3.96–3.89 (m, 1H), 3.86 (d, J = 3.9 Hz, 2H), 3.44–3.33 (m, 2H), 2.47–2.30 (m, 1H), 2.12 (dt, J = 13.7, 7.0 Hz, 1H), 1.93–1.79 (m, 1H), 1.76–1.58 (m, 1H). MS (ESI) $(M + H)^+$ = 389.1. HPLC analysis: 99.0% purity.

4.2.2. Compound ***4***

1-*(3-(4-chloro-1H-benzo[d]imidazol-2-yl)piperidin-1-yl)-2-(3,5-dichlorophenyl) ethanone* was synthesized using 1-(2-(3, 5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and 3-chlorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55 (s, 1H), 7.47 (d, J = 8.6 Hz, 1H), 7.20 (dd, J = 8.6, 1.8 Hz, 1H), 7.16 (d, J = 5.4 Hz, 1H), 7.01 (d, J = 1.1 Hz, 2H), 4.39 (dd, J = 13.6, 4.2 Hz, 1H), 3.74 (d, J = 13.8 Hz, 1H), 3.68 (d, J = 6.9 Hz, 2H), 3.66–3.59 (m, 1H), 3.48–3.38 (m, 1H), 3.36–3.26 (m, 1H), 2.70–2.59 (m, 1H), 2.18–2.07 (m, 1H), 1.65–1.52 (m, 1H), 1.45–1.31 (m, 1H). MS (ESI) $(M + H)^+$ = 423.6. HPLC analysis: >99.9% purity.

4.2.3. Compound ***5***

2-(3,5-dichlorophenyl)-1-(3-(5-fluoro-1H-benzo[d]imidazol-2-yl) piperidin-1-yl)ethanone was synthesized using 1-(2-(3, 5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and 4-fluorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.48 (dd, J = 8.6, 4.5 Hz, 1H), 7.26 (dd, J = 9.0, 1.8 Hz, 1H), 7.19 (s, 1H), 7.04 (s, 2H), 7.00 (td, J = 9.2, 2.2 Hz, 1H), 4.29 (dd, J = 13.7, 5.3 Hz, 1H), 3.84 (dd, J = 14.1, 3.2 Hz, 1H), 3.72 (d, J = 6.6 Hz, 2H), 3.65–3.56 (m, 1H), 3.54–3.49 (m, 1H), 3.34–3.22 (m, 1H), 2.61 (dt, J = 10.4, 6.0 Hz, 1H), 2.17 (qd, J = 8.6, 3.9 Hz, 1H), 1.67–1.52 (m, 1H), 1.50–1.44 (m, 1H). MS (ESI) $(M + H)^+$ = 407.8. HPLC analysis: 99.8% purity.

4.2.4. Compound ***6***

1-(3-(5-chloro-1H-benzo[d]imidazol-2-yl) piperidin-1-yl)-2-(3,5-dichlorophenyl) ethanone was synthesized using 1-(2-(3, 5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and 4-chlorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.52 (s, 1H), 7.44 (d, J = 7.3 Hz, 1H), 7.19–7.11 (m, 2H), 7.00 (s, 2H), 4.35 (d, J = 11.2 Hz, 1H), 3.77–3.70 (m, 1H), 3.68 (d, J = 6.9 Hz, 2H), 3.65–3.56 (m, 1H), 3.44 (t, J = 11.2 Hz, 1H), 3.27 (s, 1H), 2.65–2.54 (m, 1H), 2.12 (dd, J = 16.0, 7.0 Hz, 1H), 1.62–1.49 (m, 1H), 1.44–1.34 (m, 1H). MS (ESI) $(M + H)^+$ = 423.6. HPLC analysis: 98.5% purity.

4.2.5. Compound ***7***

1-(3-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)piperidin-1-yl)-2-(3,5-dichlorophenyl)ethanone was synthesized using 1-(2-(3, 5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and 6-chloropyridine-2,3-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.92 (d, J = 8.2 Hz, 1H), 7.35–7.31 (m, 1H), 7.21 (d, J = 8.3 Hz, 1H), 7.19–7.15 (m, 2H), 4.18 (dd, J = 13.5, 5.4 Hz, 1H), 3.93 (dd, J = 13.7, 2.7 Hz, 1H), 3.72 (d, J = 8.1 Hz, 2H), 3.62–3.47 (m, 2H), 3.30–3.18 (m, 1H), 2.56 (ddd, J = 13.2, 11.0, 6.7 Hz, 1H), 2.17 (ddd, J = 13.2, 8.5, 4.1 Hz, 1H), 1.64–1.54 (m, 1H), 1.55–1.43 (m, 1H). MS (ESI) $(M + H)^+$ = 424.5. HPLC analysis: 95.0% purity.

4.2.6. Compound ***8***

1-(3-(5-chloro-6-fluoro-1H-benzo[d]imidazol-2-yl)piperidin-1-yl)-2-(3,5-dichlorophenyl)ethanone was synthesized using 1-(2-(3, 5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and 4-chloro-5-fluorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55 (d, J = 6.4 Hz, 1H), 7.32 (d, J = 9.0 Hz, 1H), 7.14 (s, 1H), 7.00 (s, 2H), 4.31 (dd, J = 14.7, 5.1 Hz, 1H), 3.75 (dd, J = 12.5, 10.2 Hz, 1H), 3.69 (d, J = 4.5 Hz, 2H), 3.66–3.56 (m, 1H), 3.51–3.42 (m, 1H), 3.31–3.23 (m, 1H), 2.66–2.54 (m, 1H), 2.18–2.10 (m, 1H), 1.63–1.53 (m, 1H), 1.39 (dt, J = 12.9, 9.2 Hz, 1H). MS (ESI) $(M + H)^+$ = 441.6. HPLC analysis: 96.6% purity.

4.2.7. Compound ***9***

1-(3-(5,6-dichloro-1H-benzo[d]imidazol-2-yl)piperidin-1-yl)-2-(3,5-dichlorophenyl)ethanone was synthesized using 1-(2-(3, 5-dichlorophenyl)acetyl)piperidine-3-carboxylic acid and 4,5-dichlorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.66 (s, 2H), 7.16 (s, 1H), 7.01 (s, 2H), 4.36 (dd, J = 13.3, 3.8 Hz, 1H), 3.76–3.70 (m, 1H), 3.70–3.67 (m, 1H), 3.64 (d, J = 13.6 Hz, 1H), 3.46 (t, J = 10.2 Hz, 1H), 3.38–3.27 (m, 1H), 2.64 (dd, J = 13.4, 4.7 Hz, 1H), 2.15 (dd, J = 13.5, 9.9 Hz, 1H), 1.66–1.54 (m, 1H), 1.43–1.31 (m, 1H). MS (ESI) $(M + H)^+$ = 457.4. HPLC analysis: 99.6% purity.

4.2.8. Compound ***10***

5-chloro-2-(1-(3,5-dichlorophenethyl)piperidin-3-yl)-1H-imidazo[4,5-b]pyridine was synthesized using 1-(3,5-dichlorophenethyl) piperidine-3-carboxylic acid and 6-chloropyridine-2,3-diamine. ^{1}H NMR (500 MHz, MeOD) δ 8.08 (d, J = 8.3 Hz, 1H), 7.47 (d, J = 8.3 Hz, 1H), 7.44–7.35 (m, 3H), 4.10 (d, J = 11.4 Hz, 1H), 3.86–3.71 (m, 2H), 3.56–3.47 (m, 3H), 3.26–3.14 (m, 3H), 2.42 (d, J = 11.4 Hz, 1H), 2.29–2.18 (m, 1H), 2.18–2.05 (m, 1H), 2.04–1.90 (m, 1H). MS (ESI) $(M + H)^+$ = 424.5. HPLC analysis: >99.9% purity.

4.2.9. Compound ***11***

3-(5-chloro-1H-benzo[d]imidazol-2-yl)-N-(3,5-dichlorobenzyl) cyclohexanamine was synthesized using 3-((*tert*-butoxycarbonyl) (3,5-dichlorobenzyl)amino) cyclohexane carboxylic acid and 4-chlorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.50–7.43 (br, 1H), 7.38 (d, J = 6.3 Hz, 1H), 7.42–7.35 (m, 3H), 7.16 (dd, J = 8.5, 1.7 Hz, 1H), 3.85 (s, 2H), 3.10 (s, 1H), 2.85 (s, 1H), 2.33 (d, J = 12.4 Hz, 2H), 2.01 (s, 2H), 2.00–1.90 (m, 1H), 1.87–1.78 (m, 1H), 1.74 (dd, J = 21.6, 9.0 Hz, 1H), 1.63 (dd, J = 21.6, 9.0 Hz, 1H), 1.51–1.37 (m, 2H). MS (ESI) $(M + H)^+$ = 409.5. HPLC analysis: 95.6% purity.

4.2.10. Compound ***12***

3-*(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl) cyclohexan-amine* was synthesized using 3-((*tert*-butoxycarbonyl) (3,5-dichlorobenzyl)amino)cyclohexane carboxylic acid and 6-chloropyridine-2,3-diamine. ^{1}H NMR (500 MHz, MeOD) δ 8.21 (t, J = 8.6 Hz, 1H), 7.67–7.60 (m, 3H), 7.60–7.54 (m, 1H), 4.34 (s, 2H), 3.59–3.52 (m, 1H), 3.51–3.41 (m, 1H), 2.83–2.73 (m, 1H), 2.35 (d, J = 10.5 Hz, 1H), 2.27 (d, J = 11.7 Hz, 1H), 2.16 (d, J = 11.5 Hz, 1H), 2.07–1.93 (m, 1H), 1.71 (m, 3H). MS (ESI) $(M + H)^+$ = 410.5; HPLC analysis: 96.0% purity.

4.2.11. Compound ***13***

(1R,3S)-3-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl) cyclohexanamine was synthesized using (1S,3R)-3-((*tert*-butoxycarbonyl) (3,5-dichlorobenzyl)amino)cyclohexane carboxylic acid and 6-chloropyridine-2,3-diamine. ^{1}H NMR (500 MHz, CDCl3) δ 8.05–7.88 (br, 1H), 7.26 (s, 1H), 7.25–7.23 (m, 2H), 7.22 (s, 1H), 3.83 (dd, J = 33.3, 13.8 Hz, 2H), 3.20 (t, J = 10.4, 1H), 2.74 (t, J = 10.4, 1H), 2.42 (d, J = 12.5 Hz, 1H), 2.18–2.10 (m, 1H), 2.09–1.97 (m, 1H), 1.96–1.87 (m, 1H), 1.74–1.62 (m, 2H), 1.49 (td, J = 12.0, 3.4 Hz, 1H), 1.27 (dd, J = 16.4, 9.3 Hz, 1H). MS (ESI) $(M + H)^+$ = 410.5; HPLC analysis: >99.9% purity.

4.2.12. Compound ***26***

3-(6-chloro-1H-benzo[d]imidazol-2-yl)-N-(3, 5-dichlorobenzyl) propan-1-amine in HCl salt form. The key intermediate *tert*-butyl (3-(6-chloro-1H-benzo[*d*]imidazol-2-yl)propyl)carbamate for the synthesis of **26** was synthesized using 4-((*tert*-butoxycarbonyl) amino) butanoic acid and 4-chlorobenzene-1,2-diamine. The target compound was synthesized using *tert*-butyl (3-(6-chloro-1H-benzo [*d*]imidazol-2-yl)propyl)carbamate. ^{1}H NMR (500 MHz, MeOD) δ 7.91 (d, J = 8.3 Hz, 1H), 7.42–7.36 (m, 3H), 7.33–7.27 (m, 2H), 4.01 (s, 2H), 3.01 (t, J = 7.2 Hz, 2H), 2.81 (t, J = 6.6 Hz, 2H), 2.12–2.08 (m, 2H). MS: (ESI) $(M + H)^+$ = 369.6. HPLC analysis: 97.5% purity.

*4.3. General procedures for the synthesis of compounds **14**–**25** and **31** (Scheme 2)*

(a) Compound 2-bromo-5-chloro-1H-imidazo[4,5-b]pyridine **35** (1 eq) was dissolved in pyridine, substituted amine **36** (3 eq) was added and the reaction mixture was microwave irradiated at 100 °C for 30 min. The reaction mixture was concentrated in vacuum and the residue partitioned between saturated sodium bicarbonate and DCM. The organic extract was dried over anhydrous sodium sulfate, and concentrated in vacuum. The residue was purified through flash chromatography on silica gel eluted with MeOH/DCM (0.5% ammonia hydroxide) to give intermediate **37.**

(b) When necessary, the intermediate **37** obtained from above step was dissolved in DCM, the solution was cooled to 0 °C, then TFA (5 eq) was added dropwise. The mixture was stirred at r.t. overnight, the solvent was removed in vacuum. HCl in methanol was added to the residue, and the solvent was removed to give the HCl salt intermediate.

(c) The HCl salt intermediate obtained above was dissolved in CH_3OH, DIPEA (1.2 eq) was added and the solution was stirred at r.t. for 10 min. Then 3,5-dichlorobenzaldehyde(1.2 eq) was added, following by adding AcOH (2 eq) and $NaBH_3CN$ (2eq). The mixture was stirred at r.t. overnight. The solvent was removed under reduced pressure and the residue was extracted by DCM. The organic phase was dried over sodium sulfate. Solvent removal in vacuum and purification through flash chromatography on silica gel with MeOH-DCM (0.5% ammonia hydroxide) elution gave the target compound.

*4.3.1. Compound **14***

1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl)piperidin-3-amine was synthesized using *tert*-butyl piperidin-3-ylcarbamate. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.65–7.37 (br, 1H), 7.19 (s, 1H), 7.12 (s, 2H), 7.08–6.84 (br, 1H), 4.20–3.95 (br, 1H), 3.91–3.77 (br, 1H), 3.71 (s, 2H), 3.34–2.87 (m, 1H), 2.77 (s, 1H), 2.02–1.91 (m, 1H), 1.90–1.77 (m, 1H), 1.68–1.50 (m, 1H), 1.50–1.33 (m, 2H). MS (ESI) $(M + H)^+ = 411.7$; HPLC analysis: 99.6% purity.

*4.3.2. Compound **15***

(R)-1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl) piperidin-3-amine was synthesized using (R)-*tert*-butyl piperidin-3-ylcarbamate. ^{1}H NMR (500 MHz, CDCl3) δ 7.64–7.43 (br, 1H), 7.21 (t, J = 1.8 Hz, 1H), 7.15–7.09 (m, 2H), 7.07–6.78 (br, 1H), 4.32–4.02 (br, 1H), 4.02–3.78 (br, 1H), 3.73 (s, 2H), 3.39–2.89 (m, 2H), 2.89–2.63 (br, 1H), 2.09–1.92 (m, 1H), 1.91–1.74 (m, 1H), 1.73–1.52 (m, 1H), 1.54–1.32 (m, 1H). MS (ESI) $(M + H)^+ = 412.0$; HPLC analysis: > 99.9% purity.

*4.3.3. Compound **16***

(S)-1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl) piperidin-3-amine was synthesized using (S)-*tert*-butyl piperidin-3-ylcarbamate. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.64–7.43 (br, 1H), 7.21 (t, J = 1.8 Hz, 1H), 7.18–7.11 (m, 2H), 7.08–6.91 (br, 1H), 4.19–3.95 (br, 1H), 3.93–3.77 (br, 1H), 3.75 (s, 2H), 3.33–2.97 (m, 2H), 2.85–2.71 (br, 1H), 2.03–1.92 (m, 1H), 1.91–1.79 (m, 1H), 1.69–1.55 (m, 1H), 1.55–1.40 (m, 1H). MS (ESI) $(M + H)^+ = 411.4$; HPLC analysis: 95.6% purity.

*4.3.4. Compound **17***

(S)-5-chloro-2-(3-((3,5-dichlorobenzyl)oxy)piperidin-1-yl)-1H-imidazo[4,5-b]pyridine was synthesized using (S)-3-((3,5-dichlorobenzyl)oxy)piperidine. ^{1}H NMR (500 MHz, MeOD) δ 7.69 (d, J = 8.3 Hz, 1H), 7.34 (d, J = 8.3 Hz, 1H), 7.24–7.21 (m, 1H), 7.17–7.14 (m, 2H), 4.67 (d, J = 12.7 Hz, 1H), 4.55 (d, J = 12.7 Hz, 1H), 4.56 (s, 1H), 4.54 (s, 1H), 3.99 (dd, J = 14.0, 3.7 Hz, 1H), 3.95–3.86 (m, 2H), 3.78 (d, J = 13.8 Hz, 1H), 3.61 (ddd, J = 13.6, 10.5, 3.0 Hz, 1H), 2.15–2.04 (m, 2H), 2.01–1.93 (m, 1H), 1.82–1.73 (m, 1H). MS (ESI) $(M + H)^+ = 412.8$. HPLC analysis: 97.5% purity.

*4.3.5. Compound **18***

3,5-dichloro-N-((1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl) piperidin-3-yl)methyl)aniline was synthesized using 3,5-dichloro-*N*-(piperidin-3-ylmethyl)aniline. ^{1}H NMR (500 MHz, MeOD) δ 7.70 (d, J = 7.8 Hz, 1H), 7.39–7.27 (m, 1H), 6.63–6.49 (m, 3H), 4.05 (d, J = 9.2 Hz, 1H), 3.96–3.89 (m, 1H), 3.49–3.41 (m, 1H), 3.12 (dd, J = 29.3, 15.1 Hz, 2H), 2.18–1.91 (m, 3H), 1.85–1.69 (m, 1H), 1.58–1.42 (m, 1H), 1.37–1.21 (m, 1H). MS (ESI) $(M + H)^+ = 412.0$; HPLC analysis: 95.0% purity.

*4.3.6. Compound **19***

(S)-1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorophenyl) piperidine-3-carboxamide was synthesized using (R)-*N*-(3,5-dichlorophenyl)piperidine-3-carboxamide. ^{1}H NMR (500 MHz, DMSO) δ 10.4 (s, 1H), 7.74–7.66 (m, 1H), 7.49 (d, J = 8.0 Hz, 1H), 7.30 (t, J = 1.7 Hz, 1H), 7.03–6.88 (m, 1H), 4.27 (d, J = 12.3 Hz, 1H), 4.12 (d, J = 12.9 Hz, 1H), 3.28–3.17 (m, 1H), 3.11 (t, J = 11.4 Hz, 1H), 2.63 (tt, J = 10.9, 3.5 Hz, 1H), 2.05 (d, J = 10.7 Hz, 1H), 1.82 (d, J = 13.2 Hz, 1H), 1.74 (qd, J = 12.4, 3.5 Hz, 1H), 1.55 (qd, J = 12.5, 4.2 Hz, 1H). MS (ESI) $(M + H)^+ = 426.0$; HPLC analysis: > 99.9% purity.

*4.3.7. Compound **20***

(S)-3,5-dichloro-N-(1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl) piperidin-3-yl)benzamide was synthesized using (R)-3,5-dichloro-*N*-(piperidin-3-yl)benzamide. ^{1}H NMR (500 MHz, DMSO) δ 8.65 (d, J = 7.8 Hz, 1H), 7.85 (d, J = 1.8 Hz, 2H), 7.81 (t, J = 1.8 Hz, 1H), 7.46 (d, J = 8.0 Hz, 1H), 6.94 (d, J = 7.8 Hz, 1H), 4.14 (dd, J = 12.6, 3.7 Hz, 1H), 4.06–3.91 (m, 2H), 3.21–3.06 (m, 2H), 2.02–1.92 (m, 1H), 1.92–1.82 (m, 1H), 1.70–1.55 (m, 2H). MS (ESI) $(M + H)^+ = 426.0$; HPLC analysis: > 99.9% purity.

*4.3.8. Compound **21***

1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl)piperidin -4-amine was synthesized using *tert*-butyl piperidin-4-ylcarbamate. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80–7.43 (br, 1H), 7.35–7.21 (m, 3H), 7.13–6.76 (br, 1H), 4.37–3.97 (m, 2H), 3.81 (s, 2H), 3.29–3.06 (br, 2H), 2.78 (s, 1H), 2.00–1.92 (br, 2H), 1.53–1.40 (br, 2H). MS (ESI) $(M + H)^+ = 411.7$; HPLC analysis: 95.5% purity.

*4.3.9. Compound **22***

1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl)azetidin-3-amine was synthesized using *tert*-butyl azetidin-3-ylcarbamate. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80–7.43 (br, 1H), 7.35–7.21 (m, 3H), 7.13–6.76 (br, 1H), 4.37–3.97 (m, 2H), 3.81 (s, 2H), 3.29–3.06 (br, 2H), 2.78 (s, 1H), 2.00–1.92 (br, 2H), 1.53–1.40 (br, 2H). MS (ESI) $(M + H)^+ = 383.6$; HPLC analysis: > 99.9% purity.

*4.3.10. Compound **23***

(R)-1-(5-chloro-3H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl) pyrrolidin-3-amine was synthesized using *tert*-butyl (R)-pyrrolidin-3-ylcarbamate. ^{1}H NMR (500 MHz, MeOD) δ 7.43–7.37 (m, 4H), 6.96 (d, J = 8.0 Hz, 1H), 3.86 (s, 2H), 3.74–3.70 (m, 2H), 3.54–3.48 (m, 2H), 3.48–3.43 (m, 1H), 2.29 (s, 1H), 2.00–2.17 (m, 1H). MS (ESI) $(M + H)^+ = 397.5$. HPLC analysis: 95.2% purity.

4.3.11. Compound **24**

(S)-1-(5-chloro-3H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl) pyrrolidin-3-amine was synthesized using *tert*-butyl (S)-pyrrolidin-3-ylcarbamate. ^{1}H NMR (500 MHz, MeOD) δ 7.43–7.37 (m, 4H), 6.95 (d, J = 8.0 Hz, 1H), 3.85 (s, 2H), 3.70–3.64 (m, 2H), 3.52–3.48 (m, 2H), 3.48–3.41 (m, 1H), 2.30 (s, 1H), 2.00–1.97 (m, 1H). MS (ESI) $(M + H)^+ = 397.5$. HPLC analysis: 97.3% purity.

4.3.12. Compound **25**

1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N-(3,5-dichlorobenzyl)azepan-3-amine was synthesized using *tert*-butyl azepan-3-ylcarbamate. ^{1}H NMR (500 MHz, MeOD) δ 7.43–7.40 (m, 2H), 7.439 (s, 1H), 7.34–7.31 (m, 1H), 6.95–6.91(d, J = 8.0, 1H) 4.0–3.9 (m, 1H), 3.88 (s, 2H) 3.76–3.66 (m, 1H), 3.60–3.49 (m, 2H), 3.00–2.89 (br, 1H), 1.99–1.73 (m, 4H), 1.48–1.41 (m, 2H). MS (ESI) $(M + H)^+ = 425.9$. HPLC analysis: 96.0% purity.

4.3.13. Compound **31**

N1-(5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)-N2-(3,5-dichlorobenzyl)-N1-methylethane-1,2-diamine was synthesized using *tert*-butyl (2-(methylamino)ethyl)carbamate. ^{1}H NMR (500 MHz, MeOD) δ 7.45 (d, J = 7.7 Hz, 1H), 7.33–7.28 (m, 3H), 6.98 (d, J = 8.0 Hz, 1H), 3.93 (s, 2H), 3.71–3.67 (m, 2H), 3.19 (s, 3H), 3.06–3.00 (m, 2H). MS (ESI) $(M + H)^+ = 385.6$. HPLC analysis: 98.3% purity.

4.4. Procedures for the synthesis of compound **27** (Scheme 3)

(a) DMF (1.0 mL) solution of aldehyde (1.0 mmol) was added dropwise to a mixture of 6-chloropyridine-2, 3-diamine (1.0 mmol), and sodium bisulfite (1.0 mmol) in DMF (1.0 mL) was added over a 10 min period at 100 °C. After 2 h, the reaction mixture was concentrated, the residue was purified by flash column chromatography (DCM/MeOH) to give *tert*-butyl (3-(5-chloro-3H-imidazo [4,5-b] pyridin-2-yl) propyl) carbamate **38.**

(b,c) Following the general procedures in Scheme 2 to remove Boc with TFA, followed with the reductive amination reaction with 3, 5-dichlorobenzaldehyde yields the target compound **27**: *3-(5-chloro-3H-imidazo [4,5-b] pyridin-2-yl)-N-(3,5-dichlorobenzyl)propan-1-amine.* ^{1}H NMR (500 MHz, MeOD) δ 7.90 (d, J = 8.3 Hz, 1H), 7.21–7.26 (m, 3H), 7.28 (d, J = 8.2 Hz, 1H), 3.90 (s, 2H), 3.04 (t, J = 7.2 Hz, 2H), 2.83 (t, J = 6.6 Hz, 2H), 2.16–2.10 (m, 2H). MS (ESI) $(M + H)^+ = 370.7$. HPLC analysis: 97.8% purity.

4.5. Procedures for the synthesis of compound **28** (Scheme 4)

(a) Compound 5-chloro-1H-imidazo[4,5-b]pyridine-2(3H)-thione **39** (1 eq), 2-(2-bromoethyl) isoindoline-1,3-dione (1 eq) were dissolved in CH_3CN, K_2CO_3 (4 eq) was added and the resulting mixture was stirred at room temperature overnight. The mixture was extracted with DCM, washed with brine, concentrated under vacuum and purified via flash column chromatography (10% MeOH in DCM) to yield the intermediate **40:** 2-(2-((5-chloro-1H-imidazo [4,5-b]pyridin-2-yl)thio)ethyl)isoindoline-1,3-dione. MS (ESI) $[M+H]^+ = 361.3$

(b) 2-(2-((5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)thio)ethyl) isoindoline-1,3-dione was dissolved in MeOH, hydrazine (2 eq) was added, the mixture was then microwave irradiated at 70 °C for 20 min. The solvent was evaporated off and residue was purified via flash column chromatography directly eluting with 20% MeOH in DCM to yield intermediate **41:** 2-((5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)thio)ethanamine.MS (ESI) $[M+H]^+ = 229.7$

(c) 2-((5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)thio)ethanamine was dissolved in MeOH, 3,5-dichlorobenzaldehyde (1.1 equiv) was added, then AcOH was added (2 eq). After the mixture was stirred for 30 min, $NaBH_3CN$ (2 equiv) was added. The mixture was stirred overnight and the solvent was removed under vacuum. The residue was dissolved in DCM and washed with sat. sodium bicarbonate and brine. The organic layer was dried over Na_2SO_4 and concentrated under vacuum. The residue was purified by flash column chromatography (DCM/MeOH), yielding compound **28:** *2-((5-chloro-1H-imidazo[4,5-b]pyridin-2-yl)thio)-N-(3,5-dichlorobenzyl) ethanamine* as white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.47 (d, J = 8.0 Hz, 1H), 7.26 (s, 1H), 7.15 (s, 2H), 7.00 (d, J = 8.0 Hz, 1H), 4.82 (s, 2H), 3.75 (t, J = 6.9 Hz, 1H), 2.81 (t, J = 6.9 Hz, 1H). MS (ESI) $[M+H]^+ = 388.8$. HPLC analysis: 97.8% purity.

4.6. General procedures for the synthesis of compounds **29** and **30** (Scheme 5)

(a) To the solution of *tert*-butyl 3,5-dichlorobenzyl-2-isothiocyanatoethylcarbamate **42** (1 eq) in anhydrous acetonitrile was added the aromatic diamine **33** (1 eq). The mixture was refluxed for 20 h. After the solvent was removed, the residue was recrystallized in EtOAc/hexane twice to give the thiourea.

(b) To the solution of the thiourea (1 eq) in anhydrous acetonitrile (warm to dissolve), DIC (2 eq) was added. The solution was refluxed for 10 h. After cooling down, the white precipitate was collected by filtration to give the carbamate.

(c) To the solution of carbamate (1 eq) obtained above in dioxane, excess 4 N HCl in dioxane was added. The precipitate was formed in 5 min. The mixture was continued to stir overnight. The solid was collected and washed with dioxane to obtain the white solid as target compound.

4.6.1. Compound **29**

N1-(6-chloro-1H-benzo[d]imidazol-2-yl)-N2-(3,5-dichlorobenzyl) ethane-1,2-diamine in HCl form was synthesized using 4-chlorobenzene-1,2-diamine. ^{1}H NMR (500 MHz, MeOD) δ 7.46–7.40 (m, 3H), 7.18 (s, 1H), 7.13 (d, J = 8.3 Hz, 1H), 6.98 (d, J = 8.3 Hz, 1H), 4.01 (s, 2H), 3.62–3.56 (m, 2H), 3.10–3.04 (m, 2H). MS (ESI) $(M + H)^+ = 370.6$. HPLC analysis: 98.5% purity.

4.6.2. Compound **30**

N1-(5-chloro-3H-imidazo[4,5-b]pyridin-2-yl)-N2-(3,5-dichlorobenzyl) ethane-1,2-diamine was synthesized using 6-chloropyridine-2,3-diamine. ^{1}H NMR (500 MHz, MeOD) δ 7.43–7.39 (m, 3H), 7.30 (s, 1H), 6.97 (d, J = 8.1 Hz, 1H), 3.97 (s, 2H), 3.65–3.60 (m, 2H), 3.04–2.98 (m, 2H). MS (ESI) $(M + H)^+ = 371.7$. HPLC analysis: 98.6% purity.

4.7. Thermal shift assay

Thermal shift assays were performed in white 96-well plates (Bio-Rad) using the fluorescent dye SYPRO Orange (Invitrogen). 100 μl of 9.6 mg/mL protein was diluted in 1.2 mL assay buffer (100 mM HEPES, pH 7.5, 100 mM NaCl, 10 mM ATP-Mg) containing 2 μl of SYPRO Orange DMSO solution sold as "5000×". 1 μL of 20 mM stock solution of test compounds in DMSO (100 μM final

concentration) was pipetted into 9 μl of assay buffer and mixed with 10 μl of the above protein solution in each well. The 96-well plate was sealed with a transparent film (Bio-Rad) and centrifuged for 3 min at 800 *rpm* to ensure that samples were at the bottom of their wells and any bubble was removed. Temperature-dependent fluorescence increase reporting protein denaturation was measured in triplicate using a DNA Engine Opticon 2 instrument (Bio-Rad) from 25 to 90 °C in steps of 0.2 °C (excitation 470–505 nm; emission 540–700 nm). A T_m for each well was determined with the First Derivative method, which used Opticon Monitor software (Bio-Rad) to smooth the raw fluorescence data (converting each data point into a moving average of itself and adjacent points) and then finding the temperature at which the upward slope of the fluorescence versus temperature curve was the steepest - that is, the temperature at which the first derivative of this curve was maximal.

4.8. *T. brucei* growth inhibition assay

The assay is performed in the same procedure as described in a previous publications [8,14]: *T. brucei* (bloodstream form strain 427 from Dr. K. Stuart, Seattle Biomedical Research Institute, Seattle, WA) was cultured in HMI-9 medium containing 10% fetal bovine serum, penicillin, and streptomycin at 37 °C with 5% CO_2. Drug sensitivity of the *T. brucei* strain was determined in 96 well microtiter plates in triplicate with an initial inoculum of 1×10^4 trypomastigotes per well. Compound stock solutions were prepared in DMSO at 20 mM and added in serial dilutions for a final volume of 200 μl/well. Parasite growth was quantified at 48 h by the addition of Alamar Blue (ThermoFisher Scientific, Waltham, MA). Pentamidine isethionate (Aventis, Dagenham, UK) was included in each assay as a positive control.

4.9. Enzyme inhibition assay

Compounds were assayed against *T. brucei* methionyl-tRNA synthetase (*Tb*MetRS) in triplicate using an ATP-depletion chemiluminescence assay in a 96-well format as previously described [12]. Briefly, compounds were pre-incubated for 15 min at room temperature with 200 μg/mL bulk *E. coli* tRNA (Sigma-Aldrich or the equivalent activity of Roche), 50 nM *Tb*MetRS, 0.1 U/mL pyrophosphatase, 0.2 mM spermine, 0.1 mg/mL bovine serum albumin, 2.5 mM dithiothreitol, 25 mM HEPES-KOH pH 7.9, 10 mM $MgCl_2$, 50 mM KCl, and 2% DMSO. Then 100 nM ATP and 32 μM L-methionine was added to start the reaction. After a 120 min incubation, an equal volume (50 μl) of Kinase-Glo® (Promega) was added and the relative light units were quantified.

Percent inhibition = 100 × (test compound − AVG low control)/(AVG high control − AVG low control). Where low control is everything except compounds and high control is everything except compounds and L-methionine. IC_{50} values were calculated by non-linear regression, sigmoidal-dose response, in Prism 3.0.

4.10. Mammalian cell growth inhibition assays

The assay is performed in the same procedure as described in previous publication [9]: The human lymphocytic cell line CRL-1855 (American Type Culture Collection) were grown in RPMI medium (Lonza, Walkersville, MD) with 10% fetal bovine serum, penicillin and streptomycin at 37 °C with 5% CO_2. The human hepatocellular cell line HepG2 (American Type Culture Collection) were grown in DMEM/F-12 medium (Lonza, Walkersville, MD) with 10% fetal bovine serum, penicillin and streptomycin. Cells (5 × 103/well) were added to 96-well plates and incubated with serial dilutions of compounds in quadruplicate for 48 h. At that time, cell viability was quantified by addition of Alamar Blue, and plates were incubated for an additional 4 h at 5% CO_2, 37 °C. Absorbance readings (OD570–600) were used to calculate viability referenced against cells grown with no inhibitors.

4.11. PK studies in mice

The assay is performed in the same procedure as described in previous publications [14,16–18]: Compounds were administered by oral gavage at a concentration of 50 mg/kg in a volume of 0.2 mL of dosing solution (7% tween80, 3% ethanol, 5% DMSO, 0.9% saline). Tail blood was then collected at time points of 30, 60, 120, 240, 260 and 480 min into a heparinized capillary tube and 30 μl was then spotted onto a Whatman FTA DMPK-C Card (GE, Fairfield, CT.) Blood samples were analyzed by extracting the dried blood spots [19] in acetonitrile for measurements of compound concentrations by liquid chromatography/tandem mass spectrometry.

4.12. Distribution of compounds between mouse plasma and brain

This experiment is adapted from a previous reported procedure [18]: Mice (n = 3) were injected IP with 5 mg/kg of the test compound dissolved in 400 μl of dosing solution (7% tween 80, 3% EtOH, 5% DMSO, 0.9% saline) and sacrificed at 60 min for collection of plasma and brains. The brains were removed, weighed, and frozen to be later homogenized in acetonitrile using a Dounce homogenizer. Blood was briefly centrifuged to separate out the plasma for determining plasma concentrations of the compound. Concentrations of compound in the plasma and brain tissue was determined by liquid chromatography/tandem mass spectrometry analysis relative to a standard compound amount. The concentration of compound in the brain was obtained by dividing the moles of compound in the brain by the brain volume (obtained from the brain weight assuming 1 g is 1 mL) and correcting for the brain vasculature volume of 3% by weight.

All animal study protocols were approved by the Institutional Animal Care and Use Committee of the University of Washington (IACUC no. 4248-01).

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R01AI084004 and R01AI097177. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Crystallography performed in support of the work benefited from remote access to resources at the Stanford Synchrotron Radiation Light source supported by the U.S. Department of Energy Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515 and by the National Institutes of Health (P41GM103393).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.ejmech.2016.10.024.

References

[1] P.J. Hotez, D.H. Molyneux, A. Fenwick, J. Kumaresan, S.E. Sachs, J.D. Sachs, L. Savioli, Control of neglected tropical diseases, N. Engl. J. Med. 357 (2007) 1018–1027.

[2] R. Brun, J. Blum, F. Chappuis, C. Burri, Human african trypanosomiasis, Lancet 375 (2010) 148–159.

[3] Mapping the risk of human African trypanosomiasis. http://www.who.int/trypanosomiasis_african/country/risk_AFRO/en/, 2016 (accessed 07.05.16).

[4] Trypanosomiasis, human African (sleeping sickness). http://www.who.int/mediacentre/factsheets/fs259/en/, 2016 (accessed 07.05.16).

[5] J. Rodgers, Human African trypanosomiasis, chemotherapy and CNS disease, J. Neuroimmunol. 211 (2009) 16–22.

[6] P. Mäser, S. Wittlin, M. Rottmann, T. Wenzler, M. Kaiser, R. Brun, Antiparasitic agents: new drugs on the horizon, Curr. Opin. Pharmacol. 12 (2012) 562–566.

[7] K. Sheppard, J. Yuan, M.J. Hohn, B. Jester, K.M. Devine, D. Söll, From one amino acid to another: tRNA-dependent amino acid biosynthesis, Nucleic Acids Res. 36 (2008) 1813–1825.

[8] S. Shibata, J.R. Gillespie, A.M. Kelley, A.J. Napuli, Z. Zhang, K.V. Kovzun, R.M. Pefley, J. Lam, F.H. Zucker, W.C. Van Voorhis, E.A. Merritt, W.G. Hol, C.L.M.J. Verlinde, E. Fan, F.S. Buckner, Selective inhibitors of methionyl-tRNA synthetase have potent activity against *Trypanosoma brucei* infection in mice, Antimicrob. Agents Chemother. 55 (2011) 1982–1989.

[9] S. Shibata, J.R. Gillespie, R.M. Ranade, C.Y. Koh, J.E. Kim, J.U. Laydbak, F.H. Zucker, W.G. Hol, C.L. Verlinde, F.S. Buckner, E. Fan, Urea-Based inhibitors of *Trypanosoma brucei* Methionyl-tRNA synthetase: selectivity and *in vivo* characterization, J. Med. Chem. 55 (2012) 6342–6351.

[10] R.L. Jarvest, J.M. Berge, V. Berry, H.F. Boyd, M.J. Brown, J.S. Elder, A.K. Forrest, A.P. Fosberry, D.R. Gentry, M.J. Hibbs, D.D. Jaworski, P.J. O'Hanlon, A.J. Pope, S. Rittenhouse, R.J. Sheppard, C. Slater-Radosti, A. Worby, Nanomolar inhibitors of *Staphylococcus aureus* methionyl tRNA synthetase with potent antibacterial activity against gram-positive pathogens, J. Med. Chem. 45 (2002) 1959–1962.

[11] R.L. Jarvest, S.A. Armstrong, J.M. Berge, P. Brown, J.S. Elder, M.J. Brown, R.C. Copley, A.K. Forrest, D.W. Hamprecht, P.J. O'Hanlon, D.J. Mitchell, S. Rittenhouse, D.R. Witty, Definition of the heterocyclic pharmacophore of bacterial methionyl tRNA synthetase inhibitors: potent antibacterially active non-quinolone analogues, Bioorg. Med. Chem. Lett. 14 (2004) 3937–3941.

[12] L. Pedró- Rosa, F.S. Buckner, R.M. Ranade, C. Eberhart, F. Madoux, J.R. Gillespie, C.Y. Koh, S. Brown, J. Lohse, C.L.M.J. Verlinde, E. Fan, T. Bannister, L. Scampavia, W.G.J. Hol, T. Spicer, P. Hodder, Identification of potent inhibitors of the *Trypanosoma brucei* methionyl-tRNA synthetase via high throughput orthogonal screening, J. Biomol. Screen. 20 (2015) 122–130.

[13] C.Y. Koh, J.E. Kim, S. Shibata, R.M. Ranade, M. Yu, J. Liu, J.R. Gillespie, F.S. Buckner, C.L.M.J. Verlinde, E. Fan, W.G. Hol, Distinct states of Methionyl-tRNA synthetase indicate inhibitor binding by conformational selection, Structure 20 (2012) 1681–1699.

[14] Z. Zhang, C.Y. Koh, R.M. Ranade, S. Shibata, J.R. Gillespie, M.A. Hulverson, W. Huang, J. Nguyen, N. Pendem, M.H. Gelb, C.L.M.J. Verlinde, W.G.J. Hol, F.S. Buckner, E. Fan, 5-Fluoroimidazo[4,5-b]pyridine as a privileged fragment that conveys bioavailability to potent Trypanosomal Methionyl-tRNA synthetase inhibitors, ACS Infect. Dis. 2 (2016) 399–404.

[15] R.M. Ranade, Z. Zhang, J.R. Gillespie, S. Shibata, C.L.M.J. Verlinde, W.G.J. Hol, E. Fan, F.S. Buckner, Inhibitors of methionyl-tRNA synthetase have potent activity against Giardia intestinalis trophozoites, Antimicrob. Agents Chemother. 59 (2015) 7128–7131.

[16] J.M. Kraus, C.L.M.J. Verlinde, M. Karimi, G.I. Lepesheva, M.H. Gelb, F.S. Buckner, Rational modification of a candidate cancer drug for use against Chagas disease, J. Med. Chem. 52 (2009) 1639–1647.

[17] P.K. Suryadevara, S. Olepu, J.W. Lockman, J. Ohkanda, M. Karimi, C.L.M.J. Verlinde, J.M. Kraus, J. Schoepe, W.C. Van Voorhis, A.D. Hamilton, F.S. Buckner, M.H. Gelb, Structurally simple inhibitors of Lanosterol 14α-demethylase are efficacious in a rodent model of acute Chagas disease, J. Med. Chem. 52 (2009) 3703–3715.

[18] H.B. Tatipaka, J.R. Gillespie, A.K. Chatterjee, N.R. Norcross, N.R.,M.A. Hulverson, R.M. Ranade, P. Nagendar, S.A. Creason, J. McQueen, N.A. Duster, A. Nagle, F. Supek, V. Molteni, T. Wenzler, R. Brun, R. Glynne, F.S. Buckner, M.H. Gelb, Substituted 2- Phenylimidazopyridines: a new class of drug leads for human African Trypanosomiasis, J. Med. Chem. 57 (2014) 828–835.

[19] M. Rowland, G.T. Emmons, Use of dried blood spots in drug development: pharmacokinetic considerations, AAPS J. 12 (2010) 290–293.

Bioorganic & Medicinal Chemistry Letters 27 (2017) 2702–2707

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Optimization of a binding fragment targeting the "enlarged methionine pocket" leads to potent *Trypanosoma brucei* methionyl-tRNA synthetase inhibitors

Wenlin Huang [a], Zhongsheng Zhang [a], Ranae M. Ranade [b], J. Robert Gillespie [b], Ximena Barros-Álvarez [a,c], Sharon A. Creason [b], Sayaka Shibata [a], Christophe L.M.J. Verlinde [a], Wim G.J. Hol [a], Frederick S. Buckner [b,*], Erkang Fan [a,*]

[a] Department of Biochemistry, University of Washington, Seattle, WA 98195, United States
[b] Department of Medicine, Division of Allergy and Infectious Diseases, and the Center for Emerging and Re-emerging Infectious Diseases (CERID), University of Washington, Seattle, WA 98109, United States
[c] Laboratorio de Enzimología de Parásitos, Facultad de Ciencias, Universidad de los Andes, Mérida, Venezuela

ARTICLE INFO

Article history:
Received 14 March 2017
Accepted 14 April 2017
Available online 17 April 2017

Keywords:
Human African trypanosomiasis
Methionyl-tRNA synthetase
Structure-based design
Trypanosoma brucei

ABSTRACT

Potent inhibitors of *Trypanosoma brucei* methionyl-tRNA synthetase were previously designed using a structure-guided approach. Compounds **1** and **2** were the most active compounds in the cyclic and linear linker series, respectively. To further improve cellular potency, SAR investigation of a binding fragment targeting the "enlarged methionine pocket" (EMP) was performed. The optimization led to the identification of a 6,8-dichloro-tetrahydroquinoline ring as a favorable fragment to bind the EMP. Replacement of 3,5-dichloro-benzyl group (the EMP binding fragment) of inhibitor **2** using this tetrahydroquinoline fragment resulted in compound **13**, that exhibited an EC_{50} of 4 nM.

Human African trypanosomiasis (HAT), commonly known as sleeping sickness, is a neglected tropical disease caused by the protozoan parasite *Trypanosoma brucei*.[1] The parasite is transmitted to humans through the bite of the tsetse fly. The disease progresses in two distinct stages: an initial acute stage where the parasitic infection is restricted to the hemolymphatic system and a late stage where the parasites cross the blood-brain barrier and reside in brain tissue.[2] Current treatment options are severely inadequate for this disease.[1,3] For the treatment of early stage infection, the two drugs, pentamidine and suramin, have toxicity and require injection.[4] The late stage infection is particularly difficult to treat, as drugs must cross the blood-brain barrier to be effective. The two drugs available for the late stage infection, melarsoprol and eflornithine, are toxic, have limited ability to cross the blood-brain barrier, and require injection.[4–6] New drugs that are safe and easy to administer are urgently needed for both stages of HAT.

We recently reported on structure-guided design of *Trypanosoma brucei* methionyl-tRNA synthetase (*Tb*MetRS) inhibitors.[7] Two series of compounds were designed and demonstrated to be potent *Tb*MetRS inhibitors. The most potent compound in each series is shown in Fig. 1. Compound **1** is a cyclic linker inhibitor with an EC_{50} of 39 nM and compound **2** is a linear linker inhibitor with an EC_{50} of 22 nM. In the previous study, the 3,5-dichlorophenyl moiety was fixed as the fragment to fill the so-called "enlarged methionine pocket" (EMP)[8] and the investigation was mainly focused on the linker part. Here, we report on the optimization of the EMP binding fragment based on the cyclic and linear linker compounds **1** and **2** in order to identify the preferred moiety for binding the EMP. This led to the identification of inhibitors with significantly enhanced potency.

Analogues of compound **1** in which the 3,5-dichlorophenyl moiety was replaced by various 3,5-disubstituted phenyl or 2,3,5-trisubstituted phenyl ring were prepared as shown in Scheme 1. The synthesis followed previously reported procedures.[7] In brief, (*S*)-*tert*-butyl piperidin-3-ylcarbamate reacted with 2-bromo-5-

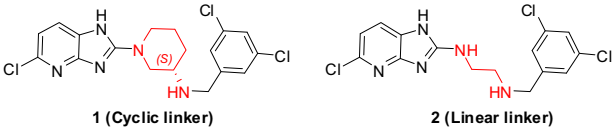
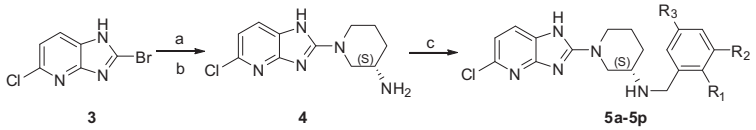


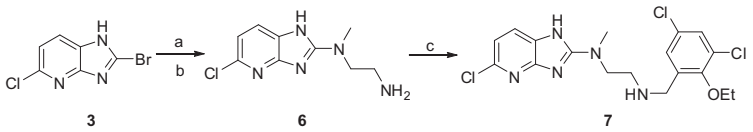
Fig. 1. Structures of compounds **1** and **2**.

* Corresponding authors.
E-mail addresses: fbuckner@uw.edu (F.S. Buckner), erkang@uw.edu (E. Fan).

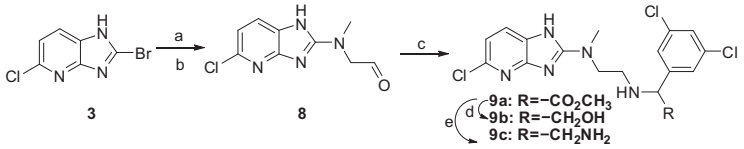
http://dx.doi.org/10.1016/j.bmcl.2017.04.048
0960-894X/



Scheme 1. Reagents and conditions: (a) (*S*)-*tert*-butyl piperidin-3-ylcarbamate, pyridine, MW, 100 °C, 30 min; (b) TFA, DCM, r.t. overnight; (c) Substituted benzaldehyde, DIPEA, $NaBH_3CN$, AcOH, CH_3OH, r.t., overnight.



Scheme 2. Reagents and conditions: (a) *tert*-butyl (2-(methylamino)ethyl) carbamate, pyridine, MW, 100 °C, 30 min; (b) TFA, DCM, r.t. overnight; (c) 3,5-dichloro-2-ethoxybenzaldehyde, DIPEA, $NaBH_3CN$, AcOH, CH_3OH, r.t., overnight.



Scheme 3. Reagents and conditions: (a) 2,2-dimethoxy-*N*-methylethanamine, pyridine, MW, 100 °C, 60 min; (b) HCl (2 M), acetone, reflux, 60 min; (c) methyl 2-amino-2-(3,5-dichlorophenyl)acetate (HCl salt), DIPEA, $NaBH_3CN$, AcOH, CH_3OH, r.t., overnight; (d) $LiAlH_4$, THF, 0 °C, 1 h; (e) $NH_3{\cdot}H_2O$, r.t. overnight.

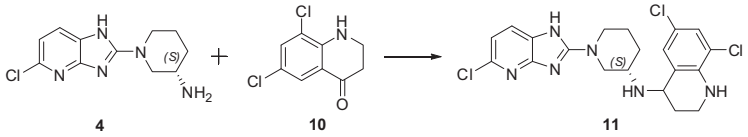
chloro-3*H*-imidazo[4,5-*b*]pyridine through nucleophilic substitution reaction, and the following Boc removal provided intermediate **4**. Reductive amination of **4** with various substituted benzaldehydes afforded the final products **5a**–**5p**.

An analogue of compound **2** in which the 3,5-dichlorophenyl moiety was replaced by 3,5-dichloro-2-ethoxy phenyl was also designed, and synthesized as shown in Scheme 2. Compound **7** was synthesized following the same procedure used for synthesizing compound **5**. Additional analogues of compound **2** were prepared through introducing substituents onto the benzylic α-position of the 3,5-dichlorophenyl ring (Scheme 3). Reagent 2,2-dimethoxy-*N*-methylethanamine reacted with 2-bromo-5-chloro-3H-imidazo [4,5-*b*]pyridine under the same microwave assisted nucleophilic substitution reaction, but with extended reaction time. The resulted intermediate containing an acetal group was hydrolyzed under acidic condition to produce the aldehyde intermediate **8**, which underwent reductive amination with methyl 2-amino-2-(3,5-dichlorophenyl)acetate to generate compound **9a**. The methyl ester group of **9a** was reduced by lithium aluminum hydride to generate compound **9b**, while ammonolysis of the ester group produced compound **9c**.

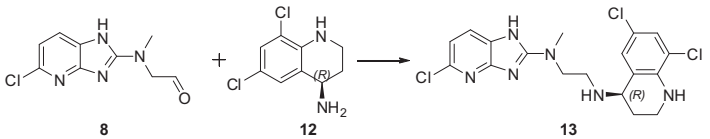
Inspired by the superior potency of bacterial MetRS and *Tb*MetRS inhibitors that contained a tetrahydroquinoline group reported previously,[9–11] a 6,8-dichloro-tetrahydroquinoline group was employed to replace the 3,5-dichlorophenyl moiety in compounds **1** and **2** to generate compounds **11** and **13**. Their synthesis is shown in Schemes 4 and 5. For compound **11**, amine and ketone were pre-reacted with $Ti(OEt)_4$ as catalyst, and the reductant $NaBH_3CN$ was added 30 min later followed by an 8 h reaction under microwave conditions (Scheme 4). Compound **13** was synthesized through reductive amination using intermediates **8** and **12** (Scheme 5). Intermediate **12** was prepared following previously published procedures.[9]

All the compounds were first evaluated for enzymatic potency against *Tb*MetRS using an ATP depletion assay as described previously.[7,12] As shown in Table 1, most of the compounds are very potent inhibitors of *Tb*MetRS, exhibiting IC_{50}s below 50 nM (the enzyme concentration used in the assay). Compounds **5k** and **5l** were found to have significantly reduced inhibitory potency compared to **1**, with IC_{50}s > 300 nM.

All the compounds were also tested for potency against *T. brucei* parasites using a growth inhibition assay as previously

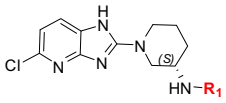


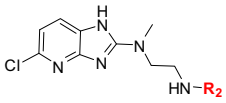
Scheme 4. Reagents and conditions: $Ti(OEt)_4$, EtOH, MW, 100 °C, 30 min; then $NaBH_3CN$, MW, 100 °C, 8 h.



Scheme 5. Reagents and conditions: DIPEA, $NaBH_3CN$, AcOH, CH_3OH, r.t., overnight.

Table 1
Inhibitory activities of compounds against *Tb*MetRS enzyme and *T. brucei* cell growth.





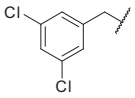
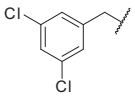
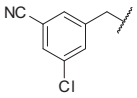
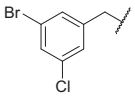
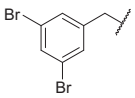
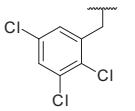
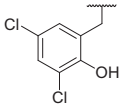
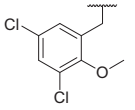
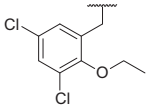
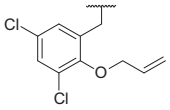
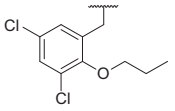
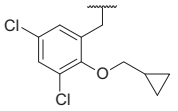
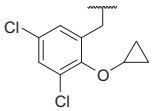
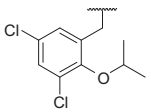
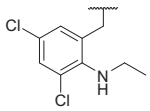
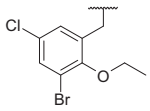
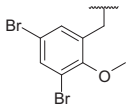
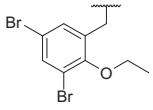
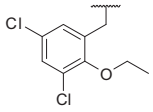
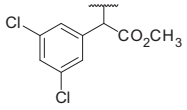
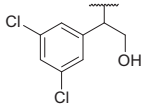
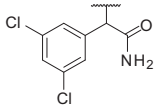
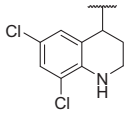
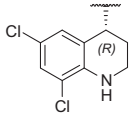
Compound number	R_1/R_2	Structure of R_1/R_2	IC_{50} (nM)[b]	EC_{50} (nM)[c]
1[a]	R_1		<50	39
2[a]	R_2		<50	22
5a	R_1		<50	111
5b	R_1		<50	71
5c	R_1		<50	85
5d	R_1		<50	354
5e	R_1		<50	241
5f	R_1		<50	357
5g	R_1		<50	111
5h	R_1		<50	326
5i	R_1		<50	827
5j	R_1		<50	511
5k	R_1		389	3115
5l	R_1		305	1703

Table 1 (*continued*)

Compound number	R_1/R_2	Structure of R_1/R_2	IC_{50} (nM)[b]	EC_{50} (nM)[c]
5m	R_1		<50	359
5n	R_1		<50	230
5o	R_1		<50	334
5p	R_1		<50	40
7	R_2		<50	314
9a	R_2		86	5273
9b	R_2		101	995
9c	R_2		75	4488
11	R_1		89	57
13	R_2		<50	4

[a] Data was published previously, included here for comparison.
[b] The values are averages of triplicate data, control for *Tb*MetRS IC_{50} assay: Met-SA[17] (±10.1%; n = 9 assays).
[c] The values are averages of triplicate data, control for *Tb*EC_{50} assay: Pentamidine (±15.5%; n = 9).

described.[7,9,13] The results are shown in Table 1 and the inhibition curves of compounds with EC_{50} < 100 nM are shown in Fig. S1 (Supporting information). Compounds with larger groups at the 3- and 5-positions, such as bromo and cyano (**5a**–**5c**), were tolerated by the enzyme, but had higher EC_{50} values by 2–3 fold. All compounds that contain the 2,3,5-trisubstituted phenyl ring except **5p** were less potent in the parasite growth inhibition assay. The ethoxy group was the best 2-position substituent as evidenced by **5g** and **5p** being more potent than their counterparts with other substituents. In fact the cellular potency of compound **5p** was nearly identical to the EC_{50} of **1** (39 nM). Larger alkoxy groups at the 2-position resulted in significant reductions in cellular activity, especially for branched alkoxy groups. In general, the larger the alkoxy group at the 2-position, the less cellular potency the compound exhibited (**5g** < **5h** < **5j** < **5i**). Compounds **5k** and **5l** that contain branched alkoxy groups at the 2-position lost cellular potency significantly. Compounds **9a**–**9c** that contain substitutions at the benzylic α-position in the case of the linear linker series exhibited weak cellular potency. The 6,8-dichloro-tetrahydroquinoline moiety was a good match for the linear linker series, but not in the cyclic linker series as indicated by compounds **11** and **13**. In the cyclic linker series, compound **11** was less potent than **1**, while in the linear linker series, compound **13** was more potent than **2**. Compound **13** was the most potent compound found in this study, exhibiting an EC_{50} of 4 nM. The calculated physicochemical properties of compounds with EC_{50} < 100 nM against *T. brucei* are presented in Table S1 (Supporting information). All these potent compounds possess ligand efficiency (LE) values >0.3, indicating a good balance of size and potency. It is noteworthy that compounds **2** and **13** show the highest LE of 0.43, and the ligand-lipophilicity efficiency (LLE) of 2.96 and 4.38, respectively. This indicates the optimization of **2** to **13** not only improved potency but also maintained LE and increased LLE.

To check for unexpected changes in the binding mode, we obtained the crystal structure of *Tb*MetRS in complex with compound **13** at 2.3 Å resolution (Fig. 2A). The binding mode of com-

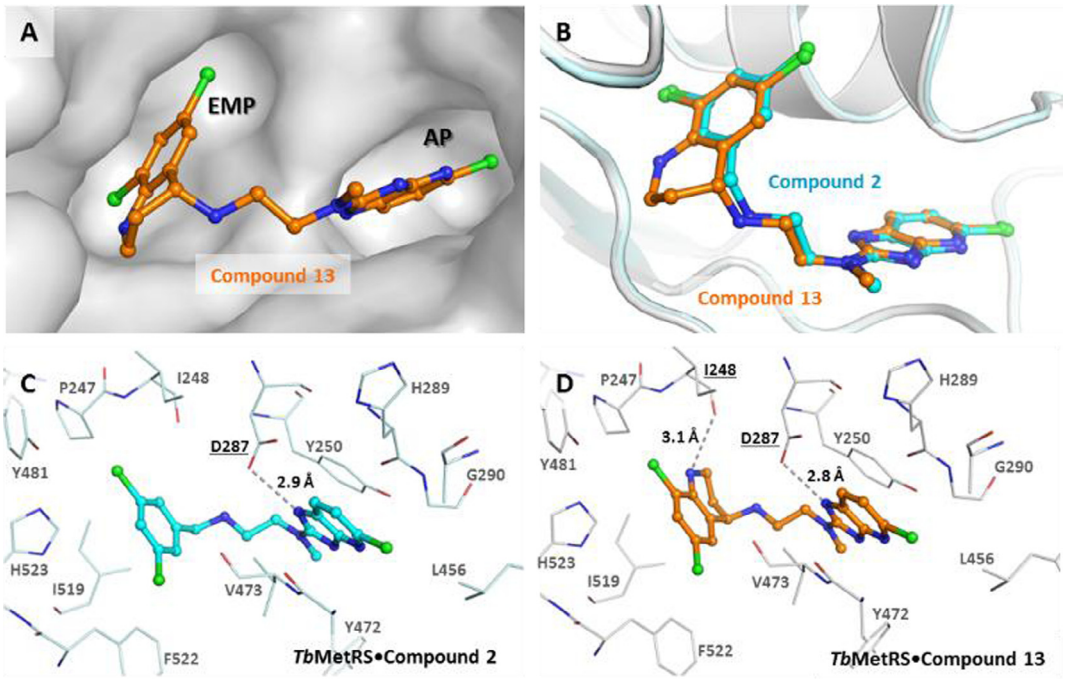


Fig. 2. Binding of compound **13** to *Tb*MetRS. (A) *Tb*MetRS·Compound **13** complex structure (PDB: 5V49). The protein surface and the two pockets, EMP and AP, where the inhibitor is bound are shown. Protein surface is colored grey, ligand carbons orange, nitrogens blue and chlorines green. (B) Superposition of *Tb*MetRS structures bound to compounds **13** and **2** (PDB: 5J59)[7]. Chlorine atoms are colored green, carbon atoms of compounds **13** and **2** are colored orange and cyan, respectively. (C and D) Hydrogen bond interactions in *Tb*MetRS·compound **2** and *Tb*MetRS·compound **13** structures are shown with dotted lines and labels for interacting residues are underlined.

Table 2
Toxicity, oral PK, and brain penetration data of select compounds.

Cpd.	CRL-8155 (μM)[b]	Hep G2 (μM)[c]	Oral PK (50 mg/kg)[d]		Brain/plasma (%)[d]
			AUC (min·μM)	C_{max} (μM)	
1[a]	29.3	49.0	6223 ± 2160	37.6 ± 22.1	0
2[a]	22.6	39.2	952 ± 331	9.7 ± 4.5	27.2 ± 7.1
5b	43.0	>50	4688 ± 2123	19.2 ± 16.4	0
5c	47.6	39.1	Not done	Not done	Not done
5p	10.5	18.7	182 ± 43	2.1 ± 1.5	4.4 ± 0.6
11	19.0	40.2	947 ± 354	14.7 ± 4.3	Not done
13	18.5	30.6	140 ± 40	1.3 ± 0.7	19.1 ± 7.5

[a] Data was published previously, included here for comparison.
[b] The values are averages of quadruplicate, control for CRL-8155 EC_{50} assay: Quinacrine (±12.5%; n = 2).
[c] The values are averages of quadruplicate, control for HepG2 EC_{50} assay: Quinacrine (±14.9%; n = 3).
[d] The values are from one experiment with 3 mice per group.

pound **13** was compared to compound **2** bound to *Tb*MetRS (Fig. 2B). The 5-chloro-imidazopyridine group of compound **13** bound in the same manner to the auxiliary pocket (AP) as the corresponding part in compound **2**, forming a hydrogen bond interaction with the catalytic residue Asp287 (Fig. 2C and D). The linkers of both compounds superimposed almost perfectly (Fig. 2B). The 6,8-dichloro-tetrahydroquinoline ring of compound **13** also bound similarly to the EMP compared to the 3,5-dichloro-benzyl group in compound **2**, as observed before for analogues with longer linkers.[9] The dichloro benzene group of compound **2** was in essentially the same position as the corresponding part in compound **13**. The availability of the compound **13** bound crystal structure could help to explain the structure-activity relationship (SAR) data generated in Table 1 and guide future inhibitor design.

Potent compounds with EC_{50} < 100 nM against *T. brucei* were selected to examine the host cell toxicity. The compounds shown in Table 2 were tested using a human lymphoblast cell line (CRL-8155) and a hepatocellular carcinoma cell line (Hep G2) following procedures described previously.[7,14] Compounds **5b** and **5c** had CC_{50}s (concentration to cause 50% cytotoxicity) close to or greater than 40 μM, whereas compounds **11** and **13** had CC_{50}s of ~20 μM, and compound **5p** had CC_{50} of ~10 μM. Overall, these compounds exhibited low toxicities to mammalian cell lines.

Compounds with good cellular potency were further tested for oral pharmacokinetic (PK) properties and/or brain penetration in mice (Table 2). The PK studies were performed following published procedures with compounds being administered by oral gavage at 50 mg/kg.[7,9,15–17] The brain permeability was tested at a dose of 5 mg/kg IP as described previously.[7,9] Compound **5b** exhibited high plasma exposure comparable to compound **1** with a C_{max} at 19.2 μM, and an AUC of 4688 min·μM. Unfortunately, like compound **1**, the brain permeability of **5b** was poor with undetectable brain levels at 60 min after IP injection. Compound **5p**, the 2,3,5-trisubstituted analogue of **1**, showed slightly improved brain/plasma ratio but less favorable oral PK compared to **1**. Compound **11** had promising PK properties with a C_{max} at 14.7 μM and AUC of 947 min·μM, whereas compound **13** showed low plasma exposure. Both the plasma and brain exposure of **13** are lower than those of **2**, but its improvement in EC_{50} over **2** may compensate for the lower exposure in an in vivo efficacy model.

In summary, a series of 3,5-disubstitued and 2,3,5-trisubstituted benzyl groups, as well as a 6,8-dichloro-tetrahydroquinoline moiety were used to replace the 3,5-dichlorobenzyl group as the EMP binding moiety based on previously discovered cyclic and linear linker *Tb*MetRS inhibitors. It was found that substituents larger than chloro at 3- and/or 5-positions do not improve potency while 2,3,5-trisubstituted benzyl groups generally resulted in decreased potency. Substituents at the benzylic α-position in the case of the linear linker series also led to a loss of potency. The 6,8-dichloro-tetrahydroquinoline moiety however, afforded compound **13** in the linear linker series with improved potency against *T. brucei* parasites. We obtained the crystal structure of *Tb*MetRS in complex with compound **13** which will help guide future inhibitor design. Compound **13** also had moderately good brain penetration in mice. This study identified potentially better fragments for binding the EMP than the 3,5-dichloro benzyl group in the cyclic and linear linker series of *Tb*MetRS inhibitors.

Acknowledgments

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under award numbers R01AI084004 and R01AI097177. We acknowledge the support of a Fulbright Fellowship to X.B.-A. We thank Stewart Turley and Robert Steinfeldt for providing support for the X-ray data collection and computing environment at the Biomolecular Structure Center of the University of Washington. Crystallography performed in support of the work benefitted from remote access to resources at the Stanford Synchrotron Radiation Lightsource supported by the U.S. Department of Energy Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515 and by the National Institutes of Health (P41GM103393).

A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmcl.2017.04.048.

References

1. Brun R, Blum J, Chappuis F, Burri C. Human African trypanosomiasis. *Lancet*. 2010;375:148.
2. Grab DJ, Kennedy PGE. Traversal of human and animal trypanosomes across the blood-brain barrier. *J Neurovirol*. 2008;14:344.
3. Rodgers J. Human African trypanosomiasis, chemotherapy and CNS disease. *J Neuroimmunol*. 2009;211:16.
4. Trypanosomiasis, human African (sleeping sickness). <http://www.who.int/mediacentre/factsheets/fs259/en/.; 2016 [Accessed 12.06.16].
5. Croft SL, Barrett MP, Urbina JA. Chemotherapy of trypanosomiases and leishmaniasis. *Trends Parasitol*. 2005;21:508.
6. Wilkinson SR, Kelly JM. Trypanocidal drugs: mechanisms, resistance and new targets. *Expert Rev Mol Med*. 2009;11.
7. Huang W, Zhang Z, Barros- Álvarez X, et al. Structure-guided design of novel *Trypanosoma brucei* methionyl-tRNA synthetase inhibitors. *Eur J Med Chem*. 2016;124:1081.
8. Koh CY, Kim JE, Shibata S, et al. Distinct states of methionyl-tRNA synthetase indicate inhibitor binding by conformational selection. *Structure*. 2012;20:1681.
9. Zhang Z, Koh CY, Ranade RM, et al. 5-Fluoroimidazo[4,5-*b*]pyridine Is a privileged fragment that conveys bioavailability to potent Trypanosomal methionyl-tRNA synthetase inhibitors. *ACS Infect Dis*. 2016;2:399.
10. Jarvest RL, Berge JM, Berry V, et al. Nanomolar inhibitors of *Staphylococcus aureus* methionyl tRNA synthetase with potent antibacterial activity against Gram-positive pathogens. *J Med Chem*. 2002;45:1959.
11. Jarvest RL, Armstrong SA, Berge JM, et al. Definition of the heterocyclic pharmacophore of bacterial methionyl tRNA synthetase inhibitors: potent antibacterially active non-quinolone analogues. *Bioorg Med Chem Lett*. 2004;14:3937.
12. Pedró-Rosa L, Buckner FS, Ranade RM, et al. Identification of potent inhibitors of the Trypanosoma brucei methionyl-tRNA synthetase via high throughput orthogonal screening. *J Biomol Screening*. 2015;20:122.
13. Shibata S, Gillespie JR, Kelley AM, et al. Selective inhibitors of methionyl-tRNA synthetase have potent activity against Trypanosoma brucei infection in mice. *Antimicrob Agents Chemother*. 2011;55:1982.
14. Shibata S, Gillespie JR, Ranade RM, et al. Urea-based inhibitors of *Trypanosoma brucei* methionyl-tRNA synthetase: selectivity and in vivo characterization. *J Med Chem*. 2012;55:6342.
15. Kraus JM, Verlinde CLMJ, Karimi M, Lepesheva GI, Gelb MH, Buckner FS. Rational modification of a candidate cancer drug for use against Chagas disease. *J Med Chem*. 2009;52:1639–1647.
16. Suryadevara PK, Olepu S, Lockman JW, et al. Structurally simple inhibitors of Lanosterol 14α-demethylase are efficacious in a rodent model of acute Chagas disease. *J Med Chem*. 2009;52:3703.
17. Tatipaka HB, Gillespie JR, Chatterjee AK, et al. Substituted 2-phenylimidazopyridines: a new class of drug leads for human African Trypanosomiasis. *J Med Chem*. 2014;57:828.

Chapter 5

Concluding remarks

The work done as part of this dissertation has provided a structural platform for future and ongoing research in the development of drugs against serious human infectious diseases. The tRNA synthetase structures solved and described in this thesis fed into different stages of the structure guided drug design process (SGDD; outlined in Figure 6 on page 21).

In an early-on SGDD stage, the description and analysis of the *M. tuberculosis* MetRS crystal structure in complex with Met-AMP [106] (Chapter 2) revealed differences with other MetRSs including the human enzymes, which could become useful in the development of drugs against tuberculosis. Due to the fact that an extra pocket in the active site region occurs in the human cytosolic enzyme compared to *M. tuberculosis* MetRS, and not the other way around, the design of inhibitors which have higher affinity for the *M. tuberculosis* enzyme than the human enzymes might be a challenge. Further studies, including HTS screening of *M. tuberculosis* MetRS followed by structure determinations of enzyme•hit complexes, may be required to arrive at leads which can be developed further.

The crystal structure of *L. donovani* TyrRS in complex with the tyrosyl adenylate analog TyrSA [107] (Chapter 3), which revealed the presence of an extra pocket shared with other important human pathogens and absent in both human TyrRS variants, could be key in the design of specific inhibitors in the treatment of visceral leishmaniasis and serious infectious diseases caused by other

pathogenic protozoa. This research is a case of both the early-on structure determination of a protein target and an analysis of a potentially druggable structural difference with the human counterparts that could drive the development of selective compounds against various parasites.

The value of collaborative work among various research groups, most of them within University of Washington, is reflected in the studies described in Chapter 4. These studies are examples of later stages in the SGDD process where detailed analysis of crystal structures contributed to the process of optimizing compounds against *Tbru*MetRS. A total of 57 crystal structures were determined upon soaking *Tbru*MetRS•Met crystals with multiple compounds derived from our collaborative drug design project. The corroboration of *Tbru*MetRS as the target for NEU-1053 that was identified from a compound library phenotypic screening against a different type of proteins and the description of the compound binding mode [108] shows the power of crystallography in the discovery of new binding modes for further optimization. The optimization and assistance in the design of new lead compounds [109, 110] against *Tbru*MetRS reflects the importance of structural information in improving not only the potency but other crucial pharmacological proprieties of inhibitors in the drug development process against diseases of affecting the most vulnerable *worldwide*.

Even though arriving at compounds which are endowed by the numerous pharmacological properties required to be useful as safe and effective drugs targeting tRNA synthetases is still somewhat in the future for the trypanosomatid diseases described in this thesis, the development of resistance is an important point to be taken into account. It is of interest that the collaborative effort of the UW groups targeting tRNA synthetases has resulted in structural information from a number of different tRNA synthetases from trypanosomatids, including TyrRS from *Leishmania major* [105] and *L. donovani* (Chapter 3), HisRS from *T. cruzi* [156], and MetRS from *T. brucei* ([90, 91],

Chapter 4). With homology modeling approaches where needed, models can be obtained from HisRS, TyrRS and MetRS enzymes from the same species. This would provide a basis for arriving at multiple compounds inhibiting different essential enzymes in the same pathogenic trypanosomatid. A multi-compound approach in treating patients would decrease considerably the probability of resistance arising. It is hoped that the investigations described in this thesis may contribute to such a vision.

References

[1] WHO. (2018) Infectious diseases. http://www.who.int/topics/infectious_diseases/en/.

[2] Alsan, M. M., Westerhaus, M., Herce, M., Nakashima, K., and Farmer, P. E. (2011) Poverty, Global Health, and Infectious Disease: Lessons from Haiti and Rwanda, *Infectious Disease Clinics of North America 25*, 611-622.

[3] WHO. (2017) Global Tuberculosis Report 2017. http://www.who.int/tb/publications/global_report/en/.

[4] CDC. (2011) TB elimination. https://www.cdc.gov/tb/publications/factsheets/general/tb.pdf.

[5] WHO, and Vaccines, S. W. G. o. B. (2017) Report on BCG vaccine use for protection against mycobacterial infections including tuberculosis, leprosy, and other nontuberculous mycobacteria (NTM) infections. http://www.who.int/immunization/sage/meetings/2017/october/1_BCG_report_revised_version_online.pdf.

[6] Koul, A., Arnoult, E., Lounis, N., Guillemont, J., and Andries, K. (2011) The challenge of new drug discovery for tuberculosis, *Nature 469*, 483-490.

[7] Brennan, P. J. (2003) Structure, function, and biogenesis of the cell wall of Mycobacterium tuberculosis, *Tuberculosis 83*, 91-97.

[8] Korf, J., Stoltz, A., Verschoor, J., De Baetselier, P., and Grooten, J. (2005) The Mycobacterium tuberculosis cell wall component mycolic acid elicits pathogen-associated host innate immune responses, *European Journal of Immunology 35*, 890-900.

[9] Davis, J. M., and Ramakrishnan, L. (2009) The Role of the Granuloma in Expansion and Dissemination of Early Tuberculous Infection, *Cell 136*, 37-49.

[10] Schaberg, T., Rebhan, K., and Lode, H. (1996) Risk factors for side-effects of isoniazid, rifampin and pyrazinamide in patients hospitalized for pulmonary tuberculosis, *Eur. Resp. J. 9*, 2026-2030.

[11] Hannaert, V., Bringaud, F., Opperdoes, F. R., and Michels, P. A. (2003) Evolution of energy metabolism and its compartmentation in Kinetoplastida, *Kinetoplastid biology and disease 2*, 11.

[12] Fernandes, A. P., Nelson, K., and Beverley, S. M. (1993) Evolution of Nuclear Ribosomal RNAs in Kinetoplastic Protozoa: Perspectives on the Age and Origins of Parasitism, *Proceedings of the National Academy of Sciences of the United States of America 90*, 11608-11612.

[13] Shaw, J. M., Feagin, J. E., Stuart, K., and Simpson, L. (1988) Editing of kinetoplastid mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG initiation codons, *Cell 53*, 401-411.

[14] Opperdoes, F. R., Baudhuin, P., Coppens, I., De Roe, C., Edwards, S. W., Weijers, P. J., and Misset, O. (1984) Purification, morphometric analysis, and characterization of the glycosomes (microbodies) of the protozoan hemoflagellate Trypanosoma brucei, *The Journal of Cell Biology 98*, 1178-1184.

[15] Cuervo, P., Domont, G. B., and De Jesus, J. B. (2010) Proteomics of trypanosomatids of human medical importance, *Journal of Proteomics 73*, 845-867.

[16] Hol, W. G. J. (2015) Three-dimensional structures in the design of therapeutics targeting parasitic protozoa: reflections on the past, present and future, *Acta Crystallographica Section F 71*, 485-499.

[17] Horn, D. (2014) Antigenic variation in African trypanosomes, *Mol. Biochem. Parasitol. 195*, 123-129.
[18] Kipnis, T. L., and Dasilva, W. D. (1989) EVASION OF TRYPANOSOMA-CRUZI FROM COMPLEMENT LYSIS, *Brazilian J. Med. Biol. Res. 22*, 1-16.
[19] Girones, N., Cuervo, H., and Fresno, M. (2005) Trypanosoma cruzi-induced molecular mimicry and Chagas' disease, *Curr.Top.Microbiol.Immunol. 296*, 89-123.
[20] Ramirez-Toloza, G., and Ferreira, A. (2017) Trypanosoma cruzi Evades the Complement System as an Efficient Strategy to Survive in the Mammalian Host: The Specific Roles of Host/Parasite Molecules and Trypanosoma cruzi Calreticulin, *Front. Microbiol. 8*, 13.
[21] Fairlamb, A. H., Gow, N. A., Matthews, K. R., and Waters, A. P. (2016) Drug resistance in eukaryotic microorganisms, *Nat Microbiol 1*, 16092.
[22] initiative, D. f. N. D. (2016) DNDi 2016 Leishmaniasis Fact Sheet.
[23] WHO. (2017) Leishmaniasis Fact Sheet, April 2017. http://www.who.int/mediacentre/factsheets/fs375/en/.
[24] WHO. (2010) Control of the Leishmaniasis. http://apps.who.int/iris/bitstream/10665/44412/1/WHO_TRS_949_eng.pdf In *Report of a meeting of the WHO Expert Committee on the Control of Leishmaniases, 22 -- 26 March 2010, Geneva*.
[25] Murray, H. W., Berman, J. D., Davies, C. R., and Saravia, N. G. (2005) Advances in leishmaniasis, *The Lancet 366*, 1561-1577.
[26] Sundar, S., and Chakravarty, J. (2013) Leishmaniasis: an update of current pharmacotherapy, *Expert Opin Pharmacother 14*, 53-63.
[27] Moore, E., and Lockwood, D. (2010) Treatment of visceral leishmaniasis, *Journal of Global Infectious Diseases 2*, 151-158.
[28] WHO. (2017) Human African trypanosomiasis (sleeping sickness) Fact Sheet, January 2017. http://www.who.int/mediacentre/factsheets/fs259/en/.
[29] initiative, D. f. N. D. (2016) DNDi 2016 Human African trypanosomiasis Fact Sheet.
[30] Simarro, P. P., Cecchi, G., Franco, J. R., Paone, M., Diarra, A., Ruiz-Postigo, J. A., Fèvre, E. M., Mattioli, R. C., and Jannin, J. G. (2012) Estimating and Mapping the Population at Risk of Sleeping Sickness (The Risk of Sleeping Sickness), *6*, e1859.
[31] initiative, D. f. N. D. (2017) DNDi 2017 Chagas disease.
[32] WHO. (2017) Chagas Disease (American trypanosomiasis) Fact Sheet, March 2017. http://www.who.int/mediacentre/factsheets/fs340/en/.
[33] Tarleton, R. L., Reithinger, R., Urbina, J. A., Kitron, U., and Gürtler, R. E. (2007) The Challenges of Chagas Disease— Grim Outlook or Glimmer of Hope? (Neglected Diseases), *PLoS Medicine 4*, e332.
[34] Brener, Z., and Gazzinelli, R. T. (1997) Immunological control of Trypanosoma cruzi infection and pathogenesis of Chagas' disease, *Int. Arch. Allergy Immunol. 114*, 103-110.
[35] Barrett, M. P., Burchmore, R. J. S., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., and Krishna, S. (2003) The trypanosomiases, *Lancet 362*, 1469-1480.
[36] Tanowitz, H. B., Kirchhoff, L. V., Simon, D., Morris, S. A., Weiss, L. M., and Wittner, M. (1992) CHAGAS-DISEASE, *Clinical Microbiology Reviews 5*, 400-419.
[37] Kohanski, M. A., Dwyer, D. J., and Collins, J. J. (2010) How antibiotics kill bacteria: from targets to networks, *Nat. Rev. Microbiol. 8*, 423-435.
[38] Wilson, D. N. (2014) Ribosome-targeting antibiotics and mechanisms of bacterial resistance, *Nat. Rev. Microbiol. 12*, 35-48.

[39] Pham, J. S., Dawson, K. L., Jackson, K. E., Lim, E. E., Pasaje, C. F., Turner, K. E., and Ralph, S. A. (2014) Aminoacyl-tRNA synthetases as drug targets in eukaryotic parasites, *Int J Parasitol Drugs Drug Resist 4*, 1-13.
[40] Kim, S., Lee, S. W., Choi, E. C., and Choi, S. Y. (2003) Aminoacyl-tRNA synthetases and their inhibitors as a novel family of antibiotics, *Applied Microbiology and Biotechnology 61*, 278-288.
[41] Ibba, M., and Soll, D. (2000) Aminoacyl-tRNA synthesis, *Annual Review of Biochemistry 69*, 617-650.
[42] Francklyn, C. S., First, E. A., Perona, J. J., and Hou, Y.-M. (2008) Methods for kinetic and thermodynamic analysis of aminoacyl-tRNA synthetases, *Methods 44*, 100-118.
[43] Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D. (1991) Class II Aminoacyl Transfer RNA Synthetases: Crystal Structure of Yeast Aspartyl-tRNA Synthetase Complexed with tRNA^{Asp}, *Science 252*, 1682-1689.
[44] Cusack, S. (1997) Aminoacyl-tRNA synthetases, *Current Opinion in Structural Biology 7*, 881-889.
[45] Ochsner, U. A., Sun, X., Jarvis, T., Critchley, I., and Janjic, N. (2007) Aminoacyl-tRNA synthetases: essential and still promising targets for new anti-infective agents, *Expert Opin Investig Drugs 16*, 573-593.
[46] Tolkunova, E., Park, H., Xia, J., King, M. P., and Davidson, E. (2000) The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript, *Journal of Biological Chemistry 275*, 35063-35069.
[47] Mudge, S. J., Williams, J. H., Eyre, H. J., Sutherland, G. R., Cowan, P. J., and Power, D. A. (1998) Complex organisation of the 5 '-end of the human glycine tRNA synthetase gene, *Gene 209*, 45-50.
[48] Nagao, A., Suzuki, T., Katoh, T., Sakaguchi, Y., and Suzuki, T. (2009) Biogenesis of glutaminyl-mt tRNA(Gln) in human mitochondria, *Proceedings of the National Academy of Sciences of the United States of America 106*, 16209-16214.
[49] Alfonzo, J. D., Blanc, V., Estévez, A. M., Rubio, M. A. T., and Simpson, L. (1999) C to U editing of the anticodon of imported mitochondrial tRNA Trp allows decoding of the UGA stop codon in Leishmania tarentolae, *EMBO Journal 18*, 7056-7062.
[50] Charriere, F., O'Donoghue, P., Helgadottir, S., Marechal-Drouard, L., Cristodero, M., Horn, E. K., Soll, D., and Schneider, A. (2009) Dual targeting of a tRNAAsp requires two different aspartyl-tRNA synthetases in Trypanosoma brucei, *J Biol Chem 284*, 16210-16217.
[51] Tan, T. H. P., Bochud-Allemann, N., Horn, E. K., and Schneider, A. (2002) Eukaryotic-Type Elongator tRNA Met of Trypanosoma brucei Becomes Formylated after Import into Mitochondria, *Proceedings of the National Academy of Sciences of the United States of America 99*, 1152-1157.
[52] Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2012) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Advanced Drug Delivery Reviews 64*, 4-17.
[53] Tao, J., and Schimmel, P. (2000) Inhibitors of aminoacyl-tRNA synthetases as novel anti-infectives, pp 1767-1775, Taylor & Francis.

[54] Stefanska, A. L., Fulston, M., Houge-Frydrych, C. S., Jones, J. J., and Warr, S. R. (2000) A potent seryl tRNA synthetase inhibitor SB-217452 isolated from a Streptomyces species, *The Journal of antibiotics 53*, 1346.
[55] Vondenhoff, G. H. M., and Van Aerschot, A. (2011) Aminoacyl-tRNA synthetase inhibitors as potential antibiotics, *Eur. J. Med. Chem. 46*, 5227.
[56] Fang, P., Yu, X., Jeong, S. J., Mirando, A., Chen, K., Chen, X., Kim, S., Francklyn, C. S., and Guo, M. (2015) Structural basis for full-spectrum inhibition of translational functions on a tRNA synthetase, *Nat Commun 6*, 6402.
[57] Tan, M., Wang, M., Zhou, X.-L., Yan, W., Eriani, G., and Wang, E.-D. (2013) The Yin and Yang of tRNA: proper binding of acceptor end determines the catalytic balance of editing and aminoacylation, *Nucleic acids research 41*, 5513.
[58] Rock, F. L., Mao, W. M., Yaremchuk, A., Tukalo, M., Crepin, T., Zhou, H. C., Zhang, Y. K., Hernandez, V., Akama, T., Baker, S. J., Plattner, J. J., Shapiro, L., Martinis, S. A., Benkovic, S. J., Cusack, S., and Alley, M. R. K. (2007) An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site, *Science 316*, 1759-1761.
[59] Silvian, L., Wang, J., and Steitz, T. (1999) Insights into editing from an Ile-tRNA synthetase structure with tNA(Ile) and mupirocin, *Science 285*.
[60] Nakama, T., Nureki, O., and Yokoyama, S. (2001) Structural basis for the recognition of isoleucyl-adenylate and an antibiotic, mupirocin, by isoleucyl-tRNA synthetase, *J Biol Chem 276*, 47387-47393.
[61] Khan, S., Sharma, A., Belrhali, H., Yogavel, M., and Sharma, A. (2014) Structural basis of malaria parasite lysyl-tRNA synthetase inhibition by cladosporin, *J Struct Funct Genomics 15*, 63-71.
[62] Zhou, H., Sun, L., Yang, X. L., and Schimmel, P. (2013) ATP-directed capture of bioactive herbal-based medicine on human tRNA synthetase, *Nature 494*, 121-124.
[63] Brick, P., Bhat, T. N., and Blow, D. M. (1989) Structure of tyrosyl-tRNA synthetase refined at 2.3 Å resolution: Interaction of the enzyme with the tyrosyl adenylate intermediate, *J. Mol. Biol. 208*, 83-98.
[64] Fidock, D. A., Nomura, T., Talley, A. K., Cooper, R. A., Dzekunov, S. M., Ferdig, M. T., Ursos, L. M. B., Bir Singh Sidhu, A., Naudé, B., Deitsch, K. W., Su, X.-Z., Wootton, J. C., Roepe, P. D., and Wellems, T. E. (2000) Mutations in the P. falciparum Digestive Vacuole Transmembrane Protein PfCRT and Evidence for Their Role in Chloroquine Resistance, *Molecular Cell 6*, 861-871.
[65] Thomas, D. G., Wilson, J. M., Day, M. J., and Russell, A. D. (1999) Mupirocin resistance in staphylococci: development and transfer of isoleucyl-tRNA synthetase-mediated resistance in vitro, *Journal of Applied Microbiology 86*, 715-722.
[66] O'Dwyer, K., Spivak, A. T., Ingraham, K., Min, S., Holmes, D. J., Jakielaszek, C., Rittenhouse, S., Kwan, A. L., Livi, G. P., Sathe, G., Thomas, E., Van Horn, S., Miller, L. A., Twynholm, M., Tomayko, J., Dalessandro, M., Caltabiano, M., Scangarella-Oman, N. E., and Brown, J. R. (2015) Bacterial resistance to leucyl-tRNA synthetase inhibitor GSK2251052 develops during treatment of complicated urinary tract infections, *Antimicrobial agents and chemotherapy 59*, 289.
[67] Kuntz, I. D. (1992) Structure-Based Strategies for Drug Design and Discovery, *Science 257*, 1078-1082.

[68] Tari, L. (2012) *Chapter 1: The Utility of Structural Biology in Drug Discovery*, Vol. 841, Humana Press, Totowa, NJ, Totowa, NJ.
[69] Blundell, T. L. (1996) Structure-based drug design, *Nature 384*, 23-26.
[70] D. Andricopulo, A. (2014) Structure-Based Drug Design to Overcome Drug Resistance: Challenges and Opportunities, *Current Pharmaceutical Design 20*, 687-693.
[71] initiative, D. f. N. D. (2018) Target Product Profile – Sleeping Sickness.
[72] Kim, C. U., Lew, W., Williams, M. A., Liu, H., Zhang, L., Swaminathan, S., Bischofberger, N., Chen, M. S., Mendel, D. B., Tai, C. Y., Laver, W. G., and Stevens, R. C. (1997) Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity, *Journal of the American Chemical Society 119*, 681.
[73] Wlodawer, A., and Erickson, J. W. (1993) Structure-based inhibitors of HIV-1 protease, *Annual review of biochemistry 62*, 543.
[74] King, N. M., Prabu-Jeyabalan, M., Nalivaika, E. A., Wigerinck, P., de Bethune, M.-P., and Schiffer, C. A. (2004) Structural and Thermodynamic Basis for the Binding of TMC114, a Next-Generation Human Immunodeficiency Virus Type 1 Protease Inhibitor, *The Journal of Virology 78*, 12012.
[75] Coteron, J. M., Marco, M., Esquivias, J., Deng, X., White, K. L., White, J., Koltun, M., El Mazouni, F., Kokkonda, S., Katneni, K., Bhamidipati, R., Shackleford, D. M., Angulo-Barturen, I., Ferrer, S. B., Jimenez-Diaz, M. B., Gamo, F.-J., Goldsmith, E. J., Charman, W. N., Bathurst, I., Floyd, D., Matthews, D., Burrows, J. N., Rathod, P. K., Charman, S. A., and Phillips, M. A. (2012) Structure-Guided Lead Optimization of Triazolopyrimidine-Ring Substituents Identifies Potent Plasmodium falciparum Dihydroorotate Dehydrogenase Inhibitors with Clinical Candidate Potential, *J. Med. Chem. 54*.
[76] Phillips, M. A., Lotharius, J., Marsh, K., White, J., Dayan, A., White, K. L., Njoroge, J. W., El Mazouni, F., Lao, Y., Kokkonda, S., Tomchick, D. R., Deng, X., Laird, T., Bhatia, S. N., March, S., Ng, C. L., Fidock, D. A., Wittlin, S., Lafuente-Monasterio, M., Benito, F. J., Alonso, L. M. S., Martinez, M. S., Jimenez-Diaz, M. B., Bazaga, S. F., Angulo-Barturen, I., Haselden, J. N., Louttit, J., Cui, Y., Sridhar, A., Zeeman, A.-M., Kocken, C., Sauerwein, R., Dechering, K., Avery, V. M., Duffy, S., Delves, M., Sinden, R., Ruecker, A., Wickham, K. S., Rochford, R., Gahagen, J., Iyer, L., Riccio, E., Mirsalis, J., Bathhurst, I., Rueckle, T., Ding, X., Campo, B., Leroy, D., Rogers, M. J., Rathod, P. K., Burrows, J. N., and Charman, S. A. (2015) A long-duration dihydroorotate dehydrogenase inhibitor (DSM265) for prevention and treatment of malaria, *Sci. Transl. Med. 7*.
[77] Yuthavong, Y., Tarnchompoo, B., Vilaivan, T., Chitnumsub, P., Kamchonwongpaisan, S., Charman, S. A., McLennan, D. N., White, K. L., Vivas, L., Bongard, E., Thongphanchang, C., Taweechai, S., Vanichtanankul, J., Rattanajak, R., Arwon, U., Fantauzzi, P., Yuvaniyama, J., Charman, W. N., and Matthews, D. (2012) Malarial dihydrofolate reductase as a paradigm for drug development against a resistance-compromised target, *Proceedings of the National Academy of Sciences of the United States of America 109*, 16823.
[78] Wilson, D. N. (2014) Ribosome-targeting antibiotics and mechanisms of bacterial resistance, *Nat Rev Microbiol 12*, 35-48.

[79] Jensen, M., and Mehlhorn, H. (2009) Seventy-five years of Resochin ® in the fight against malaria, *Founded as Zeitschrift für Parasitenkunde 105*, 609-627.
[80] Tu, Y. (2011) The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine, *Nature Medicine 17*, 1217-1220.
[81] Su, X.-Z., and Miller, L. (2015) The discovery of artemisinin and the Nobel Prize in Physiology or Medicine, *Science China Life Sciences 58*, 1175-1179.
[82] Deng, X., Gujjar, R., El Mazouni, F., Kaminsky, W., Malmquist, N. A., Goldsmith, E. J., Rathod, P. K., and Phillips, M. A. (2010) Structural Plasticity of Malaria Dihydroorotate Dehydrogenase Allows Selective Binding of Diverse Chemical Scaffolds, *J. Biol. Chem. 284*.
[83] Verlinde, C. L., and Hol, W. G. (1994) Structure-based drug design: progress, results and challenges, *Structure 2*, 577-587.
[84] Jianping, D., Kalyan, D., Henri, M., Luc, K., Koen, A., Paul, A. J. J., Stepen, H. H., and Edward, A. (1995) Structure of HIV-1 RT/TIBO R 86183 complex reveals similarity in the binding of diverse nonnucleoside inhibitors, *Nature Structural Biology 2*, 407.
[85] Critchley, I. A., Green, L. S., Young, C. L., Bullard, J. M., Evans, R. J., Price, M., Jarvis, T. C., Guiles, J. W., Janjic, N., and Ochsner, U. A. (2009) Spectrum of activity and mode of action of REP3123, a new antibiotic to treat Clostridium difficile infections, *Journal of Antimicrobial Chemotherapy 63*, 954-963.
[86] Critchley, I. A., Young, C. L., Stone, K. C., Ochsner, U. A., Guiles, J., Tarasow, T., and Janjic, N. (2005) Antibacterial activity of REP8839, a new antibiotic for topical use, *Antimicrobial agents and chemotherapy 49*, 4247.
[87] Green, L. S., Bullard, J. M., Ribble, W., Dean, F., Ayers, D. F., Ochsner, U. A., Janjic, N., and Jarvis, T. C. (2009) Inhibition of Methionyl-tRNA Synthetase by REP8839 and Effects of Resistance Mutations on Enzyme Activity, *Antimicrobial Agents and Chemotherapy 53*, 86-94.
[88] Shibata, S., Gillespie, J. R., Kelley, A. M., Napuli, A. J., Zhang, Z., Kovzun, K. V., Pefley, R. M., Lam, J., Zucker, F. H., Van Voorhis, W. C., Merritt, E. A., Hol, W. G., Verlinde, C. L., Fan, E., and Buckner, F. S. (2011) Selective inhibitors of methionyl-tRNA synthetase have potent activity against Trypanosoma brucei Infection in Mice, *Antimicrob Agents Chemother 55*, 1982-1989.
[89] Larson, E. T., Kim, J. E., Zucker, F. H., Kelley, A., Mueller, N., Napuli, A. J., Verlinde, C., Fan, E. K., Buckner, F. S., Van Voorhis, W. C., Merritt, E. A., and Hol, W. G. J. (2011) Structure of Leishmania major methionyl-tRNA synthetase in complex with intermediate products methionyladenylate and pyrophosphate, *Biochimie 93*, 570-582.
[90] Koh, C. Y., Kim, J. E., Shibata, S., Ranade, R. M., Yu, M. Y., Liu, J. Y., Gillespie, J. R., Buckner, F. S., Verlinde, C., Fan, E., and Hol, W. G. J. (2012) Distinct States of Methionyl-tRNA Synthetase Indicate Inhibitor Binding by Conformational Selection, *Structure 20*, 1681-1691.
[91] Koh, C. Y., Kim, J. E., Wetzel, A. B., de van der Schueren, W. J., Shibata, S., Ranade, R. M., Liu, J., Zhang, Z., Gillespie, J. R., Buckner, F. S., Verlinde, C. L., Fan, E., and Hol, W. G. (2014) Structures of Trypanosoma brucei methionyl-tRNA synthetase with urea-based inhibitors provide guidance for drug design against sleeping sickness, *PLoS Negl Trop Dis 8*, e2775.
[92] Shibata, S., Gillespie, J. R., Ranade, R. M., Koh, C. Y., Kim, J. E., Laydbak, J. U., Zucker, F. H., Hol, W. G., Verlinde, C. L., Buckner, F. S., and Fan, E. (2012) Urea-based

inhibitors of Trypanosoma brucei methionyl-tRNA synthetase: selectivity and in vivo characterization, *J Med Chem 55*, 6342-6351.
[93] Zhang, Z., Koh, C. Y., Ranade, R. M., Shibata, S., Gillespie, J. R., HuIverson, M. A., Huang, W., Nguyen, J., Pendem, N., Gelb, M. H., Verlinde, C., Hol, W. G. J., Buckner, F. S., and Fan, E. (2016) 5-Fluoroimidazo 4,5-b pyridine Is a Privileged Fragment That Conveys Bioavailability to Potent Trypanosomal Methionyl-tRNA Synthetase Inhibitors, *ACS Infect. Dis. 2*, 399-404.
[94] Ranade, R., Gillespie, J., Shibata, S., Verlinde, C., Fan, E., Hol, W., and Buckner, F. (2013) Induced Resistance to Methionyl-tRNA Synthetase Inhibitors in Trypanosoma brucei Is Due to Overexpression of the Target, *Antimicrobial Agents and Chemotherapy 57*, 3021.
[95] Faghih, O., Zhang, Z. S., Ranade, R. M., Gillespie, J. R., Creason, S. A., Huang, W. L., Shibata, S., Barros-Alvarez, X., Verlinde, C., Hol, W. G. J., Fan, E. K., and Buckner, F. S. (2017) Development of Methionyl-tRNA Synthetase Inhibitors as Antibiotics for Gram-Positive Bacterial Infections, *Antimicrobial Agents and Chemotherapy 61*, 18.
[96] Greenwood, R. C., and Gentry, D. R. (2002) Confirmation of the antibacterial mode of action of SB-219383, a novel tyrosyl tRNA synthetase inhibitor from a Micromonospora sp, *Journal of Antibiotics 55*, 423-426.
[97] Qiu, X., Janson, C. A., Smith, W. W., Green, S. M., McDevitt, P., Johanson, K., Carter, P., Hibbs, M., Lewis, C., Chalker, A., Fosberry, A., Lalonde, J., Berge, J., Brown, P., Houge-Frydrych, C. S., and Jarvest, R. L. (2001) Crystal structure of Staphylococcus aureus tyrosyl-tRNA synthetase in complex with a class of potent and specific inhibitors, *Protein Science 10*, 9.
[98] Stefanska, A. L., Coates, N. J., Mensah, L. M., Pope, A. J., Ready, S. J., and Warr, S. R. (2000) SB-219383, a novel tyrosyl tRNA synthetase inhibitor from a Micromonospora sp I. Fermentation, isolation and properties, *Journal of Antibiotics 53*, 345-350.
[99] Houge-Frydrych, C. S., Readshaw, S. A., and Bell, D. J. (2000) SB-219383, a novel tyrosyl tRNA synthetase inhibitor from a Micromonospora sp. II. Structure determination, *The Journal of antibiotics 53*, 351.
[100] Brown, P., Eggleston, D. S., Haltiwanger, R. C., Jarvest, R. L., Mensah, L., O'Hanlon, P. J., and Pope, A. J. (2001) Synthetic analogues of SB-219383. Novel C-glycosyl peptides as inhibitors of tyrosyl tRNA synthetase, *Bioorganic & Medicinal Chemistry Letters 11*, 711-714.
[101] Chen, L. W., Wang, P. F., Tang, D. J., Tao, X. X., Man, R. J., Qiu, H. Y., Wang, Z. C., Xu, C., and Zhu, H. L. (2016) Metronidazole containing pyrazole derivatives potently inhibit tyrosyl-tRNA synthetase: design, synthesis, and biological evaluation, *Chem Biol Drug Des 88*, 592-598.
[102] Jarvest, R. L., Berge, J. M., Houge-Frydrych, C. S. V., Janson, C., Mensah, L. M., O'Hanlon, P. J., Pope, A., Saldanha, A., and Qiu, X. Y. (1999) Interaction of tyrosyl aryl dipeptides with S-aureus tyrosyl tRNA synthetase: Inhibition and crystal structure of a complex, *Bioorganic & Medicinal Chemistry Letters 9*, 2859-2862.
[103] Wang, X. D., Deng, R. C., Dong, J. J., Peng, Z. Y., Gao, X. M., Li, S. T., Lin, W. Q., Lu, C. L., Xiao, Z. P., and Zhu, H. L. (2013) 3-Aryl-4-acyloxyethoxyfuran-2(5H)-ones as inhibitors of tyrosyl-tRNA synthetase: Synthesis, molecular docking and antibacterial evaluation, *Bioorganic & Medicinal Chemistry 21*, 4914-4922.
[104] Bhatt, T. K., Khan, S., Dwivedi, V. P., Banday, M. M., Sharma, A., Chandele, A., Camacho, N., Ribas de Pouplana, L., Wu, Y., Craig, A. G., Mikkonen, A. T., Maier, A.

G., Yogavel, M., and Sharma, A. (2011) Malaria parasite tyrosyl-tRNA synthetase secretion triggers pro-inflammatory responses, *Nat Commun 2*, 530.
[105] Larson, E. T., Kim, J. E., Castaneda, L. J., Napuli, A. J., Zhang, Z., Fan, E., Zucker, F. H., Verlinde, C. L., Buckner, F. S., Van Voorhis, W. C., Hol, W. G., and Merritt, E. A. (2011) The double-length tyrosyl-tRNA synthetase from the eukaryote Leishmania major forms an intrinsically asymmetric pseudo-dimer, *J Mol Biol 409*, 159-176.
[106] Barros-Alvarez, X., Turley, S., Ranade, R. M., Gillespie, J. R., Duster, N. A., Verlinde, C., Fan, E., Buckner, F. S., and Hol, W. G. J. (2018) The crystal structure of the drug target Mycobacterium tuberculosis methionyl-tRNA synthetase in complex with a catalytic intermediate, *Acta crystallographica. Section F, Structural biology communications 74*, 245-254.
[107] Barros-Álvarez, X., Kerchner, K. M., Koh, C. Y., Turley, S., Pardon, E., Steyaert, J., Ranade, R. M., Gillespie, J. R., Zhang, Z., Verlinde, C. L. M. J., Fan, E., Buckner, F. S., and Hol, W. G. J. (2017) Leishmania donovani tyrosyl-tRNA synthetase structure in complex with a tyrosyl adenylate analog and comparisons with human and protozoan counterparts, *Biochimie 138*, 124-136.
[108] Devine, W. G., Diaz-Gonzalez, R., Ceballos-Perez, G., Rojas, D., Satoh, T., Tear, W., Ranade, R. M., Barros-Álvarez, X., Hol, W. G. J., Buckner, F. S., Navarro, M., and Pollastri, M. P. (2017) From Cells to Mice to Target: Characterization of NEU-1053 (SB-443342) and Its Analogues for Treatment of Human African Trypanosomiasis, *ACS Infect. Dis. 3*, 225-236.
[109] Huang, W., Zhang, Z., Barros-Alvarez, X., Koh, C. Y., Ranade, R. M., Gillespie, J. R., Creason, S. A., Shibata, S., Verlinde, C., Hol, W. G. J., Buckner, F. S., and Fan, E. (2016) Structure-guided design of novel Trypanosoma brucei Methionyl-tRNA synthetase inhibitors, *Eur J Med Chem 124*, 1081-1092.
[110] Huang, W. L., Zhang, Z. S., Ranade, R. M., Gillespie, J. R., Barros-Alvarez, X., Creason, S. A., Shibata, S., Verlinde, C., Hol, W. G. J., Buckner, F. S., and Fan, E. (2017) Optimization of a binding fragment targeting the "enlarged methionine pocket" leads to potent Trypanosoma brucei methionyl-tRNA synthetase inhibitors, *Bioorganic & Medicinal Chemistry Letters 27*, 2702-2707.
[111] Sheppard, K., Akochy, P. M., and Soll, D. (2008) Assays for transfer RNA-dependent amino acid biosynthesis, *Methods 44*, 139-145.
[112] Schimmel, P., Tao, J. S., and Hill, J. (1998) Aminoacyl tRNA synthetases as targets for new anti-infectives, *Faseb Journal 12*, 1599-1609.
[113] Sassanfar, M., Kranz, J. E., Gallant, P., Schimmel, P., and Shiba, K. (1996) A eubacterial Mycobacterium tuberculosis tRNA synthetase is eukaryote-like and resistant to a eubacterial-specific antisynthetase drug, *Biochemistry 35*, 9995-10003.
[114] Deniziak, M. A., and Barciszewski, J. (2001) Methionyl-tRNA synthetase, *Acta Biochim. Pol. 48*, 337-350.
[115] Gentry, D. R., Ingraham, K. A., Stanhope, M. J., Rittenhouse, S., Jarvest, R. L., O'Hanlon, P. J., Brown, J. R., and Holmes, D. J. (2003) Variable sensitivity to bacterial methionyl-tRNA synthetase inhibitors reveals subpopulations of Streptococcus pneumoniae with two distinct methionyl-tRNA synthetase genes, *Antimicrobial Agents and Chemotherapy 47*, 1784-1789.

[116] Brown, J. R., Gentry, D., Becker, J. A., Ingraham, K., Holmes, D. J., and Stanhope, M. J. (2003) Horizontal transfer of drug-resistant aminoacyl-transfer-RNA synthetases of anthrax and Gram-positive pathogens, *EMBO Rep. 4*, 692-698.
[117] Pedro-Rosa, L., Buckner, F. S., Ranade, R. M., Eberhart, C., Madoux, F., Gillespie, J. R., Koh, C. Y., Brown, S., Lohse, J., Verlinde, C. L., Fan, E., Bannister, T., Scampavia, L., Hol, W. G., Spicer, T., and Hodder, P. (2015) Identification of potent inhibitors of the Trypanosoma brucei methionyl-tRNA synthetase via high-throughput orthogonal screening, *J Biomol Screen 20*, 122-130.
[118] Murillo, A. C., Li, H. Y., Alber, T., Baker, E. N., Berger, J. M., Cherney, L. T., Cherney, M. M., Yoon Song, C., Eisenberg, D., Garen, C. R., Goulding, C. W., Hung, L. W., Ioerger, T. R., Jacobs, W. R., M.N.G.James, Kim, C., Krieger, I., Lott, J. S., Sankaranarayanan, R., Segelke, B. W., Terwilliger, T. C., F.Wang, Wang, S., and Sacchettini, J. C. (2007) High Throughput Crystallography of TB Drug Targets, *Infectious Disorders - Drug Targets 7*, 127-139.
[119] Fang, Z., van der Merwe, R. G., Warren, R. M., Schubert, W. D., and van Pittius, N. C. G. (2015) Assessing the progress of Mycobacterium tuberculosis H37Rv structural genomics, *Tuberculosis 95*, 131-136.
[120] Terwilliger, T. C., Park, M. S., Waldo, G. S., Berendzen, J., Hung, L. W., Kim, C. Y., Smith, C. V., Sacchettini, J. C., Bellinzoni, M., Bossi, R., De Rossi, E., Mattevi, A., Milano, A., Riccardi, G., Rizzi, M., Roberts, M. M., Coker, A. R., Fossati, G., Mascagni, P., Coates, A. R. M., Wood, S. P., Goulding, C. W., Apostol, M. I., Anderson, D. H., Gill, H. S., Eisenberg, D. S., Taneja, B., Mande, S., Pohl, E., Lamzin, V., Tucker, P., Wilmanns, M., Colovos, C., Meyer-Klaucke, W., Munro, A. W., McLean, K. J., Marshall, K. R., Leys, D., Yang, J. K., Yoon, H. J., Lee, B. I., Lee, M. G., Kwak, J. E., Han, B. W., Lee, J. Y., Baek, S. H., Suh, S. W., Komen, M. M., Arcus, V. L., Baker, E. N., Lott, J. S., Jacobs, W., Alber, T., and Rupp, B. (2003) The TB structural genomics consortium: a resource for Mycobacterium tuberculosis biology, *Tuberculosis 83*, 223-249.
[121] Ingvarsson, H., and Unge, T. (2010) Flexibility and communication within the structure of the Mycobacterium smegmatis methionyl-tRNA synthetase, *FEBS J 277*, 3947-3962.
[122] Nakanishi, K., Ogiso, Y., Nakama, T., Fukai, S., and Nureki, O. (2005) Structural basis for anticodon recognition by methionyl-tRNA synthetase, *Nat. Struct. Mol. Biol. 12*, 931-932.
[123] Alexandrov, A., Vignali, M., LaCount, D. J., Quartley, E., de Vries, C., De Rosa, D., Babulski, J., Mitchell, S. F., Schoenfeld, L. W., Fields, S., Hol, W. G., Dumont, M. E., Phizicky, E. M., and Grayhack, E. J. (2004) A facile method for high-throughput co-expression of protein pairs, *Mol. Cell. Proteomics 3*, 934-938.
[124] Choi, R., Kelley, A., Leibly, D., Hewitt, S. N., Napuli, A., and Van Voorhis, W. (2011) Immobilized metal-affinity chromatography protein-recovery screening is predictive of crystallographic structure success, *Acta Crystallogr. F-Struct. Biol. Cryst. Commun. 67*, 998-1005.
[125] Mehlin, C., Boni, E., Buckner, F. S., Engel, L., Feist, T., Gelb, M. H., Haji, L., Kim, D., Liu, C., Mueller, N., Myler, P. J., Reddy, J. T., Sampson, J. N., Subramanian, E., Van Voorhis, W. C., Worthey, E., Zucker, F., and Hol, W. G. J. (2006) Heterologous expression of proteins from Plasmodium falciparum: Results from 1000 genes, *Mol. Biochem. Parasitol. 148*, 144-160.

[126] Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry 7*, 7.
[127] Teng, T. Y. (1990) MOUNTING OF CRYSTALS FOR MACROMOLECULAR CRYSTALLOGRAPHY IN A FREESTANDING THIN-FILM, *J. Appl. Crystallogr. 23*, 387-391.
[128] Haas, D. J., and Rossmann, M. G. (1970) CRYSTALLOGRAPHIC STUDIES ON LACTATE DEHYDROGENASE AT-75 DEGREES C, *Acta Crystallographica Section B-Structural Crystallography and Crystal Chemistry B 26*, 998-&.
[129] Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, In *Macromolecular Crystallography, Part A* (Carter, C., and Sweet, R., Eds.), pp 307–326, Academic Press.
[130] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.
[131] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.
[132] Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr D Biol Crystallogr 53*, 16.
[133] Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr 66*, 12-21.
[134] Schrödinger, L. L. C. The PyMOL Molecular Graphics System, Version 1.7. .
[135] Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state, *J Mol Biol 372*, 774-797.
[136] Ojo, K. K., Ranade, R. M., Zhang, Z. S., Dranow, D. M., Myers, J. B., Choi, R., Hewitt, S. N., Edwards, T. E., Davies, D. R., Lorimer, D., Boyle, S. M., Barrett, L. K., Buckner, F. S., Fan, E. K., and Van Voorhis, W. C. (2016) Brucella melitensis Methionyl-tRNA-Synthetase (MetRS), a Potential Drug Target for Brucellosis (vol 11, e0160350, 2016), *Plos One 11*, 3.
[137] Serre, L., Verdon, G., Choinowski, T., Hervouet, N., Risler, J. L., and Zelwer, C. (2001) How methionyl-tRNA synthetase creates its amino acid recognition pocket upon L-methionine binding, *J. Mol. Biol. 306*, 863-876.
[138] Crepin, T., Schmitt, E., Blanquet, S., and Mechulam, Y. (2004) Three-dimensional structure of methionyl-tRNA synthetase from Pyrococcus abyssi, *Biochemistry 43*, 2635-2644.
[139] Khan, S., Sharma, A., Belrhali, H., and Yogavel, M. (2014) Structural basis of malaria parasite lysyl-tRNA synthetase inhibition by cladosporin, *Journal of Structural and Functional Genomics 15*, 9.
[140] Khan, S. (2016) Recent advances in the biology and drug targeting of malaria parasite aminoacyl-tRNA synthetases, *Malar J 15*, 203.
[141] Anand, S., and Madhubala, R. (2016) Twin Attributes of Tyrosyl-tRNA Synthetase of Leishmania donovani: A HOUSEKEEPING PROTEIN TRANSLATION ENZYME AND A MIMIC OF HOST CHEMOKINE, *J Biol Chem 291*, 17754-17771.

[142] Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G., Wohlkonig, A., Ruf, A., Muyldermans, S., Hol, W. G., Kobilka, B. K., and Steyaert, J. (2014) A general protocol for the generation of Nanobodies for structural biology, *Nat Protoc 9*, 674-693.
[143] Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., and Sternberg, M. J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis, *Nat Protoc 10*, 845-858.
[144] Smart, O. S., Womack, T. O., Sharff, A., Flensburg, C., Keller, P., and al, e. (2011) Grade Web Server, version 1.001. .
[145] Kobayashi, T., Takimura, T., Sekine, R., Kelly, V. P., Kamata, K., Sakamoto, K., Nishimura, S., and Yokoyama, S. (2005) Structural snapshots of the KMSKS loop rearrangement for amino acid activation by bacterial tyrosyl-tRNA synthetase, *J Mol Biol 346*, 105-117.
[146] Tsunoda, M., Kusakabe, Y., Tanaka, N., Ohno, S., Nakamura, M., Senda, T., Moriguchi, T., Asai, N., Sekine, M., Yokogawa, T., Nishikawa, K., and Nakamura, K. T. (2007) Structural basis for recognition of cognate tRNA by tyrosyl-tRNA synthetase from three kingdoms, *Nucleic Acids Res 35*, 4289-4300.
[147] Fricker, P., Gastreich, M., and Rarey, M. (2004) Automated Generation of Structural Molecular Formulas under Constraints, *Journal of Chemical Information and Computer Sciences 44*, 14.
[148] Stierand, K., Maass, P. C., and Rarey, M. (2006) Molecular complexes at a glance: automated generation of two-dimensional complex diagrams, *Bioinformatics 22*, 1710-1716.
[149] Sajish, M., and Schimmel, P. (2015) A human tRNA synthetase is a potent PARP1-activating effector target for resveratrol, *Nature 519*, 370-+.
[150] Bonnefond, L., Frugier, M., Touze, E., Lorber, B., Florentz, C., Giege, R., Sauter, C., and Rudinger-Thirion, J. (2007) Crystal structure of human mitochondrial tyrosyl-tRNA synthetase reveals common and idiosyncratic features, *Structure 15*, 1505-1516.
[151] Hartmann, M. D., Shkolnaya, L. A., Bourenkov, G. P., Strizhov, N. I., and Bartunik, H. D. (2006) The Structure of Tyrosyl-tRNA Synthetase from Mycobacterium Tuberculosis, PDB ID: 2JAN.
[152] Kobayashi, T., Nureki, O., Ishitani, R., Yaremchuk, A., Tukalo, M., Cusack, S., Sakamoto, K., and Yokoyama, S. (2003) Structural basis for orthogonal tRNA specificities of tyrosyl-tRNA synthetases for genetic code expansion, *NATURE STRUCTURAL BIOLOGY 10*
[153] Diaz, R., Luengo-Arratta, S. A., Seixas, J. D., Amata, E., Devine, W., Cordon-Obras, C., Rojas-Barros, D. I., Jimenez, E., Ortega, F., Crouch, S., Colmenarejo, G., Fiandor, J. M., Martin, J. J., Berlanga, M., Gonzalez, S., Manzano, P., Navarro, M., and Pollastri, M. P. (2014) Identification and Characterization of Hundreds of Potent and Selective Inhibitors of Trypanosoma brucei Growth from a Kinase-Targeted Library Screening Campaign (Discovery of T. brucei Growth Inhibitors by Kinase-Targeted HTS), *8*, e3253.
[154] Jarvest, R. L., Armstrong, S. A., Berge, J. M., Brown, P., Elder, J. S., Brown, M. J., Copley, R. C. B., Forrest, A. K., Hamprecht, D. W., O'Hanlon, P. J., Mitchell, D. J., Rittenhouse, S., and Witty, D. R. (2004) Definition of the heterocyclic pharmacophore of bacterial methionyl tRNA synthetase inhibitors: potent antibacterially active non-quinolone analogues, *Bioorganic & Medicinal Chemistry Letters 14*, 3937-3941.
[155] Jarvest, R. L., Berge, J. M., Berry, V., Boyd, H. F., Brown, M. J., Elder, J. S., Forrest, A. K., Fosberry, A. P., Gentry, D. R., Hibbs, M. J., Jaworski, D. D., Hanlon, P. J., Pope, A.

J., Rittenhouse, S., Sheppard, R. J., Slater-Radosti, C., and Worby, A. (2002) Nanomolar inhibitors of Staphylococcus aureus methionyl tRNA synthetase with potent antibacterial activity against gram-positive pathogens, *Journal of medicinal chemistry 45*, 1959.

[156] Koh, C. Y., Siddaramaiah, L. K., Ranade, R. M., Nguyen, J., Jian, T. Y., Zhang, Z. S., Gillespie, J. R., Buckner, F. S., Verlinde, C., Fan, E. K., and Hol, W. G. J. (2015) A binding hotspot in Trypanosoma cruzi histidyl-tRNA synthetase revealed by fragment-based crystallographic cocktail screens, *Acta Crystallogr. Sect. D-Struct. Biol. 71*, 1684-1698.

Appendix A

The following table summarizes the result of the efforts done as part of this dissertation in obtaining *Tbru*MetRS•Inhibitor crystal structures that assisted in collaborative design and optimization projects of *Tbru*MetRS inhibitors. A total of 57 crystal structures obtained upon soaking *Tbru*MetRS•Met crystals with multiple inhibitors and with inhibitors and ATP were solved by the current writer and are listed below. If deposited, the PDB deposition codes are listed, as well as the corresponding reference if the structural information was included in a publication.

The list includes also the failed attempts where no density or very poor density for the inhibitor was found and no inhibitor could be modeled bound to *Tbru*MetRS. Essentially all compounds were synthesized by the group of Dr. Erkang Fan at the University of Washington, except for NEU-1053, which was kindly supplied by Dr. Michael Pollastri of Northeastern University.

#	Compound identifier	Density for compound		Comments
		YES	NO	
1	1753		X	
2	1754	X		
3	1762		X	
4	1766	X		
5	1710	X		
6	1788	X		
7	1789		X	
8	1801	X		
9	1805	X		
10	1809	X		
11	1811	X		

#	Compound identifier	Density for compound		Comments
		YES	NO	
12	1829	X		
13	1835	X		
14	1851		X	
15	1852		X	
16	1854	X		
17	1856	X		PDB ID: 5J58. Huang *et al*. (2016) [106]
18	1862	X		
19	1868		X	
20	1869	X		
21	1870	X		
22	1893	X		PDB ID: 5J59. Huang *et al*. (2016) [106]
23	1907	X		
24	1917	X		PDB ID: 5V49. Huang *et al*. (2017) [107]
25	1926	X		
26	1938	X		
27	1961	X		
28	1962	X		
29	1963	X		
30	1986	X		
31	1987	X		
32	1992	X		
33	2035	X		
34	2047	X		
35	2048	X		
36	2050	X		
37	2056	X		
38	2062	X		
39	2063	X		
40	2067	X		
41	2071		X	
42	2072		X	
43	2074		X	
44	2093	X		PDB ID: 6CML (coordinates not released to date)
45	2098	X		

#	Compound identifier	Density for compound		Comments
		YES	NO	
46	2102	X		
47	2105	X		
48	2106		X	
49	2107			
50	2108	X		
51	2119		X	
52	2128		X	
53	2129	X		
54	2198	X		
55	2201	X		
56	2204	X		
57	2205	X		
58	NEU-1053	X		PDB ID: 5TQU. Devine *et al*. (2017) [105]
59	70786556	X		PDB ID: 5J5A. Huang *et al*. (2016) [106] Soak, collection and initial refinement by Cho Yeow Koh. Final refinement, validation and structural analysis by the author of this dissertation.
60	1575 + ATP	X		
61	1614 + ATP	X		
62	1709 + ATP	X		
63	1710 + ATP	X		
64	1717 + ATP	X		
65	1856 + ATP	X		
66	1893 + ATP	X		
67	1899 + ATP	X		
68	1917 + ATP	X		
69	1938 + ATP	X		
70	NEU-1053 + ATP	X		