Targeting *Trypanosoma brucei* FPPS by Fragment-based drug discovery

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Pharmazie der

Philipps-Universität Marburg

vorgelegt von

Lena Münzker

aus **Mistelbach**

Marburg/Lahn 2019

Erstgutachter: **Prof. Dr. Gerhard Klebe**

Zweitgutachter: **Dr. Wolfgang Jahnke**

Eingereicht am **01.08.2019**

Tag der mündlichen Prüfung am **12.09.2019**

Hochschulkennziffer: 1180

ERKLÄRUNG

Ich versichere, dass ich meine Dissertation „Targeting *Trypanosoma brucei* FPPS by Fragment-based drug discovery" selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den.....

.....

(Unterschrift mit Vor- und Zuname)

Table of Contents

List of Figures

List of Tables

List of Schemes

Summary

Trypanosoma brucei (*T. brucei*) is the causative agent of the Human African Trypanosomiasis (HAT), which is a neglected disease with an endemic occurrence in 36 sub-Saharan African countries. The current standard of care suffers from low efficacy and severe side effects. Therefore, new drugs with better safety and efficacy profiles are urgently needed. Nitrogen-containing bisphosphonates, a current treatment for bone diseases, have been shown to block the growth of the *T. brucei* parasites by inhibiting farnesyl pyrophosphate synthase (FPPS); however, due to their particular pharmacokinetic properties they are not well suited for parasitic therapy. Recently, an additional allosteric site was discovered at the surface of human FPPS that, based on sequence analysis, is likely also present in *T. brucei* FPPS. The high unmet medical need combined with the discovery of a potential new target site prompted a fragment-based drug discovery approach to identify non-bisphosphonate binders on *T. brucei* FPPS, which is presented in this work.

Fragment screening was performed by NMR and X-ray crystallography. To this end, a robust *T. brucei* FPPS crystallization system was established enabling high-throughput determination of crystal structures up to 1.67 Å resolution. Structural superimpositions revealed that the allosteric site found on human FPPS is in fact present in *T. brucei* FPPS. This observation enabled subsequent protein-observed NMR and crystal soaking experiments with established human FPPS binders resulting in three protein-ligand complex structures with bound fragments in the previously unknown allosteric site. For most of the tested binders, K_d by SPR was outside of experimental range for *T. brucei* FPPS and only for one fragment the K_d on *T. brucei* FPPS was determined three orders of magnitude higher than the IC_{50} value on human FPPS. Crystal structural analysis revealed a different binding mode on human and *T. brucei* FPPS with reduced protein-ligand interactions on *T. brucei* FPPS, which explains the significantly reduced binding affinity.

Encouraged by the detection of first allosteric binders on *T. brucei* FPPS, fragment pools were screened by ligand-observed NMR and identified hits were followed-up by single compound ligand-observed NMR and protein-observed NMR resulting in 25 validated fragment hits for *T. brucei* FPPS. Validated hits were followed-up by crystal soaking and co-crystallization experiments and seven protein-ligand complex structures were solved using PanDDA. Out of the seven fragments, four fragments were bound in the active site, one fragment was detected in the allosteric site that was identified as part of this thesis, and two fragments were bound in surface-exposed binding sites. Notably, an active site bound fragment with a four atom long flexible linker adopted an orthogonal binding mode along α_D when compared to the other three ligands. Sixteen fragment analogues of the elongated flexible active site fragment were tested by SAR using additional test compounds retrieved from catalogue and archive, and one crystal structure with a fragment analogue was solved and was surprisingly found in the allosteric site.

In addition to the NMR fragment screen, an X-ray screen was performed at XChem (Diamond, UK) and at EMBL/ESRF (Grenoble, FR) resulting in seven protein-ligand structures. One fragment was positioned in the active site, three fragments in the allosteric site, two fragments in a cryptic site between helices α_I and α_H and one fragment at the opposite side of the allosteric site close to α_G and α_F. Fragment binding was further validated in protein-observed NMR.

As fragments identified by such screening approaches typically exhibit low binding affinities usually in µM to mM range, structure-based fragment optimisation based on a fragment merging and growing approach was performed. In total, ten compounds were synthesised and subjected to protein-observed NMR and X-ray structural analysis. Strikingly, a fragment merger based on *T. brucei* and *T. cruzi* active-site binders bound in a new binding site close to the SARM instead to the active site.

Taken together, this work presents high-resolution structures of *T. brucei* FPPS and identified 19 compounds binding to seven different sites thereby paving the way for future studies aiming to identify high-affinity non-bisphosphonate inhibitors for *T. brucei* FPPS with pharmacokinetic properties that are suitable for parasitic indications.

Zusammenfassung

Trypanosoma brucei (*T. brucei*) ist der Erreger der Afrikanischen Trypanosomiasis, einer vernachlässigten Krankheit mit einem endemischen Auftreten in 36 Subsahara-Ländern. Da derzeitige Behandlungmöglichkeiten eine geringe Wirksamkeit und starke Nebenwirkungen aufweisen, sind neue verbesserte Medikamente dringend erforderlich. Stickstoffhaltige Bisphosphonate, die derzeit bei Knochenerkrankungen eingesetzt werden, blockieren das Wachstum der *T. brucei* Parasiten durch Hemmung der Farnesylpyrophosphat-Synthase (FPPS), sind aber aufgrund ihrer speziellen pharmakokinetischen Eigenschaften nicht für die parasitäre Therapie geeignet. Vor kurzem wurde im humanen FPPS eine zusätzliche allosterische Tasche entdeckt, die basierend auf einer Sequenzanalyse wahrscheinlich auch in *T. brucei* FPPS vorhanden ist.

Die hohe medizinische Notwendigkeit in Kombination mit der Entdeckung einer potenziellen neuen Bindungstasche führte zu einem fragmentbasierten Ansatz zur Entwicklung von neuen nicht-Bisphosphonat-basierten Verbindungen für *T. brucei* FPPS, welcher in dieser Arbeit vorgestellt wird. Kristallographisches und NMR-basiertes Fragmentscreening wurde durchgeführt. Zu diesem Zweck wurde ein robustes *T. brucei* FPPS Kristallisationssystem etabliert, das die *High throughput*-Bestimmung von Kristallstrukturen mit einer Auflösung von bis zu 1,67 Å ermöglicht. Strukturelle Überlagerungen zeigten, dass die auf dem humanen FPPS gefundene allosterische Tasche, tatsächlich in *T. brucei* FPPS existiert. Diese Entdeckung ermöglichte anschließende NMR- und Kristallisationsxperimente mit etablierten humanen FPPS-Verbindungen, die zu drei Protein-Liganden-Komplexstrukturen mit gebundenen Fragmenten in der bisher unbekannten allosterischen Tasche führten.

Für die meisten der getesteten Verbindungen lag der durch SPR (*Surface plasmon resonance*) bestimmte K_d-Wert für *T. brucei* auβerhalb des experimentellen Testbereichs. Nur für eine Verbindung wurde der K_d-Wert auf *T. brucei* FPPS bestimmt und war um drei Größenordnungen höher als der IC_{50}-Wert auf dem humanen FPPS. Die Kristallstrukturanalyse ergab unterschiedliche Bindungsmodi auf den beiden FPPS Enzymen und zeigte reduzierte Protein-Ligand-Interaktionen auf *T. brucei* FPPS, wodurch die deutlich unterschiedliche Bindungsaffinität erklärt werden kann.

Ermutigt durch die Entdeckung erster Verbindungen die in der allosterischen Tasche auf *T. brucei* FPPS binden, wurden Fragment-Pools durch Ligand-NMR im Screen getestet und identifizierte Hits durch Experimente mit einzelnen Verbindungen in Liganden-NMR und Protein-NMR weiterverfolgt. Dadurch wurden 25 validierte Fragmente auf *T. brucei* FPPS entdeckt, die durch Soaking- und Ko-kristallisationsexperimente weiterverfolgt wurden. Kristalldaten wurden mit PanDDA (*Pan-Dataset Density Analysis*) analysiert und sieben Protein-Liganden-Komplexstrukturen wurden gelöst. Von den sieben Fragmenten waren vier

Fragmente in der katalytischen Tasche gebunden, ein Fragment wurde in der allosterischen Tasche entdeckt, die als Teil dieser Arbeit identifiziert wurde, und zwei Fragmente waren an Bindungsstellen an der Oberfläche gebunden. Insbesondere ein Fragment mit einem vieratomigen flexiblen Linker, das in der katalytischen Tasche gebunden war, war orthogonal zu den anderen drei Liganden und entlang der Helix α_D gebunden. Sechzehn Fragmentanaloge des länglichen flexiblen Fragmentes in der katalytischen Tasche wurden zur Optimierung der Struktur-Aktivitätsbeziehung getestet und eine Kristallstruktur mit einem Fragmentanalogen in der allosterischen Tasche wurde gelöst. Zusätzlich wurden Fragment-Screens durch Kristallisation bei XChem (Diamond, UK) und am EMBL/ESRF (Grenoble, FR) durchgeführt, die zu sieben Protein-Ligand-Strukturen führten. Ein Fragment war in der katalytischen Tasche gebunden, drei Fragmente in der allosterischen Tasche, zwei Fragmente in einer kryptischen Tasche zwischen den Helices α_I und α_H und ein Fragment auf der gegenüberliegenden Seite der allosterischen Tasche nahe den Helices α_G und α_F. Die Fragmentbindung wurde zusätzlich in Protein-NMR validiert.

Da Fragmente typischerweise eine geringe Bindungsaffinität bis in den mM-Bereich aufweisen, wurde eine strukturbasierte Fragment-Optimierung durch Medizinalchemie durchgeführt. Insgesamt wurden zehn Verbindungen synthetisiert und in Protein-NMR und in Kristallisationsexperimente getestet. Auffallend war ein fusioniertes Fragment basierend auf *T. brucei* und *T. cruzi* FPPS Verbindungen. Dieses Fragment ist in einer neuen Bindungsstelle nahe dem zweiten DDxxD Motiv gebunden.

Zusammenfassend stellt diese Arbeit hochauflösende Kristallstrukturen von *T. brucei* FPPS vor und zusätzlich wurden 19 Verbindungen entdeckt, die an sieben verschiedene Bindungsstellen der *T. brucei* FPPS binden. Dadurch wurde der Weg für zukünftige Studien zur Entdeckung von hochaffinen Nicht-Bisphosphonat-Inhibitoren mit pharmakokinetischen Eigenschaften für parasitäre Indikationen auf *T. brucei* FPPS geebnet.

Abbreviations and Acronyms

[w/v]	Weight per volume
[w/w]	Weight per weight
°	Degree
24-SMT	24-sterol methyl transferase
Å	Angström (10^{-10} m = 0.1 nm)
AceDRG	A stereo-chemical description generator for ligands
AIMLESS	Scale together multiple observations of reflections
AMBER	Assisted model building with energy refinement
approx.	Approximately
ATP	Adenosine triphosphate
autoPROC	A framework for automated data processing
bp	Base pair
BSA	Bovine serum albumin
BSF	Bloodstream form
CATT	Card agglutination test for trypanosomiasis
$CC_{1/2}$	Half-dataset correlation coefficient
CCD	Charge coupled device
CCP4	Collaborative Computational Project 4
clogP	The calculated logarithm of the octanol-water partition coefficient
COOT	Crystallographic object-oriented toolkit
CPMG	Carr Purcell Meiboom Gill
CRIMS	Crystallization Information Management System
CSP	Chemical shift perturbations
Da	Dalton
DIALS	Diffraction integration for advanced light source
DIMPLE	A pipeline for the rapid generation of difference maps from protein crystals with putatively bound ligands
DMAPP	Dimethyl allyl pyrophosphate
DNA	deoxyribonucleic acid
DND*i*	Drugs for Neglected Diseases *initiative*
DRC	Democratic Republic of the Congo
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
DTO	24,24-dimethylcholesta-5,7,25(27)-trienol
DTT	Dithiothreitol

EC_{50}	Effective concentration 50
ED_{50}	Effective dose 50
EDNA	A framework for plugin-based applications applied to X-ray experiment online data analysis
EMA	European Medicines Agency
ESRF	European synchrotron radiation facility
ETO	Ergosta-5,7,25(27)-trienol-3β-ol
FARM	First aspartate-rich motif
FBDD	Fragment-based drug discovery
FBS	Fragment-based screening
FGM	Fibroblast growth medium
FKBP	FK506 Binding Protein
FPPS	Farnesyl pyrophosphate synthase
FQ	Fit quality
GC-MS	Gas chromatography-mass spectrometry
GE	Group efficiency
GPP	Geranyl pyrophosphate
GrenADeS	Grenoble automatic data processing system
HAT	Human African Trypanosomiasis
HBA	Hydrogen-bond acceptor
HBD	Hydrogen-bond donor
H-bond	Hydrogen-bond
hkl	Miller indices
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA
HPLC	High pressure liquid chromatography
HRV 3C	Human rhinovirus 3C
HTS	High-throughput screening
iNEXT	Infrastructure for NMR, EM and X-rays for Translational Research
IPP	Isopentenyl pyrophosphate
ISPyB	Information system for protein crystallography beamlines
K	Kelvin
L. donovani	*Leishmania donovani*
LabCIP	Lab cleaning in place
LB	Luria-Bertani
LDM	lipid-depleted medium
LE	Ligand efficiency
LLE	Ligand-lipophilicity efficiency

MAD	Multi-wavelength anomalous diffraction
MASSIF	Massively Automated Sample Screening and evaluation Integrated Facility
MCSS	Multiple Copy Simultaneous Search
MDG	Methyl-α-D-Glucopyranoside
MEP	2-C-methyl-D-erythritol 4-phosphate
mins	Minutes
mmCIF	Macro-molecular Crystallographic Information File
MOE	Molecular operating environment
MVAK	Mevalonate kinase
MW	Molecular weight
MWCO	Molecular weight cut-off
N-BPs	Nitrogen-containing bisphosphonates
NECT	Nifurtimox-eflornithine combination therapy
NIBR	Novartis Institutes for BioMedical Research
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear magnetic resonance
NMT	N-myristoyltransferase
occ.	Occupancy
OD_{600}	Optical density at 600 nm
P. falciparum	*Plasmodium falciparum*
PAINS	Pan-assay interference compounds
PanDDA	Pan-Dataset Density Analysis
PCF	Procyclic form
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PFT	Protein farnesyl transferase
PGGT-I	Protein geranylgeranyltransferase-I
pH	Negative decimal logarithm of the hydrogen ion activity
PHENIX	Python-based Hierarchical Environment for Integrated Xtallography
POINTLESS	determine Laue group
PPIs	Protein-protein interactions
REFMAC5	Refinement of Macromolecular Structures, REFMAC
R_{free}	Free crystallographic R-factor
R_{meas}	Multiplicity-corrected R-factor
RNAi	RNA interference

ROTB	Rotatable bonds
rpm	Rounds per minute
SAR	Structure-activity relationship
SARM	Second aspartate-rich motif
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SOC	Super Optimal broth with Catabolite repression
SOFAST-HMQC	Band-Selective Optimized Flip Angle Short Transient – Heteronuclear Multiple Quantum Coherence
SPR	Surface Plasmon Resonance
SRA	Serum resistance-associated
T. brucei, Tbru, *T. b*	*Trypanosoma brucei*
T. gondii	*Toxoplasma gondii*
T2S	Type II secretion
TB	Terrific broth
TCI	Triple resonance inverse
TR	Trypanothione reductase
TROSY	Transverse relaxation optimized spectroscopy
TryS	Trypanothione synthetase
TSA	Thermal shift analysis
UV	Ultraviolet
VSG	Variant surface glycoproteins
WHO	World Health Organisation
wLOGSY	Water-ligand-observed *via* gradient spectroscopy
XDS	X-ray Detector Software
XDSAPP	Expert system and graphical user interface (GUI) for the automated processing of diffraction images using the *XDS* program suite
Xia2	Automated data reduction

For amino acid abbreviations, general one and three-letter codes are used.

1. Introduction

The focus of the thesis is the fragment based drug discovery process towards an inhibitor of *Trypanosoma brucei* (*T. brucei*) FPPS, a potential target for Human African Trypanosomiasis (HAT). Bisphosphonates already in clinical application for bone diseases and the identification of an allosteric pocket on human FPPS form the foundation of this project to discover non-bisphosphonate binders on *T. brucei* FPPS with pharmacokinetic properties that are appropriate for parasitic therapy. Since its first application in 1996, fragment-based drug discovery has shown its potential to identify new binding pockets on most challenging targets and has delivered two approved drugs, with more than 40 drugs in clinical studies.

In Chapter 1, the three different aspects of the project will be introduced. First, the HAT disease, the *T. brucei* parasite, current treatment options and clinical drug candidates for HAT are described. In the second part, FPPS as a potential drug discovery target is outlined, its role in the sterol biosynthesis is introduced and previous research on *T. brucei* FPPS is presented. The third part is dedicated to the fragment-based drug discovery highlighting its development, different screening techniques and hit optimisation approaches. Methods and Materials are described in Chapter 2. In Chapter 3, the development of the protein production, the crystallization screening, and optimisation phase leading to a 1.67 Å X-ray structure of *T. brucei* FPPS is present. Furthermore, additional *T. brucei* FPPS-fragment co-crystal structures of reported human allosteric binders were solved and revealed binding in a similar allosteric pocket on *T. brucei* FPPS (Chapter 4). Additional allosteric binders, non-bisphosphonate active site and fragments in new binding pockets on *T. brucei* FPPS were identified in an NMR and X-ray fragment screen and their binding mode will be discussed and first SAR by catalogue on several candidates will be shown (Chapter 5 and 6).

In Chapter 7, the structure-based fragment optimisation by medicinal chemistry using fragment growing and merging is described and a synthesised merged fragment has been developed binding in a new site. The basis for the active site fragments that were used as starting points for the synthesis and the synthetic strategies will be presented.

Chapter 8 gives a conclusion and outlook on the research of HAT treatment and Chapter 9 lists the used references. Lastly, Chapter 10 is the Appendix containing the plasmid maps, a list of compounds, SPR plots, crystallographic tables on data collection and refinement, additional density maps, and a description of the ConSurf parameters.

1.1 Human African Trypanosomiasis

1.1.1 *Trypanosoma brucei*

Trypanosoma brucei (T. brucei or T. b.) is a subspecies of the unicellular flagellate of the genus *Trypanosoma,* which causes the human African Trypanosomiasis (HAT). Three morphologically indistinguishable subspecies have been described: *T. b. rhodesiense* causing the acute form of the disease, *T. b. gambiense* resulting in the chronic progress of the disease and *T. b. brucei*, which is only infectious to animals.[1-2] All three species contain a characteristic flagellum starting from the flagellar pocket at the front end of the parasite, which is essential for viability of the parasite, a mitochondrion in which the kinetoplast resides, a nucleus and glycosomes as well as a plethora of other organelles.[3-4] Both human-infective subspecies are transmitted by the bite of the blood-feeding *Glossina*, also well-known as tsetse fly (Figure 1).[5-6]

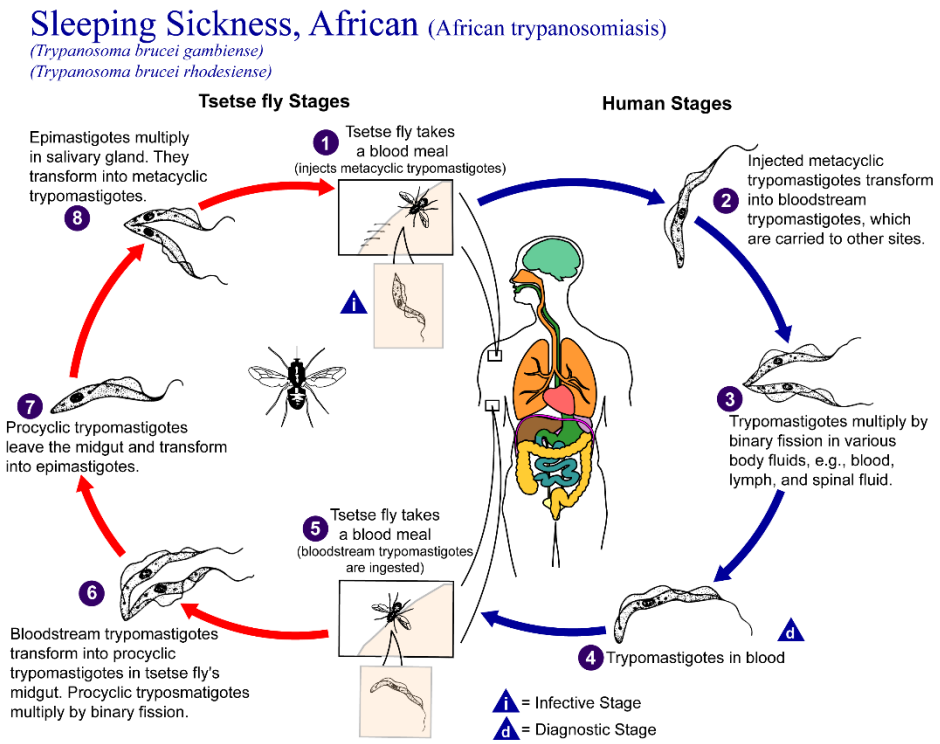


Figure 1: *T. b gambiense* and *T. b rhodesiense* life cycle. The tsetse fly stage and human stage of the parasite with its morphological changes are depicted. From Alexander J de Silva and Melanie Moser, Centers for Disease Control Public Health Image Library.

After ingestion of the stumpy bloodstream trypanosomes in the blood of a tsetse fly, the trypanosomes undergo a complex morphological change in the fly tissues through differentiation to procyclic trypomastigotes. The trypomastigotes replicate by binary fission in about 18 – 35 days until they reach their free floating final metacyclic trypomastigote form, which is infective for human beings. The metacyclic trypomastigotes and tsetse saliva are transmitted to humans by the

bite of the fly.[6-12] Only a population of 0.01 – 0.1 % of the tsetse flies carry the mature infective trypanosomes and can transmit the disease but with a blood meal every three days and being infective for the rest of the lifetime one tsetse fly can infect many people.[4, 13] Once humans are infected the parasite enters the lymphatic system and is carried to the bloodstream where metacyclic trypomastigotes undergo the transformation into bloodstream trypomastigotes. From there on, they reach other parts of the body and body fluids and further multiply by binary fission. If a tsetse fly takes a blood meal on this infected host, bloodstream trypomastigotes are taken up by the fly closing the complex life cycle of *T. brucei* (Figure 1).[13-16]

The full genome of *T. brucei* trypanosomes was recently sequenced and 10 % of the genome encode for their surrounding surface coat consisting of variant surface glycoproteins (1250 – 1500 VSG genes).[10, 17-18] Through a high level of antigenic variation characterised by constant change of the expression of one of these VSGs, trypanosomes repeatedly overcome the host immune response and this also makes it difficult to develop a vaccination.[10, 19-20]

The human serum also contains lytic factors to combat trypanosomal infections, one of which includes the human trypanolytic factor containing apo lipoprotein apoL1.[21-22] *T. b. rhodesiense* expresses serum resistance-associated (SRA) protein, which neutralizes apoL1 and therefore prevents pore formation as part of the lytic attack of the host organism.[23-24] In contrast, *T. b. gambiense* does not contain the *sra* gene and is described to resist lysis by lower levels of apoL1 uptake. This is achieved by utilizing its haptoglobin-haemoglobin receptor combined with a membrane stiffing by a *T. b. gambiense* specific glycoprotein.[22, 25-26] In contrast to *T. b. brucei*, these mechanisms have enabled the other two subspecies to combat the host immunity and to cause the dreaded HAT disease.[22]

1.1.2 Epidemiology of human African trypanosomiasis

HAT is a neglected disease in 36 sub-Saharan African countries, thus it is considered as endemic in these countries. *T. b. gambiense* infections account for 97 – 98 % of all reported HAT disease cases and they occur in West and Central Africa (Figure 2).[15] More specifically, 84 % of all cases were reported in the Democratic Republic of the Congo (DRC) in 2016.[27] In contrast, *T. b. rhodesiense* parasitic infections are found in Eastern and Southern Africa with most cases reported in Uganda.[28] Millions of people were killed since the beginning of the 20th century and the last peak in transmission was reached just before the start of the 21st century.[2, 29] Hereafter, strict control programs of people at risk especially in rural areas and health institutes for the treatment of HAT were introduced in endemic countries. In consequence, the number of infected people is steadily decreasing since. In 2016, overall 2164 new HAT cases were reported by the World Health Organisation (WHO) and the promising downwards trend had led the WHO to target HAT elimination as a public health problem by 2020, which in detail means less than 2000 reported

cases per year.[27] However, the number of occurring cases is thought to be significantly higher since the number of undetected cases should not be disregarded as one of the most neglected populations is affected.[4, 30] Around 57 million people are estimated to be still at risk for infection.[27] Additionally, through migration and tourism to HAT risky areas especially in national parks and game reserves exported cases were also reported in non-endemic countries. In total, 94 cases were reported in non-endemic countries across all five continents in the period of 2000 – 2010.[31] The United Republic of Tanzania represents the main country of origin for *T. b. rhodesiense* infections with 59 % of all cases whereas exported cases of *T. b. gambiense* infections originate mainly from the DRC and Gabon with each accounting for 23 % of all cases.[31]

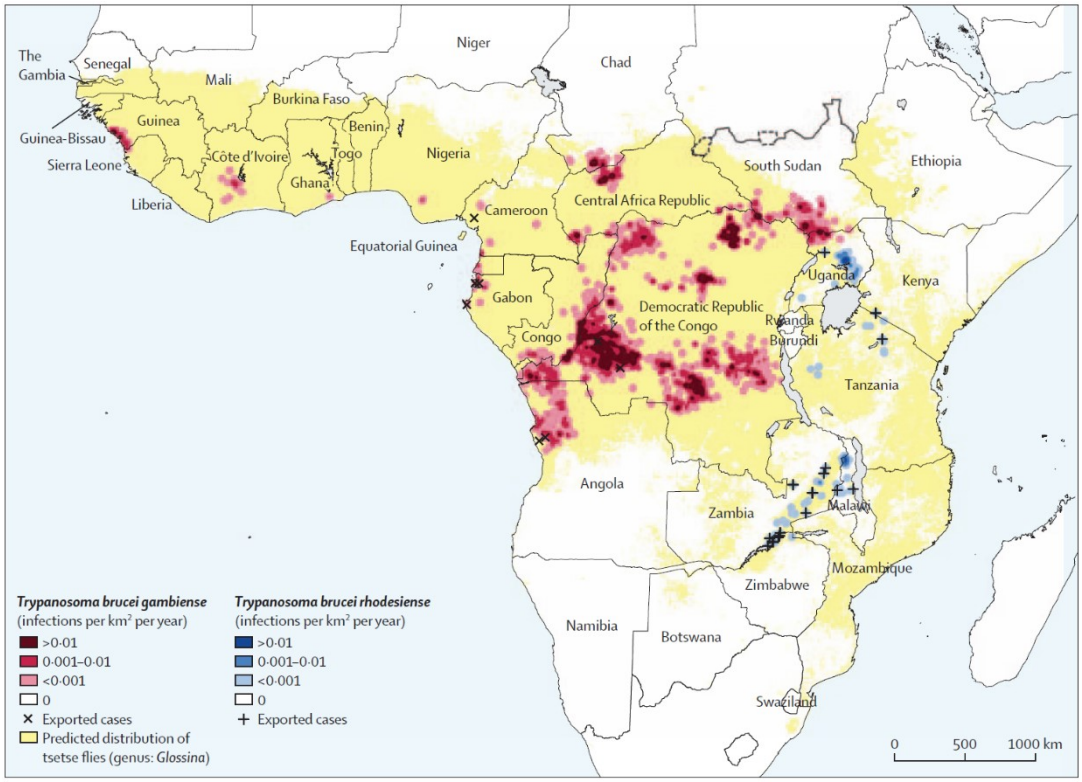


Figure 2: Local distribution of *T. b. gambiense* and *T.brucei rhodesiense* infections. *T. b. gambiense* and *T. b. rhodesiense* infections are in conditional formatting from dark red and dark blue for more than 0.01 infections per km^2 per year to white for zero cases, respectively. Yellow shaded areas show the predicted distribution of the tsetse flies. Reprinted from reference[4] with permission from Elsevier.

1.1.3 Clinical progress of the disease

Once the host gets infected, two stages are differentiated, a haemolymphatic and a meningoencephalitic stage.[10] For *T. b. rhodesiense* infections on humans, a trypanosomal chancre of 3 – 4 cm characterized by heat and local erythema might appear around the bite of the tsetse fly within 2 – 3 days. However, for *T. b. gambiense* this first clinical sign is rarely seen.[4, 32] In the so-called first stage of the disease, unspecific symptoms such as lymphadenopathy, headache, joint pain, fatigue, weight loss, malaise, and intermittent fever can occur while parasites are only found

in the blood and lymph.[10, 33-34] From here on, the disease progress highly differs between the two subspecies. Typically, rhodesiense-HAT infections progress to the second meningoencephalitic stage within weeks after infection and show distinct differences in various foci areas.[15, 35] In contrast, for gambiense-HAT, the mean duration for progression to the second stage is reported to be around 3 years with strong individual differences that can vary between a few months up to more than 15 years.[4, 36-37] Hence, the rhodesiense-HAT is classified as the acute form and patients generally die within 6 months whereas patients with gambiense-HAT can live with the disease for a number of years.[38] Nevertheless, both forms are considered fatal if untreated or diagnosed too late.[13]

The second stage is reached when the parasites have crossed the blood-brain barrier and have invaded the central nervous system resulting in severe neurological symptoms and mental changes such as motor weakness, speech disorders, walking difficulties, emotional lability, confusion, dementia and sleep disturbance with daytime somnolence and insomnia, giving the characteristic name of the diseases.[10, 39-40] Nevertheless, clinical symptoms between different patients vary considerably and the clinical signs between the two stages overlap. Thus, diagnosis and recognizing the staging of the disease have proven to be difficult.[40-41]

1.1.4 Diagnosis

Unfortunately, no effective vaccination or drug as prophylaxis to prevent HAT is on the market.[10, 42] Hence, the general strategy for limiting the disease is the reduction of the distribution of the vector *Glossina* together with a rapid diagnosis and treatment before the progress to the second stage of the disease.[15] Two different strategies of diagnosis are followed for the two parasitic subspecies. In general, parasitological diagnosis of *T. b. rhodesiense* infections is relatively easy due to a high level of parasitemia. The preparation of a fresh thick blood sample for microscopic examination allows the detection of trypanosomes with a sensitivity of 5000 trypanosomes/mL.[15, 34, 43-44] On the other hand, detection of *T. b. gambiense* infections pose a challenge and the development of the serologic card agglutination test for trypanosomes (CATT) in the late 1970s significantly improved diagnosis.[45] Nowadays, the test is also commonly used for screening of the population in areas with prevalence of gambiense-HAT.[10] The test uses the antigen type LiTat1.3 of the *T. b. gambiense* and can be performed on blood, plasma or serum samples. In the latter case, the specificity is highly improved.[45-46] After diagnosis of infection with the disease, diagnosis of the stage needs to be followed to decide for an appropriate treatment. For both HAT forms, disease staging is most commonly carried out by examination of the cerebrospinal fluid with respect to the number of white blood cell count and the detection of trypanosomes.[10, 15] Once the stage of the disease is determined, an appropriate treatment is initiated; however, available drugs are at least 30 years old, have strong limitations, and show various degrees of toxicity.[47]

1.1.5 Current treatment options for HAT

The treatment of choice for first-stage *T. b. gambiense* is pentamidine isethionate, which was first introduced in 1937 and is administered by intramuscular injection of 4 mg/kg for seven days (Figure 3).[42, 48] This mode of action is still not elucidated but it is known that the drug molecule tightly interacts with DNA and disrupts the mitochondrial functions.[42, 49-51] The therapy is generally well-tolerated, but side effects such as hypotension and nephrotoxicity and gastrointestinal problems are reported.[4, 52-53]

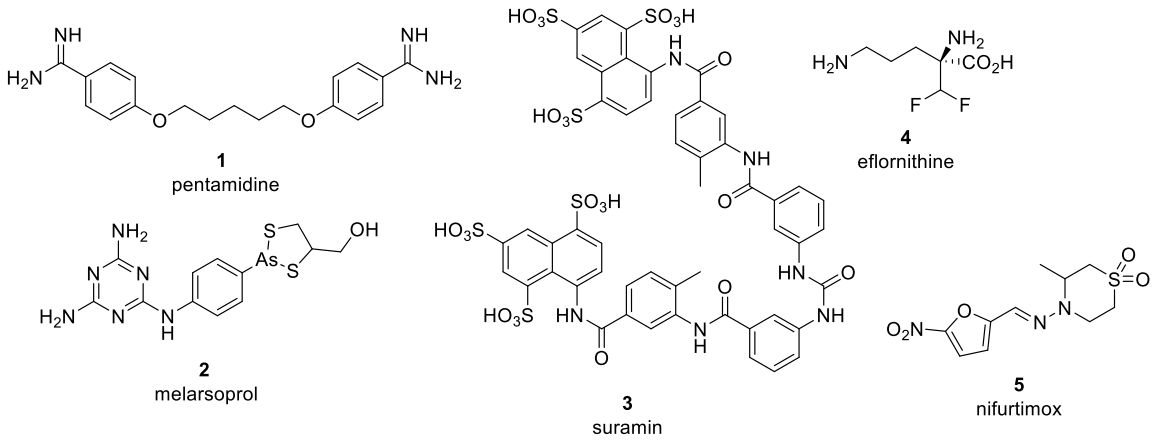


Figure 3: Currently available drugs for the treatment of HAT disease. Chemical structures of pentamidine, melarsoprol, suramin, eflornithine, nifurtimox are shown from **1** – **5.**

Suramin, also known as Bayer 205 was first synthesised in 1917 at the German pharmaceutical company Bayer and is the first-in-line drug for the treatment of the first stage of *T. b. rhodesiense* infections (Figure 3).[54] The drug needs a slow intravenous administration and the treatment program can last up to one month with 20 mg/kg injections every 3 – 7 days and thus, requires long hospital stays.[55] Side effects frequently occur but they are relatively mild and include sometimes among others renal failure, anaemia, fatigue and pyrexia.[55-56] Suramin is also effective against *T. b. gambiense* but is rarely used for the gambiense-HAT treatment due to easier administration of pentamidine.[55]

For the second stage of the disease, drugs that cross the blood-brain barrier are required.[40] Historically, melarsoprol, an organo-arsenical compound, and eflornithine, an analogue of ornithine, were the only available treatment options as monotherapies.[57-59] Eflornithine modulates the ornithine decarboxylase function and, most importantly, it is the only drug with a well-described molecular target in this context.[60-62]

The introduction of the nifurtimox-eflornithine combination therapy (NECT) was a breakthrough in the treatment of HAT and is now the most efficient and cost-effective option to treat *T. b. gambiense* infections.[63-66] Nifurtimox is only registered for the treatment of Chagas disease and therefore, treatment of patients with NECT requires previous authorization of the

ministry of health.[4, 15, 67] In 2009, NECT was included in the WHO list of essential medicines and is provided free of charge to endemic countries and after one year 88 % of all gambiense-HAT cases were already treated with NECT.[15, 68] NECT regimen consists of an oral administration of nifurtimox every 8 hours for 10 days combined with 14 intravenous infusions every 12 hours for a period of ten days.[63] Therefore, this combination therapy allowed the reduction of eflornithine intravenous infusions from 56 to 14 applications over a period of 10 days, which resulted in a reduced amount of expensive drug needed, a decreased logistic burden and shortened hospital stays.[15, 64, 69] Additionally, concerns on the higher risk of infections in field conditions than in previous highly controlled clinical trials were not confirmed and the NECT was shown to be justified in endemic countries.[70] Most importantly, the combined treatment showed reduced side effects, overall decreased mortality and lower relapse rates compared to eflornithine monotherapy.[63-64] Eflornithine monotherapy is now only the treatment of choice if nifurtimox would be contraindicative or not accessible.[15]

For the treatment of second stage rhodesiense-HAT, melarsoprol is still the only option. It requires an intravenous slow bolus injection of 2.2 mg/kg/day for 10 days.[71] However, in 5 – 10 % of treated patients an encephalopathic syndrome develops that is fatal for half of the patients. Apart from that, headache, pyrexia, gastrointestinal and skin reactions also frequently occur.[4, 72-73] New drugs are urgently needed with an improved medical profile and easier administration. For that, mainly phenotypic-based high-throughput screening (HTS) and target-based approaches are followed as discussed in the next chapters.

1.1.6 Phenotypic screening

Phenotypic-based HTS is a well-established approach in drug discovery projects for neglected diseases.[74-77] It aims to identify molecules that alter the cell's phenotype in a disease relevant way and has thus the potential to identify molecules with new modes of action without the knowledge of a specific target.[78]

In 1997, Rãz *et al.* described a widely applied viability assay on bloodstream form (BSF) *T. brucei* based on the dye Alamar Blue™ (resazurin) in a 96 well format.[79-81] Among other studies, the assay was used to screen a library with 4-aminoquinoline-based heterodimeric compounds and the most potent compound **6** (Figure 4) was determined with an IC_{50} of 0.12 µM.[82] Interestingly, the inhibitory efficiency was partly explained by the inhibition of the trypanothione reductase but other biological effects are probably included by its activity.[82] This whole cell viability assay was further developed in a 384-well format for high-throughput approaches and applied for an HTS of a 87,296 library (WEHI 2003 collection)[83] against BSF *T. b. brucei* lister 427 cells.[84-85]

Similarly, a 96-well and 384-well format for a firefly luciferase-based whole cell viability assay on BSF *T. brucei* strain 427 were developed and compared to the Alamar Blue assay.[86-87] From a firefly luciferase-based HTS, the optimisation of a hit identified the imidazopyridine lead compound **7** (Figure 4) with high potency for first stage HAT and cured *T. brucei* infected mice after 2.5 mg/kg oral administration.[88]

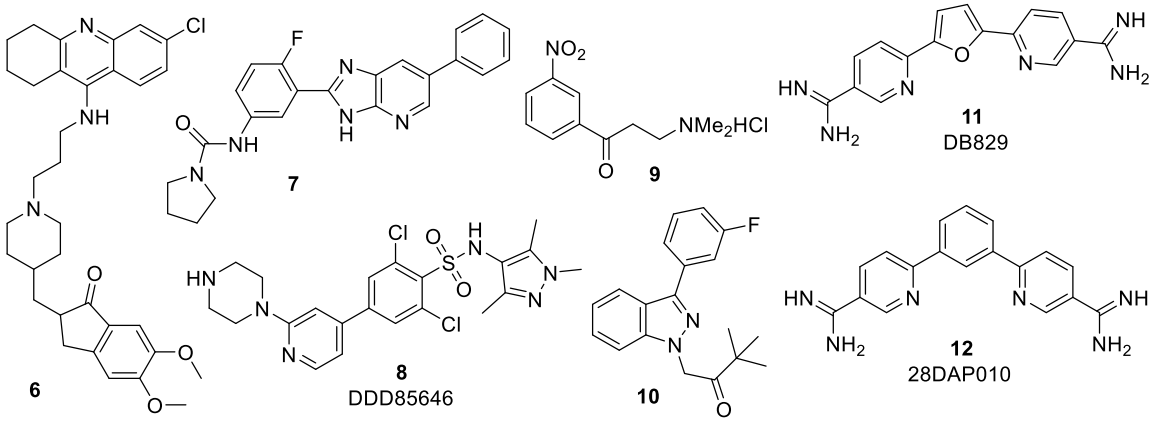


Figure 4: New molecules with activity on *T. brucei* cells or *T. brucei* enzyme targets. Compound **6** and **7** were identified in HTS phenotypic assay against whole cell *T. brucei*. Compounds **8** – **10** showed activity on N-myristoyltransferase, trypanothione reductase, trypanothione synthase, respectively. Compounds **11** and **12** exhibited activity for first and second stage disease and are inspired by the clinically used drug, pentamidine.

Recently, Faria *et al.* reported a new robust and cost-effective whole cell assay based on the cyanine dye SYBR® Green. At the same time, they screened a 4000 compound kinase focused library in the newly developed screen and the Alamar Blue screen to compare the assay performance and to identify new potent antitrypanosomal chemical scaffolds.[89]

1.1.7 Target based approaches and clinically drug inspired new molecules

In HAT drug discovery, target-based approaches are widely applied as an alternative to phenotypic screening in identifying new molecules with anti-trypanosomal activity, generally in an HTS campaign against a specific target.[90] A plethora of targets was investigated as summarized in various comprehensive reviews.[51, 90-92] The enzyme N-myristoyltransferase (NMT) was shown to be essential for *T. brucei* parasitic growth and is a validated drug target for HAT.[93-94] From a 62,000 diversity-based compound library[95] screen several hits were optimised including the pyrazolesulfonamide, DDD85646 (Figure 4) shown to cure first stage *T. brucei* infections in mice.[96-97] Another potential target is trypanothione reductase (TR) and Martyn *et al.* identified in an HTS of 134,500 compounds by measuring the activity of TR[98] compound **9** (Figure 4) with an IC_{50} of 0.34 µM against *T. brucei* and a 59-fold selectivity over human glutathione reductase.[51, 99-101]

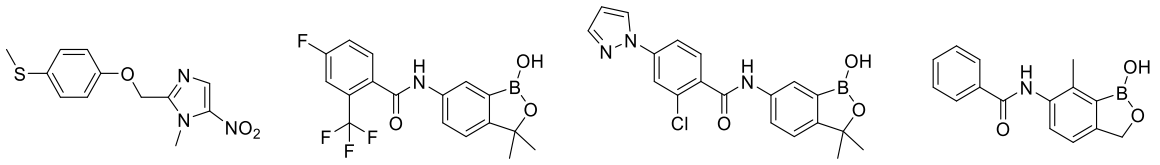
Trypanothione synthetase (TryS) is an essential target for *T. brucei* survival. In a drug discovery campaign, indazole compound **10** (Figure 4) was identified, which showed to inhibit

T. brucei TryS with an IC_{50} of 0.09 µM.[102] Recently, there is also increasing interest in rhodesian cysteine proteases as potential drug targets and series of peptide-based and peptidomimetic inhibitors showed promising results.[103-104] Other investigated targets include cathepsin B[105], Aurora kinase 1[106], pteridine reductase[107-108], fructose 1, 6-bisphosphate aldolase[109-110] and phosphofructokinase[111]. However, most importantly, farnesyltransferase[112-113] and several target enzymes of the sterol biosynthesis, including FPPS[114-118], were also studied and a comprehensive summary is given in Chapter 1.2.2.

Another area of interest has been on clinical drug-inspired molecules, such as diamidines due to first stage pentamidine therapy in the clinic. These compounds are thought to target AT-rich stretches of DNA, several studies based on amidines, and diamidines have been described elsewhere.[119-123] A bisamidine compound entered clinical trials but unfortunately, the trial was discontinued as patients developed transaminases and acute renal insufficiency.[124] Most importantly, DB829, an aza-diamidine and 28DAP010 (Figure 4), a related analogue, showed both efficacy against the first and second stage of the disease caused by both *T. brucei* subspecies.[125-126] DB829 showed curation in a monkey model at a dose of 2.5 mg/kg intramuscular treatment and 28DAP010 demonstrated similar *in vivo* efficacy in mouse models.[126-127] These compounds show high potential for further drug development but the programs were stopped due to the two compounds, fexinidazole and SCYX-7158 which exhibit a superior profile as described in the next chapter.[128]

1.1.8 Drug candidates in clinical trials

An orally available, safe and effective drug was set as the overall goal of the Drugs for Neglected Diseases *initiative* (DND*i*) by 2018. Fexinidazole, a 2-substituted 5-nitroimidazole (Figure 5), was initially developed by Hoechst in the 1980s, further abandoned and rediscovered in 2005 by DND*i* and is now the most advanced compound in the drug development process with a potential oral treatment for first and second stage gambiense-HAT.[129-130] A DND*i* phaseII/III trial conducted in the DRC and the Central African Republic with a 10-day once-a-day oral treatment regime of 1800 mg on days 1 to 4 and 1200 mg on days 5 to 10 demonstrated that the treatment is well-tolerated and effective.[131] Therefore, the fexinidazole regulatory dossier was accepted by the European Medicines Agency (EMA) under the Article 58 in 2017.[132] In 2018, the EMA's Committee for Medicinal Products for Human Use released a positive opinion of fexinidazole and the drug was registered in the DRC end of 2018.[133] Another phase II/III trial is prepared to start in mid-2019 in Malawi to assess fexinidazole also for rhodesiense-HAT.[134]



13
Fexinidazole

14
SCYX-7158
Acoziborole

15
SCYX-1608210

16
SCYX-1330682

Figure 5: Drug candidates for the treatment of HAT in clinical trials. Chemical structures of fexinidazole, SCYX-7158, SCYX-1608210 and SCYX-1330682 are shown.

SCYX-7158 (Acoziborole, Figure 5), a benzoxaborole is another candidate that was optimised from a lead series after an initial screen of a benzoxaboroles library in a *T. brucei* viability assay.[135] This benzoxaborole could potentially be administered as a single oral dose due to its 400 hours long half-life.[42] Additionally, two back-up candidates SCYX-1330682 and SCYX-1608210 (Figure 5) also showed efficacy in mouse models but the program was set on hold due to the promising results of SCYX-7158 in preclinical and phase I studies.[136-137] In the phase I study with SCYX-7158 the administration of a single 960 mg dose was determined and the penetration through the blood-brain barrier could be confirmed to enable treatment of second stage gambiense-HAT.[138] In 2016, a pivotal DND*i* phaseII/III trial (ClinicalTrials.gov Identifier: NCT03087955) was initiated for SCYX-7158 with the continued recruitment of 191 patients in 2018 and the goal of submission of a regulatory dossier to the EMA under Article 58 in 2021.[139]

In summary, it is a long route until the approval of a new drug but the development of orally available fexinidazole is a therapeutic breakthrough for a safe and effective HAT treatment. Furthermore, SCYX-7158 also shows a high potential to be introduced for treatment. Additionally, many other research programs are ongoing and one of them will be described in this thesis.

1.2 *T. brucei* FPPS as a target enzyme for HAT

1.2.1 Sterol biosynthesis

Farnesyl pyrophosphate synthase (FPPS) is part of the sterol biosynthesis, which is facilitated by a highly regulated and important metabolic pathway found in most eukaryotes, bacteria, fungi and plants.[140-143] The pathway is particularly well-studied in humans and part of the synthesis for sterol isoprenoids and nonsterol isoprenoids, such as cholesterol, dolichols, ubiquinones, isopentenyl tRNA.[144] More specifically, cholesterol is involved in the membrane structure and is further converted into bile acids, lipoproteins, steroid hormones with a vital role in a variety of cellular processes[145-146], ubiquinones participate in the electron transport[147], and dolichol is essential for glycoprotein synthesis.[148] Furthermore, isoprenoid post-translation modification of proteins,

such as Rho, Rac or Ras, plays a critical role in protein localisation and activity, and cellular signalling cascades.[149-150]

It is also well-established that the sterol biosynthesis is essential in *T. brucei* trypanosomes; however, in contrast to mammalian and similar to plants and fungi, trypanosomes synthesise ergosterol and various other 24-methyl sterols instead of cholesterol.[151-152] Ergosterol differs from cholesterol by two double bonds in the B ring, an additional double bond at C_{22} and a β-methyl at position 24.[153] Before studying the sterol biosynthesis in *T. brucei* it is noteworthy to mention that two experimentally studied different *T. brucei* forms are distinguished in the *T. brucei* life cycle: the procyclic form (PCF) (in tsetse fly) and the bloodstream form (BSF) (in host). These two forms have a remarkable different sterol content and sterol biosynthesis activity. The PCF and BSF *T. brucei* are important to be distinguished when describing the sterol biosynthesis.[154-155] Contradictive literature is reported on the necessity of the endogenous biosynthesis of sterols in BSF *T. brucei*, which is relevant for antiparasitic drugs targeting the ergosterol biosynthesis as discussed in the following chapters.

1.2.1.1 Sterol biosynthesis in procyclic form of *T. brucei*

The de novo sterol biosynthesis starts with the mevalonate pathway. Leucine, which is converted in a 5-step process into 3-hydroxy-3-methylglutaryl-Coenzym A (HMG-CoA), was described as the main carbon source in PCF of *T. brucei*.[155] Apart from that, an SCP2 thiolase, similarly to the acetoacteyl-CoA-thiolase in mammalian, was identified in *T. brucei* as essential for sterol biosynthesis from ketogenic sources, such as glucose, threonine and acetate.[17, 156-158] SCP2 thiolase catalyses the condensation of two acetyl-CoA units to acetoacetyl-CoA (Figure 6).[156] Subsequently, acetoacetyl-CoA is catalysed by 3-hydroxy-3-methylglutaryl-CoA (HMG)-synthase to form HMG-CoA with the addition of a third acetyl-CoA unit.[159-160] Then HMG-CoA is reduced to mevalonate through HMG-CoA reductase, which is the last step of the three-step mevalonate pathway.[153, 161] The early steps of the mevalonate pathway are localised in the mitochondrion. While mammalian cells have a mitochondrial and cytosolic HMG-CoA synthase, both the HMG-CoA synthase and the HMG-CoA reductase are localised in the mitochondrion in *T. brucei*.[160, 162-163]

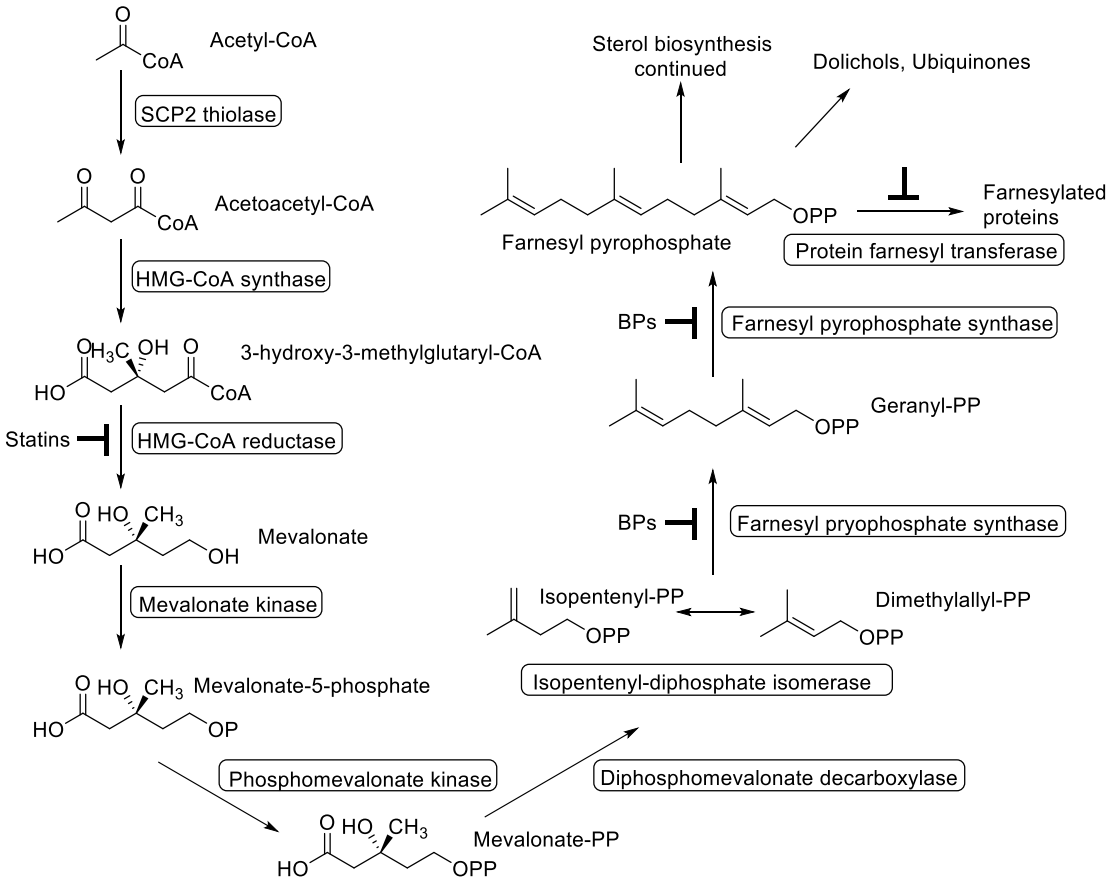


Figure 6: First part of sterol biosynthesis in *T. brucei*. The figure depicts the steps of the sterol biosynthesis from Acetyl-CoA to farnesyl pyrophosphate and the enzymes in brackets, which catalyse the steps. Potential blocking of the catalytic steps by inhibitors is marked. Adapted from [164].

The mevalonate pathway is followed by the isoprenoid pathway.[153] In the first step, mevalonate is phosphorylated to mevalonate-5-phosphate through the mevalonate kinase (MVAK) using ATP.[160, 165] The MVAK is conserved across other trypanosomatids and MVAK of *T. brucei* contains signal sequences PTS1 and PTS2, targeting it for transport to the peroxisomal matrix.[160, 166-167] Proteomic studies, a digitonin titration and indirect immunofluorescence experiments showed further evidence for a glycosomal location of MVAK but how it is imported into the glycosome for phosphorylation through MVAK is not clear.[160, 168]

An additional ATP-dependent phosphorylation step catalysed by phosphomevalonate kinase forms mevalonate-pyrophosphate. In a decarboxylation step diphosphomevalonate decarboxylase[169] forms isopentenyl-pyrophosphate (IPP).[169-171] Further isomerisation of isopentenyl pyrophosphate to dimethyl allyl pyrophosphate (DMAPP) by the isopentenyl-diphosphate isomerase enables the formation of longer isoprenoids.[153] In a first head-to-tail condensation reaction of IPP with DMAPP the 10-carbon isoprenoid geranyl pyrophosphate (GPP) is formed. Then GPP undergoes another condensation reaction with a new molecule of IPP to form farnesyl pyrophosphate (FPP). These two consecutive condensation reactions are catalysed by the enzyme farnesyl pyrophosphate synthase (FPPS).[153, 170, 172] FPPS constitutes the last enzyme of the

isoprenoid pathway and is followed by committed steps of the sterol biosynthesis.[153] For the PCF of *T. brucei* the localisation of FPPS was thoroughly studied by immunofluorescence microscopy analysis, a permeabilisation assay and Western blot.[173] In these studies, FPPS was mainly found in the cytoplasm with a potential additional localisation in the endoplasmic reticulum (ER). Additional immunofluorescence experiments with BSF of *T. brucei* support the same assumption about the localisation.[173]

FPP is a branching point and can either be the substrate for farnesylated and geranylgeranylated proteins and the formation of dolichols, ubiquinones or it is further catalysed to squalene as part of the sterol biosynthesis (Figure 7).[112, 174-176]

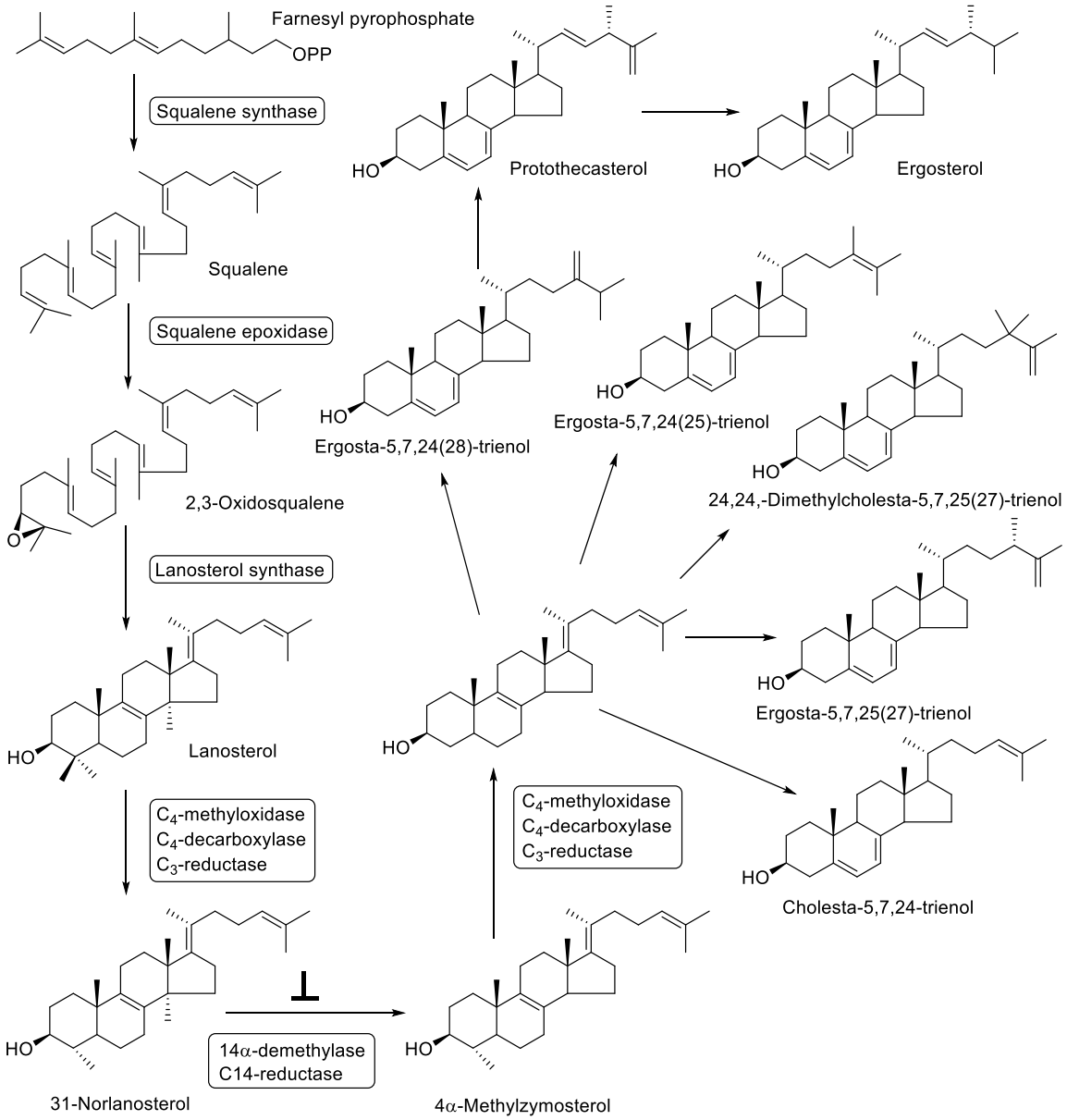


Figure 7: Second part of sterol biosynthesis in *T. brucei*. The figure depicts the steps of the sterol biosynthesis from farnesyl pyrophosphate to ergosterol, different final products in *T. brucei* and the enzymes in brackets, which catalyse the steps. Potential blocking of the catalytic step from 31-norlanosterol to 4α-methylzymosterol by inhibitors is marked. Adapted from [164] and [177].

The following reactions of the sterol biosynthesis are reported to take place in the ER.[155, 178] First, the formation of squalene is catalysed by squalene synthase and through presqualene-diphosphate, squalene is formed by the condensation of two FPP molecules.[153] Furthermore, squalene is the pre-requisite for the synthesis of 2,3-oxidosqualene catalysed by squalene epoxidase.[179]

Then, the 30-carbon chain 2,3-oxidosqualene cyclises in a highly complex reaction to lanosterol by lanosterol synthase and lanosterol is further demethylated at C_4 position to form 31-norlanosterol.[180-181] This is in contrast to mammalian, fungi and even *T. cruzi* pathways where C_{14} is first demethylated.[155, 182-183] In the next repetitive three-step process 31-norlanosterol is catalysed by sterol-14α-demethylase (CYP51)[184-185] and then through sterol C_{14}-reductase[186], 4α-methylzymosterol is formed.[187] Next, the second methyl group on C_4 is removed and zymosterol is released.[155] Zymosterol is the branching point for multiple product formation of ergosta-5,7,24(28)-trienol, ergosta-5,7,24(25)-trienol, 24,24,-dimethylcholesta-5,7,25(27)-trienol (DTO), ergosta-5,7,25(27)-trien-3β-ol (ETO) and cholesta-5,7,24-trienol (Figure 7). Methylation of the side chain to yield 24-alkyl sterols are catalysed by the AdoMet-dependent 24-sterol methyl transferase (24-SMT), which is absent in mammalian cells. Ergosta-5,7,25(27)-trienol is further catalysed to ergosterol (ergosta-5,7,22(23)-trienol) through prototheca sterol.[177] Interestingly, the genetic information for a C25(27)-reductase and a C_{22}-desaturase to form ergosterol was described to be absent. This is in contrast to other organisms, including *T. cruzi*.[155, 164] However, as proposed by Zhou *et al.* it could also be a possibility that due to an unusual sequence of these enzymes the genetic database still has its limitations.[182]

Gas chromatography-mass spectrometry (GC-MS) analysis of procyclic cells grown on fibroblast growth medium (FGM) showed equal amounts of cholesterol and endogenously synthesised sterols, mostly cholesta-5,7,24-trienol and ergosta-5,7,25(27)-trienol with only 3.5 % ergosterol of the total sterol content.[188] However, other studies of procyclic cells indicated ergosterol accounting for 14.3 % of the total sterol mass.[189]

1.2.1.2 Sterol biosynthesis in bloodstream form

For BSF of *T. brucei* it was long believed that due to their receptor-mediated uptake of cholesterol from the host, the endogenous biosynthesis of sterols is completely eliminated, which would limit the effect of antiparasitic drugs targeting the ergosterol biosynthesis.[190-191] However, gene expression of all genes for sterol biosynthesis in BSF of *T. brucei* was shown and 11 – 12 isoprene unit long dolichols and ubiquinones requiring FPP were identified by high-performance liquid chromatography (HPLC) and MS. *In vitro* investigation identified a functional 24-alkyl sterol biosynthesis.[155, 182] Nevertheless, the enzyme activity and expression levels are different in PCF and BSF *T. brucei*. The activity of HMG-CoA reductase was found to

be 3-fold lower and HMG-CoA synthase, MVAK and 24-SMT were expressed 2.5-, 3.78- and 3-fold lower in BSF than in PCF *T. brucei*.[155, 160, 189, 192] Sterol biosynthesis inhibitors were tested on both parasitic life stages and most importantly, growth inhibitory effects *in vitro* and *in vivo* were observed on PCF and BSF *T. brucei* as further discussed in the next chapter.[115, 155, 182, 192-194]

Sterols have a major structural function for membranes and cholesterol, exogenously taken up from the host, accounts for more than 95 % of the total sterol content in BSF *T. brucei.*[182] In yeast, a dual role of ergosterol as a component of the membrane and a sparking role in the signalling of cell proliferation has been described.[195-196] A similar picture emerged for *T. brucei*, cholesterol is the bulk sterol in membranes but ergosterol could replace it partly to signal for cell growth and function as a metabolic molecule. Notably, Zhou *et al.* reported a close association of cholesterol import with the regulation of the ergosterol biosynthesis. In case of limited cholesterol availability, ergosterol biosynthesis is upregulated and with high exogenous cholesterol, the endogenous ergosterol biosynthesis is downregulated but never fully eliminated.[182]

However, in an analysis of the sterol content of BSF cells cultured on FGM, which resembles the cholesterol availability in human blood, 0.01 fg/cell ergosterol were detected by High-Performance Liquid Chromatography-Ultraviolet (HPLC-UV) scanning. In contrast, in cells grown on lipid-depleted medium (LDM), 0.005-0.001 fg/cell ergosterol were detected. This can only occur via upregulation of the ergosterol biosynthesis with limited cholesterol availability.[188] In any case, the necessity of these small amounts for cell proliferation and viability was further shown by experiments with *T. brucei* 24-SMT RNA interference (RNAi) cell lines on LDM together with 25-azalanosterol. In these conditions decreased cell growth was observed and cells were rescued by the addition of 1.2 µM ergosterol. These results clearly show, contrary to previous assumptions, that ergosterol biosynthesis in BSF *T. brucei* is relevant and that enzymes of the sterol biosynthesis are valid targets for antiparasitic drug discovery. Several research projects on different sterol biosynthesis targets are reported in the literature and described in the next chapter.

1.2.2 Sterol biosynthesis inhibitors in procyclic and bloodstream form *T. brucei*

Enzymes of the sterol biosynthesis and further enzymes using substrates of the sterol biosynthesis, such as protein farnesyl transferase (PFT), are extensively investigated as drug targets and a comprehensive number of studies is presented here. HMG-CoA reductase, one of the first enzymes in the mevalonate pathway can be inhibited by statins that block the catalytic binding site of HMG-CoA.[197-198] HMG-CoA reductase is a validated target in humans for the treatment of atherosclerotic cardiovascular disease. Statins are widely prescribed for their cholesterol-lowering properties and its potential is further investigated for additional applications.[199-202] Notably, incubation of PCF and BSF *T. brucei* cells with the statin synvinolin (simvastatin) also showed a dose-dependent inhibitory growth effect. PCF *T. brucei* grown on lipoprotein-free serum and

BSF *T. brucei* grown on complete serum were incubated with 25 µM synvinolin for 40 hours and a 2-fold *in vitro* growth reduction was observed.[192] Interestingly, Yokoyama *et al.* determined an around 100-fold lower IC_{50} (2 µM) of synvinolin against BSF *T. brucei* compared to PCF *T. brucei* (150 µM).[113] Similarly, Andersson *et al.* reported the growth inhibitory effect of 0.1 – 20 µM lovastatin, another HMG-CoA reductase inhibiting drug, after 10^3 U/ml interferon-γ stimulation of BSF *T. brucei* and showed a reversed effect with the addition of 10 mM mevalonate.[203] Another HMG-CoA reductase inhibitor, compactin led to altered morphology at cytostatic concentrations in procyclic cells after 3 days and 15 µM compactin resulted in complete growth inhibition for BSF cells.[204]

Furthermore, FPPS targeting nitrogen-containing bisphosphonates (N-BPs) are broadly in clinical use for various bone resorption diseases but these compounds have also shown anti-parasitic effects as discussed in detail in the next chapter.[114-118] The C_{15}-FPP and C_{20}-GPP products, apart from their role in the sterol biosynthesis, are attached to cysteine residues of proteins, such as GTPases, for protein farnesylation and geranylgeranylation with various roles in trypanosomes.[204-205] The transfer is catalysed by the enzymes PFT and protein geranylgeranyltransferase-I (PGGT-I) in *T. brucei* and inhibitors of the mammalian PFT[206-208] were shown to inhibit *T. brucei* PFT *in vitro*.[112-113] BSF cells already showed cell deformation after 6 hours and full lysis after 15 hours with incubation of 25 µM farnesyl-O-NH-PA ester.[113] Similarly, four CaaX mimetics (FTI-277, GGT-297, L-745,631, and SCH-44342) led to cell deformation after 24 hours and cytocidal effects on BSF *T. brucei*.[209] Several isothiazole dioxides were also tested for *in vitro* inhibition and an isothiazole dioxide with an aryl-isoxazolyl group showed an IC_{50} of 2 µM for *T. brucei* PFT and no rat PFT inhibition at the same concentration. Gelb *et al.* showed *in vitro* growth inhibition of BSF *T. brucei.*[210-211]

The squalene synthase was previously broadly studied as a target against atherosclerosis in humans.[212-213] In an investigation of different quinuclidine derivatives as inhibitors of squalene synthase, growth inhibition of *T. brucei* was observed but no activity against recombinant squalene synthase could be detected suggesting a different mode of action.[214-215] However, the sterol 14α-demethylase, CYP51 was found to be inhibited in a dose-dependent manner by various azoles and inhibition correlated well with anti-parasitic effects on PCF and BSF *T. brucei*.[193] Moreover, the complex structure of CYP51 with a non-azole compound N-1-(2,4-dichlorophenyl)-2-(1H-imidazol-1-yl)ethyl)-4-(5-phneyl-1,3,4-oxadiazol-2-yl) benzamide (VNI) was solved and oral administration of 20 mg/kg VNI to *T. brucei*-infected mice showed dose-dependent decreased parasitemia.[185]

One of the last steps of the sterol biosynthesis is the formation of 24-alkyl sterols through 24-SMT, an enzyme missing in mammalians.[183, 189] Azasterols have demonstrated inhibition of 24-SMT in *Saccharomyces cerevisiae*[216], *Leishmania* spp.[217], *T. cruzi*[217-218] and *Pneumocystis carinii*[219] but an investigation of azasterol analogues as 24-SMT inhibitors against *T. brucei* found

inhibition of PCF and BSF *T. brucei* growth with no association to 24-SMT inhibition.[189, 220] Nevertheless, an azasterol lead compound showed growth inhibition with an effective concentration 50 (EC_{50}) of 12 nM and the structure-activity relationship (SAR) was further investigated.[189, 221] More recently, substrate-based and transition state analogue inhibitors, azasterols and 25-thialanosterol sulfonium salt demonstrated selective and dose-dependent inhibition of BSF *T. brucei* with effective dose 50 (ED_{50}) values between 1 to 3 µM. Treatment with 5 mg/kg of 25-thialanosterol sulfonium salt prolonged life of mice by at least two days.[188] Further, Leaver *et al.* investigated the effect of 26-fluorolanosterol (26-FL) which irreversibly inactivated the 24-SMT and showed cell growth inhibition.[222] Whereas Miller *et al.* showed 50 % growth inhibition with the addition of around 15 µM 26, 27-dehydrolanosterol, a pro-drug that leads to inactivation of 24-SMT after C_{26}-methylation, to PCF and BSF *T. brucei.*[223]

In combination, these results show that the sterol biosynthesis pathway presents a plethora of enzymes that can be inhibited by various compound classes with *in vitro* and *in vivo* anti-parasitic effects on PCF and BSF *T. brucei.* Therefore, these enzymes constitute promising targets for antiparasitic drugs.

1.2.3 Farnesyl pyrophosphate synthase

Farnesyl pyrophosphate synthase (EC 2.5.1.10), also known as farnesyl diphosphate synthase (FPPS) was originally described in 1959[224] and is a homodimeric Mg^{2+} dependent enzyme of two 32 to 44 kDa subunits in various organisms.[225-231] In humans, FPPS consists of 13 α-helices connected by interhelical loop regions.[232] The enzyme has two highly conserved aspartate-rich motifs: first aspartate-rich motif (FARM) and second aspartate-rich motif (SARM) located at opposite sides of the active site, which is mostly formed by antiparallel α-helices (Figure 8).[233-236]

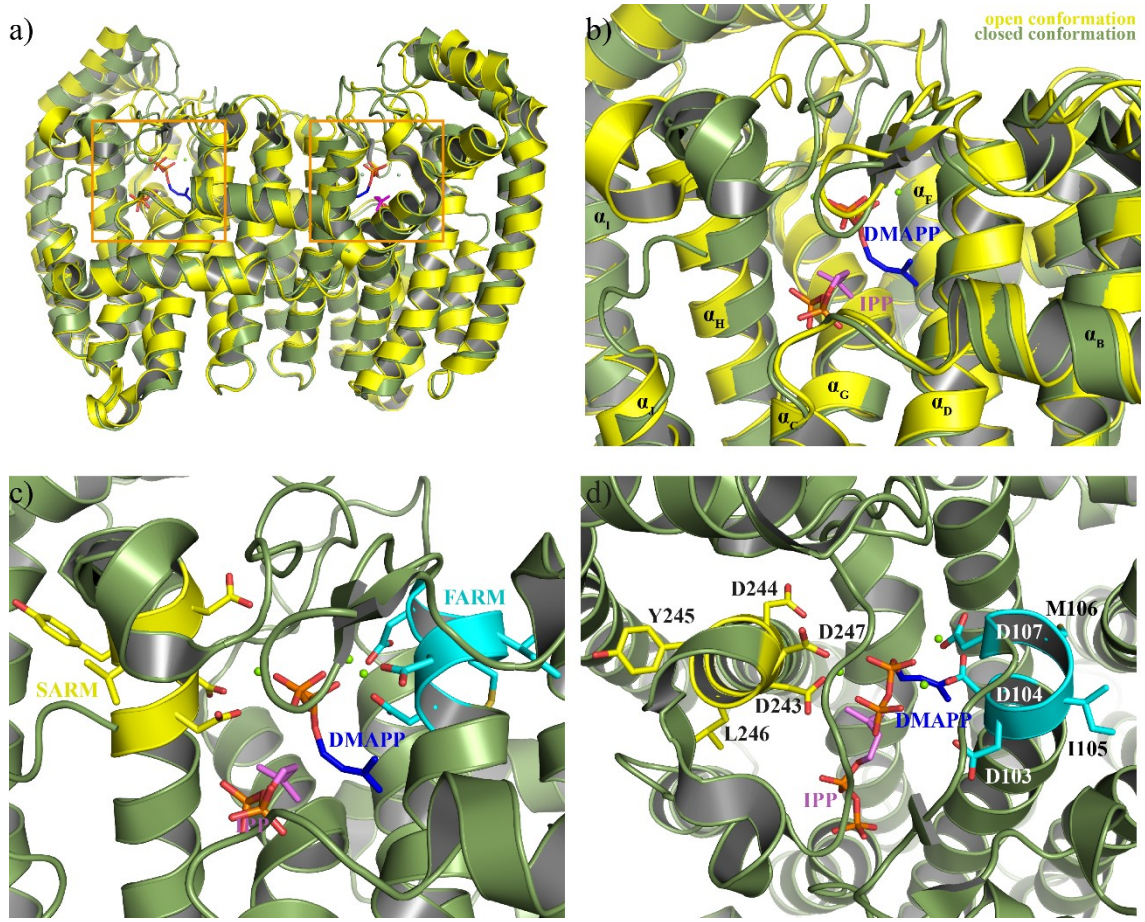


Figure 8: Location and amino acids of the active site on human FPPS. The human FPPS homodimer in open-conformation ((PDB ID 4XQR, yellow) and closed conformation (PDB ID 2F8Z, green) is depicted with IPP (PDB ID 2F8Z, pink) and DMAPP (blue, position modelled from PDB ID 1RQI). **(a)** Overall view of human FPPS homodimer with active site marked in orange. (b) Close-up on active site in open and closed conformation. **(c)** and **(d)** Side view and top view on the active site. The FARM is marked in cyan and the SARM is marked in yellow and for both motifs, residues are shown as stick representation. In **(c)** the FARM and SARM are labelled and in **(d)** the residues of both motifs are labelled.

FPPS plays a key role in the isoprenoid pathway of sterol biosynthesis and catalyses the two consecutive head-to-tail condensation reactions of IPP and DMAPP to form GPP and further GPP and IPP to release FPP. The catalysed reaction is described to follow a three-step mechanism consisting of ionisation, condensation and elimination (Figure 9).[237] First, DMAPP ionisation is facilitated by the enzyme's trinuclear Mg^{2+} cluster where all three Mg^{2+} participate in the salt bridge formation with the unesterified oxygens of the pyrophosphate. A carbenium intermediate with an allylic carbocation is formed that is stabilised by FPPS interactions and electrostatic interactions with the generated pyrophosphate. At the same time, the metal ions stabilise the negative charge of the pyrophosphate. The hydrophobic C_5 isoprenoid tail of the second substrate, IPP is in juxtaposition to the positive charge on the DMAPP and enables a quick condensation reaction with the nucleophilic double bond of IPP. For the elimination step, the non-metal-ligated DMAPP pyrophosphate oxygen serves as the catalytic base and is in the right position to abstract the proton from the IPP C_2-pro R to form the double bond of the final product. [227, 230, 238]

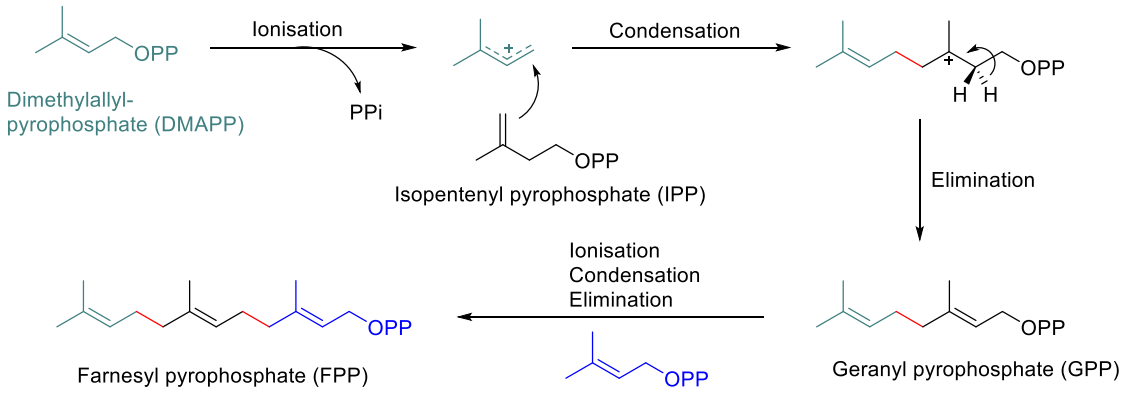


Figure 9: Chemical scheme of FPPS catalysis. The ionisation, condensation and elimination steps to form geranyl pyrophosphate (GPP) and farnesyl pyrophosphate (FPP) are depicted. Adapted from [239].

The FPPS structure and mechanism were originally studied based on crystal structures of *avian*, *Escherichia coli* (*E. coli*) and *Staphylococcus aureus* (*S. aureus*) FPPS and is now especially well studied in human FPPS.[226, 230-231, 234, 240] Unliganded human FPPS is found in an open conformation and during the catalytic reaction, the enzyme undergoes significant structural changes. Upon binding of DMAPP, the enzyme adopts a partially closed conformation where the two aspartate-rich motifs (FARM and SARM) get in closer proximity. The α_H-α_I and α_D-α_E loops close the access to the DMAPP binding site and the IPP binding site is formed. Then IPP binds and the C-terminus with its highly basic residues moves towards the IPP binding site to seal the active site and to adopt a fully closed conformation (Figure 8). This conformation properly locates the IPP and DMAPP substrates for chain elongation and shields the formed reactive carbocation intermediate from water. After catalysis, pyrophosphate is released, which destabilizes the closed conformation, and the enzyme transitions to the open conformation to release the C_{10}-isoprenoid and enable the transfer of GPP to the allylic site. In another catalytic cycle, a new IPP molecule binds and C_{15}-FPP is formed. [231]

1.2.4 Targeting human FPPS by bisphosphonates

The first synthesis of bisphosphonates (BPs) was reported in 1865[241] and bisphosphonates were introduced in the clinics in the late 1960s and early 1970s after it was shown that bisphosphonates inhibit the hydroxyapatite dissolution and further studies demonstrated inhibition of bone resorption.[242-243] Due to their high incorporation into sites of active bone remodelling and their fundamental properties to suppress bone resorption bisphosphonates are now in widespread clinical use for the therapeutic application in bone-related diseases, such as hypercalcemia of malignancy, Paget's disease, multiple myeloma and osteoporosis. [244-248]

Bisphosphonates are non-hydrolysable analogues of inorganic pyrophosphate with a replacement of the central oxygen to a methylene group and various side chains bound to the C_1 position.[114] Up to now, a great number of bisphosphonate analogues were synthesised and tested,

which can be divided into three major groups: the first generation of bisphosphonates, such as etidronate (Didronel®)[249], clodronate (Bonefos®, Clasteon®)[250], and tiludronate (Skelid®)[251], which do not possess a nitrogen group at the R2 position, the second generation bisphosphonates, which are aliphatic nitrogen-containing bisphosphonates (N-BPs), such as pamidronate (Aredia®)[252], alendronate (Fosamax®)[253], and ibandronate (Boniva®)[254] and the third generation, which contain a nitrogen heterocycle, such as zoledronate (Reclast®, Zometa®)[255], and risedronate (Actonel®, Atelvia®)[246-247, 256-257]. Remarkably, second generation aliphatic N-BPs showed a 10- to 100-fold increase in antiresorptive potency and heterocyclic N-BPs were found to be up to 10,000-fold more potent than the first generation bisphosphonates.[256, 258] The evolution, detailed investigation and SAR of different bisphosphonate compounds is described elsewhere.[225, 259-264] Some representative bisphosphonate structures are shown in Figure 10.

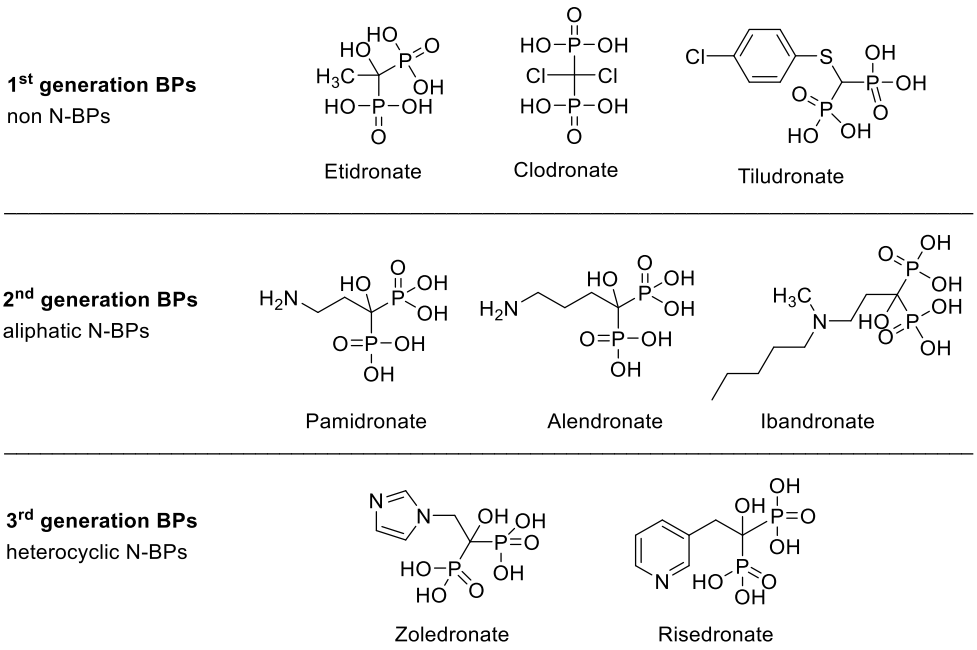


Figure 10: Representative 1st, 2nd and 3rd generation bisphosphonate structures.

Two different mechanisms of bisphosphonates were described. The first generation bisphosphonates are metabolised to non-hydrolysable β,γ-methylene containing (AppCp-type) analogues of adenosine triphosphate (ATP) by reversing the pyrophosphorylytic reactions of aminoacyl-transfer RNA synthetase and intracellular accumulation appears to provoke osteoclast apoptosis.[261, 265] In contrast, Luckman *et al.* proposed in the late 1990s that N-BPs target enzymes of the mevalonate pathway.[266] Further experiments suggested that osteoclastic bone resorption could be prevented by geranylgeraniol (GGOH) and in the same year, FPPS was described as the molecular target of N-BPs.[267-270]

Competition experiments and structural studies of FPPS complexed with N-BPs revealed that bisphosphonates bind to the GPP/DMAPP catalytic pocket.[230-231, 271] The phosphonate groups of

N-BPs are ligated through octahedral coordination of Mg^{2+} with the carboxyl groups of the aspartate-rich motifs, water molecules and the bisphosphonate oxygen atoms in the same manner as the pyrophosphate of the GPP/DMAPP.[231, 256] ITC experiments further showed the dependency of the binding on Mg^{2+} and a 30-fold increased risedronate affinity in the presence of Mg^{2+}.[232]

The interaction of bisphosphonates with the aspartate-rich motifs and the Lys257 of the α_H-α_I loop, which then have contact with the Asp243 and a phosphonate oxygen leads to structural rearrangements and the formation of the closed conformation (Figure 11).[231-232] The binding of the most potent N-BPs, risedronate and zoledronate, is almost identical and the protonated imidazole nitrogen atom of zoledronate interacts with the carbonyl oxygen of Lys200 and the hydroxyl oxygen of Thr201 at the helix α_G, which is part of the carbocation binding pocket.[231, 237] These binding interactions match with the proposed function of the cationic side chain of N-BPs to mimic the putative transition state allylic carbocation and to act as transition state analogues.[272] Many of the clinically used bisphosphonates have a hydroxyl group attached, which was shown to increase binding to calcium and create a tertiary binding interaction of the phosphate and hydroxyl group in bisphosphonates to the bone matrix and therefore, results in the extreme bone specificity.[247, 258]

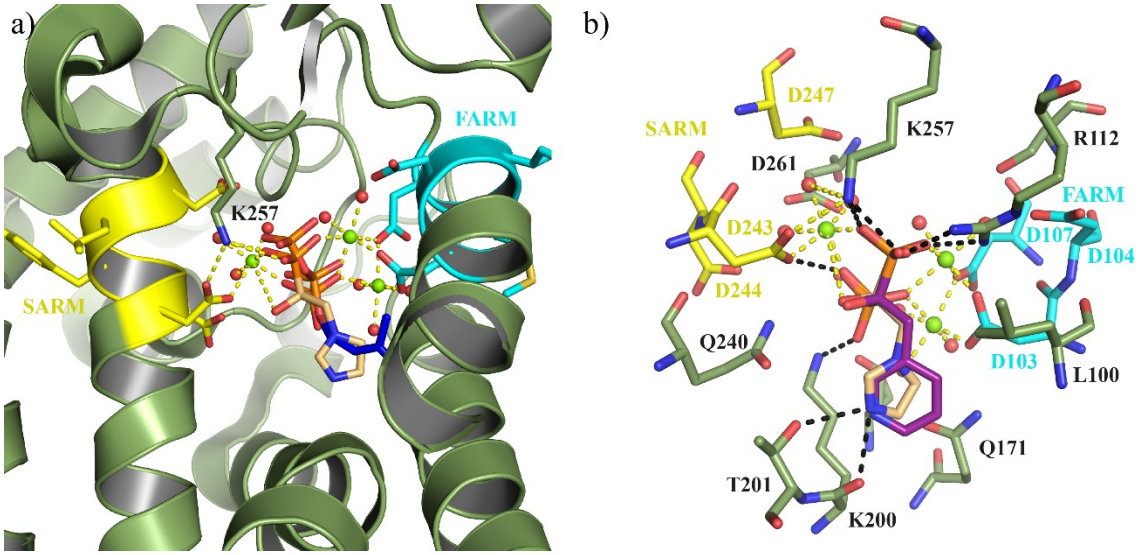


Figure 11: Bisphosphonates on human FPPS. (a) Human FPPS (PDB ID 2F8C) with zoledronate (yellow) and DMAPP (position modelled from PDB ID 1RQI, green) is shown. FARM (cyan), SARM (yellow) and K257 are labelled. **(b)** Bisphosphonate ligand binding site (PDB ID 2F8C) with zoledronate (yellow) and risedronate (violet) is depicted. Interactions of residues with zoledronate (blue dashed lines) and interactions of Mg^{2+} (yellow dashed lines) are shown.

The efficacy and clinical benefit of bisphosphonates were evaluated in various clinical trials. Bisphosphonates are administered either orally or intravenously as a standard of care for bone-related diseases.[247, 257, 273-275] However, apart from the beneficial effect on bones, emerging evidence has been collected, demonstrating that the bisphosphonate class of inhibitors has also direct and indirect anti-tumour effects. Thus, bisphosphonates are further explored for their potential anti-cancer activity.[276-279] Interestingly, bisphosphonates could also be a potential

therapeutic target for neurodegenerative diseases and bacterial diseases, and most importantly, bisphosphonates targeting FPPS have shown antiparasitic effects.[115, 280-281]

1.2.5 Bisphosphonates in antiparasitic infections

Urbina *et al.* describe one of the pioneering works for bisphosphonates in parasitic infections in 1999.[282] Pamidronate, alendronate and risedronate showed effective inhibition of *T. cruzi* amastigote growth *in vitro* with IC_{50} values of 65 µM, 65 µM and around 300 µM as assayed by uracil proliferation. Most importantly, no toxicity of bisphosphonates to host cells except for concentrations above 300 µM was observed and *in vivo* studies with pamidronate demonstrated a proof of concept.[282] In the early 2000s, Martin *et al.* published a landmark study on the screening of 19 bisphosphonates against *Leishmania donovani* (*L. donovani*), *Toxoplasma gondii* (*T. gondii)*, *T. brucei* , *T.cruzi* and *Plasmodium falciparum* (*P. falciparum*) and demonstrated low micromolar up to nanomolar inhibitory effects in all studied organisms.[115] Remarkably, in this study the *ortho*-isomer of risedronate showed an IC_{50} of 220 nM in BSF *T. brucei* and for the first time the effect of bisphosphonates on the sterol biosynthesis in *T. cruzi* was demonstrated.[115] The molecular target FPPS is now well-described and several *T. cruzi* FPPS-N-BP structures were solved to study the mechanism of action.[227, 283]

The major problem of bisphosphonates for the treatment of parasitic infections is their tight binding to bone mineral, leading to their rapid clearance from the bloodstream. Thus, to overcome this limitation, SARs were studied and modified bisphosphonates were extensively investigated. Bisphosphonates demonstrated *in vitro* and *in vivo* efficiency in a broad range of studies with *T. cruzi*, which shows their potential as an alternative therapeutic treatment of Chagas disease.[118, 284-287] Additionally, up to now, bisphosphonates were tested for antiparasitic effects *in vitro* and *in vivo* also in *Cryptosporidium parvum*[288-289], *Entamoeba histolytica*[290], *Leishmania* spp.[291-293], *T. gondii*[294-295] the cause of cryptosporidiosis, amebiasis, leishmaniasis, and toxoplasmosis respectively.

Major research efforts are also directed towards the discovery of new treatments for malaria, caused by *Plasmodium* species and responsible for 435 000 deaths globally in 2017.[296] In contrast to *T. brucei*, Plasmodium parasites depend on the 7-enzyme prokaryotic 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway to produce IPP and DMAPP and FPPS/GGPPS is a bifunctional enzyme that synthesises C_{15} and C_{20} isoprenoids. However, equally to other parasites several N-BPs have shown inhibition of FPPS/GGPPS with potent antimalarial activity.[117, 290, 297-298]

1.2.6 Targeting *T. brucei* FPPS

An IC_{50} of 220 nM for *ortho*-risedronate on BSF of *T. brucei* was first described in an *in vitro* growth inhibition assay as part of a study on various parasitic organisms as mentioned in Chapter 1.2.5.[115] *In vitro* rescue experiments with GPP, FPP and farnesol on *T. b. rhodesiense* after inhibition of growth with risedronate determined FPPS as the target and later on RNAi studies validated *T. brucei* FPPS as a potential drug target.[115-116]

The first crystal structure of *T. brucei* FPPS at 3.3 Å resolution for the apoprotein and at 2.5 Å in the presence of minodronate, a third generation N-BP, was reported by Mao *et al.* in 2004.[229] In the following years the structures of *T. brucei* FPPS in complex with several other bisphosphonates were solved at resolutions ranging from 1.94 Å to 2.69 Å and are stored as 16 PDB entries (5AEL[299], 5AFX[299], 5AHU[299], 4RXC[300], 4RXD[300], 4RXE[300], 4RYP[300], 3DYF[301], 3DYG[301], 3DYH[301], 3EFQ[302], 3EGT[301], 2P1C[303], 2OGD, 2EWG[304] and 2I19[304]) in the PDB.[305]

T. brucei FPPS has an overall 32 – 68 % identity with FPPS of other organisms.[194] More precisely, it shares a 64 % overall similarity and 44 % identity and an 89 % similarity in the catalytic pocket with human FPPS.[300] The enzyme is closest to *T. cruzi* FPPS with seven conserved regions and 13-α helices connected via interhelical loop regions.[194, 306] Ten major helices and three small helices that are characteristic for FPPS are formed. Of these helices α_B, α_D, α_E, α_F and α_G and the connecting loops between α_A-α_B, α_D-α_E and α_F-α_G are part of the homodimeric interface.[306] *T. brucei* and *T. cruzi* FPPS share an 11-mer insertion (residues 184 to 194 in *T. brucei*) on loop 2 (connecting helices α_F-α_G) and besides this, *T. brucei* FPPS also has an additional unique 6-mer insertion from residues 66 – 72 containing three asparagines on loop 1 (connecting helices α_C-α_D) (Figure 12).[194]

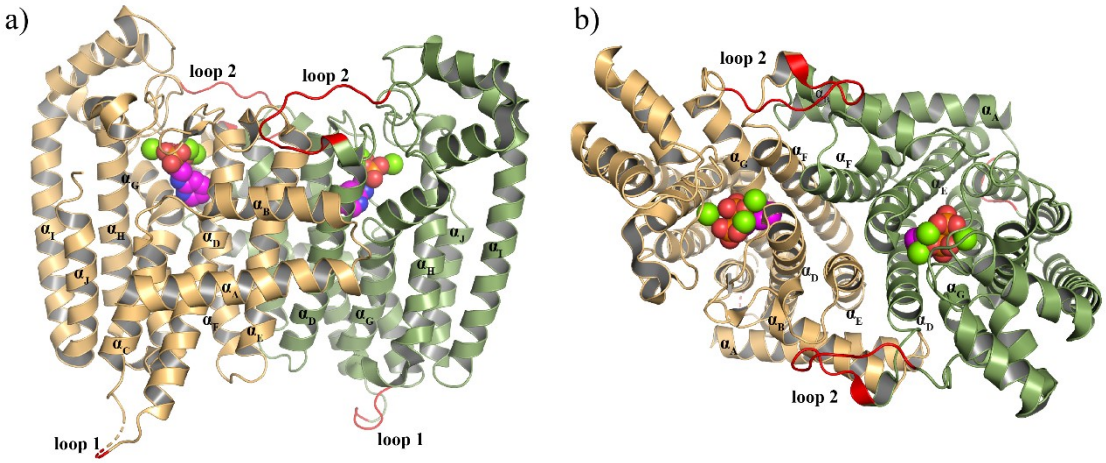


Figure 12: *T. brucei* FPPS homodimer showing loops. (a) Overall view of the *T. brucei* FPPS homodimer (PDB ID 2I19, bisphosphonate as CPK representation in magenta) with individual monomers in yellow and green. Loop 1 and loop 2 insertions are shown in red and visible helices are labelled. **(b)** Top view on FPPS homodimer. Helices at the homodimer interface are labelled.

In the catalytic site, two highly conserved aspartate-rich motifs FARM and SARM are found, which are on the C-terminal end of helix α_D as $D^{103}DIMD^{107}$ sequence and on the C-terminal end of helix α_H as $D^{255}DVMD^{259}$ sequence.[236, 306] These conserved sequences together with a conserved trinuclear cluster of magnesium ions are essential for enzyme catalysis of FPPS.[235, 307] When DMAPP is bound, the three Mg^{2+} interact with its pyrophosphate group to keep them in place. The carboxylate of Asp103 and the oxygen atom of Asp107 of the FARM coordinate directly to Mg^{2+}_A and Mg^{2+}_C. Asp255 of SARM coordinates to Mg^{2+}_B. The full octahedral geometry is completed via coordination of structural water molecules.[236, 306] Mg^{2+} is also essential in the catalysis, as the polarisation of the O-C bond on DMAPP via Mg^{2+} initiates the attack of IPP for GPP formation.[236] Apart from that, two conserved arginines C-terminal to the FARM (DDIMDXXXXRR) show implications in the catalytic activity as observed in mutational studies of *S. cerevisiae* FPPS.[306-307] Interestingly, the additional arginine on the third position of the C-terminal end of trypanosomatids FPPS was not observed to have any impact on catalysis.[306]

The amino acid sequences in the DMAPP binding site as well as the residues in the IPP binding site except for the *P. falciparum* FPPS are conserved across all protozoal FPPS enzymes.[306] Interacting residues of the alkyl part and the pyrophosphate moiety of IPP are thoroughly described by Tekwani *et al.* from a crystal structure (PDB ID 1YHM) of *T. cruzi* FPPS. The respective residues on *T. brucei* FPPS interacting with the alkyl head group are Leu100, Tyr216 and Phe251 whereas the pyrophosphate group interacts with polar mostly basic residues Asp50, Gln96, Arg113, Arg365, Lys367 and Lys47 (Figure 13 a). Lys367 is especially critical due to its interaction with both IPP phosphate groups.[306] Comparison of the amino acid sequence of protozoan FPPS also showed that Lys47, Lys212, Lys361, Lys364, and Lys367 are conserved throughout the protozoan FPPS enzymes with additional contributions to substrate binding.[226-227, 306] Two aromatic amino acid residues, His98 and Tyr99 (HYXXXDDIMD), N-terminal to the FARM are described to determine the maximum chain length of the final product catalysed by FPPS and mutation of the respective residues in avian FPPS led to >15-carbon long isoprenoids.[240, 306]

Bisphosphonates mimic the DMAPP substrate and bind to the allylic binding site on *T. brucei* FPPS as already generally described in Chapter 1.2.4. A potential steric effect of Tyr99 for longer isoprenoids was supported by a bisphosphonate-bound (Figure 14, compound **18**) *T. brucei* FPPS structure showing interaction of the phenyl ring with Tyr99.[303] Compound **18** showed hydrophobic interactions with Ala169, Met106, Thr168, Tyr99 and Thr213 and the long side chain fully occupied the allylic binding site (Figure 13 b). The two phosphonate groups closely resembled the position of zoledronate-bound human FPPS.[231, 303]

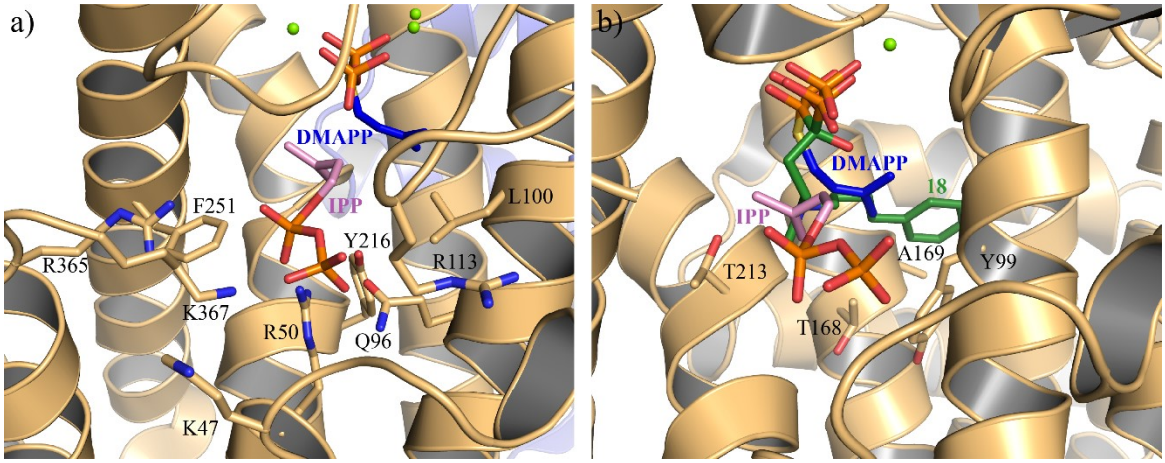


Figure 13: Active site of *T. brucei* FPPS. (a) *T. brucei* FPPS active site (PDB ID 2I19, bisphosphonate not shown) with DMAPP (blue) and IPP (pink). Position modelled from PDB ID 1RQI (protein backbone not shown). **(b)** *T. brucei* FPPS active site (PDB ID 2P1C) with bisphosphonate **18** in green. DMAPP (blue) and IPP (pink) are overlaid (position modelled from PDB ID 1RQ, protein backbone not shown).

Additionally, Yang *et al.* identified in a comprehensive screen of a 925 compound library containing 562 bisphosphonates, 56 bisphosphonates with effective *T. brucei* FPPS enzyme and cell growth inhibition.[299] The study also included imidazolium bisphosphonates with side chains of C_0 – C_{15} and an OH group or an H atom at the C_1 position (Figure 14, representative compounds of study **22**, **55** – **88, 91 – 92**). Interestingly, the most potent cell growth inhibition was achieved with C_7 – C_9 side chain lengths at which *T. brucei* FPPS inhibition was found to diminish. On one hand, this result can be explained by a better cell penetration of bisphosphonates with longer aliphatic chains; on the other hand, steric repulsion in the pocket might occur.[299] In crystallization studies of pyridinium bisphosphonates with different chain lengths (Figure 14, compounds **23** – **26**), a *T. brucei* FPPS-compound **25** complex structure (PDB ID 3EFQ) was solved that showed the side-chain terminus bent back as the compound did not fit due to Tyr99.[301]

Liu *et al.* solved an FPPS-bisphosphonate (Figure 14, compound **45**) complex structure that had two bisphosphonates bound in the active site and showed electron density for four coordinated Mg^{2+} instead of three.[300] The first molecule is bound in the allylic site interacting with three Mg^{2+} as in all other reported structures and the second one forms interaction with the third and fourth coordinated Mg^{2+}. The overall adopted conformation highly resembled the apo-state of the protein and did not induce a conformational change leading to a closure of the active site with the rotation of the helix α_2 (between helices α_I-α_H), a movement normally seen upon bisphosphonate binding.[300]

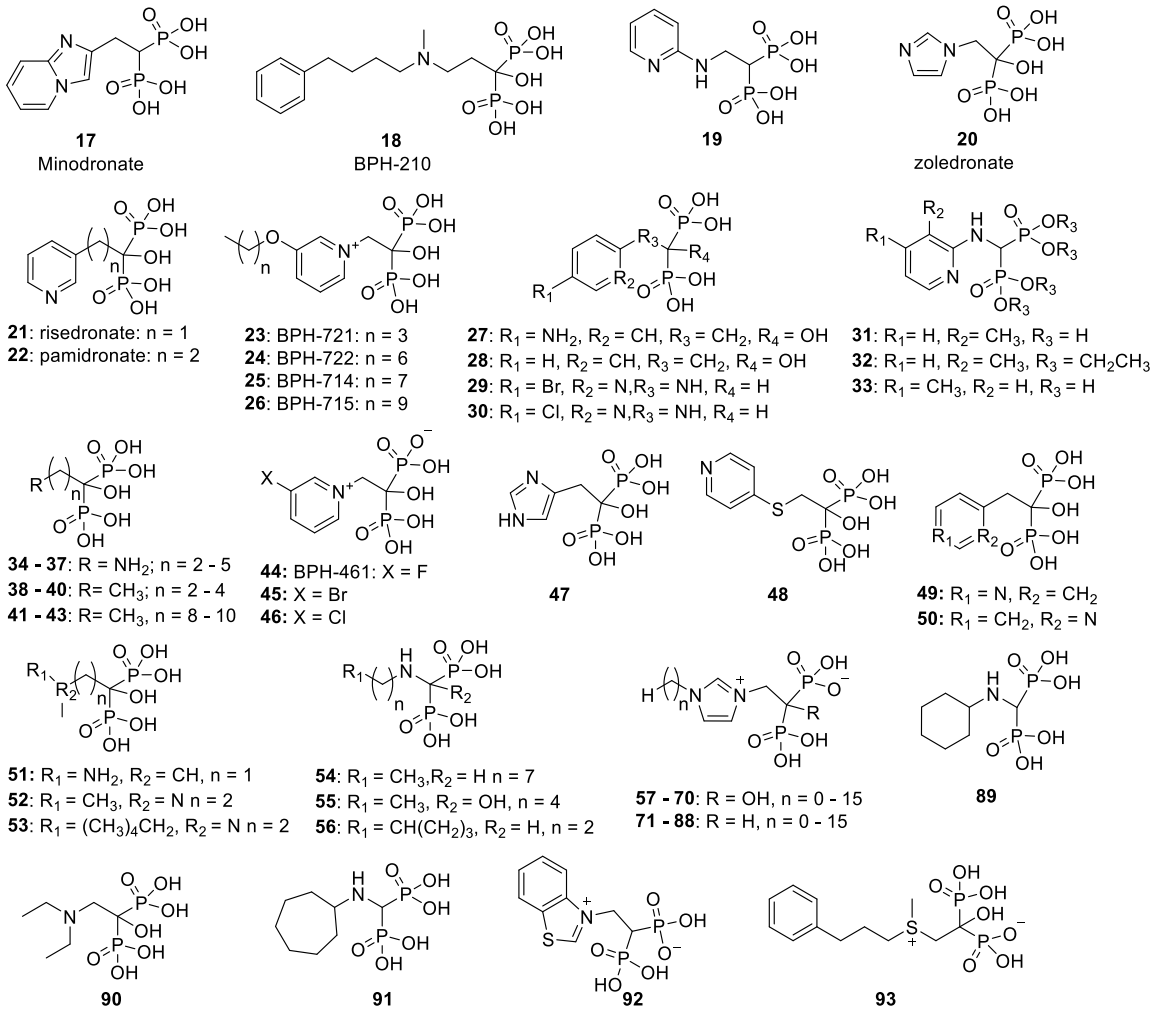


Figure 14: Chemical structures of bisphosphonates tested on *T. brucei* FPPS.

In the combination of two studies, the *in vitro* growth inhibition against BSF *T. brucei* of 31 bisphosphonates (Figure 14, compounds **21** – **22**, **27** – **43**, **47** – **48**, **50** – **56**, **89** – **90**) was determined and found to be in the range of 220 nM for *ortho*-risedronate to up to >> 100 µM for pamidronate.[115-116] Noteworthy, IC_{50} values < 1 µM were only observed for heterocyclic N-BPs. Compounds **21** and **31** have a therapeutic index (TI) of >200:1 together with a nanomolar IC_{50} and therefore, were of great perspective for additional investigation.[115-116] These two compounds together with four others (Figure 14, compounds **33** – **34**, **42** – **43**) were tested independently for *in vivo* FPPS inhibitory activity in *T. brucei* infected mice. The treatment of 10 mg/kg for 5 days resulted in toxicity but risedronate was also tested in different dosing schemes and two times 5 mg/kg per day for 5 days showed 60 % survival, which clearly shows that bisphosphonates can also inhibit parasitic proliferation *in vivo*.[194] Additional bisphosphonate analogues tested on *T. brucei* FPPS are described as part of a large-scale study on diverse bisphosphonate analogues such as phosphonium and sulfonium containing compounds and compound **93** (Figure 14) with an IC_{50} of 180 nM on *T. brucei* FPPS was identified.[308]

Apart from bisphosphonates, 16 different diamidines and related compounds (Figure 15, three representative diamidine structures, compounds **94** – **96**), based on the clinically used pentamidine, were also investigated for activity on *T. brucei* FPPS.[123] Diamidines are actually thought to target AT-rich DNA but cell growth inhibition and target effect did not correlate well and previous studies showed inhibition of *T. brucei* FPPS by diamidines.[300, 309] The IC_{50} values on *T. brucei* FPPS varied largely from the low nM to low mM range. However, the correlation of *T. brucei* FPPS inhibition in combination with the experimental data for k-DNA binding together with one computed descriptor showed good prediction for cell growth inhibition. These results present not only compounds with *T. brucei* FPPS activity but also the potential for multitarget inhibitors.[123]

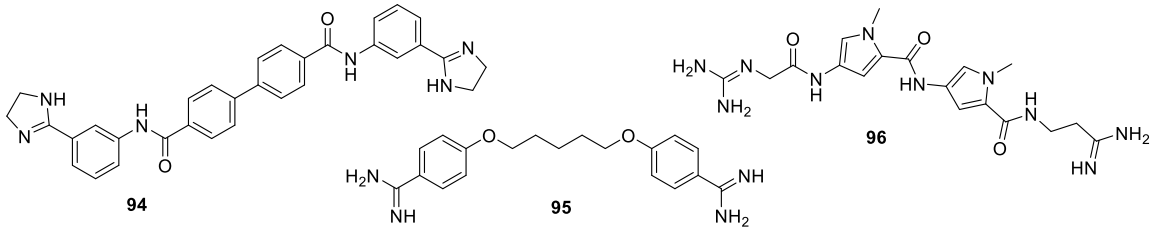


Figure 15: Three representative diamidines tested on *T. brucei* FPPS for inhibitory activity.

1.2.7 Identification of an allosteric site on human FPPS

In the search for new classes of non-bisphosphonate inhibitors without inherent bone affinity, Jahnke *et al.* discovered an allosteric site on human FPPS and new potent non-bisphosphonate inhibitors in a fragment-based approach.[310] An NMR fragment screen based on a 400-fragment library was performed and competition experiments using zoledronate showed that identified hits do not compete for binding to the active site. In X-ray crystallographic studies with FPPS-fragment binary and FPPS-fragment-zoledronate ternary complexes, the binding location was determined to be close to the C-terminus and near the IPP binding site. The identified allosteric site is formed by C-terminal helix α_J and three helices α_C,α_G and α_H. In some complexes further residues of the α_B-α_C, α_H-α_I and C-terminal loop and the N-terminal helix α_A also participate in the binding interaction (Figure 16). In the binary fragment bound complex the enzyme was found in the open state and in the ternary complex in the partially closed conformation. The partially closed state brings the α_H and α_J helices more adjacent to the α_C and α_G helices and reduces the size of the site.[310]

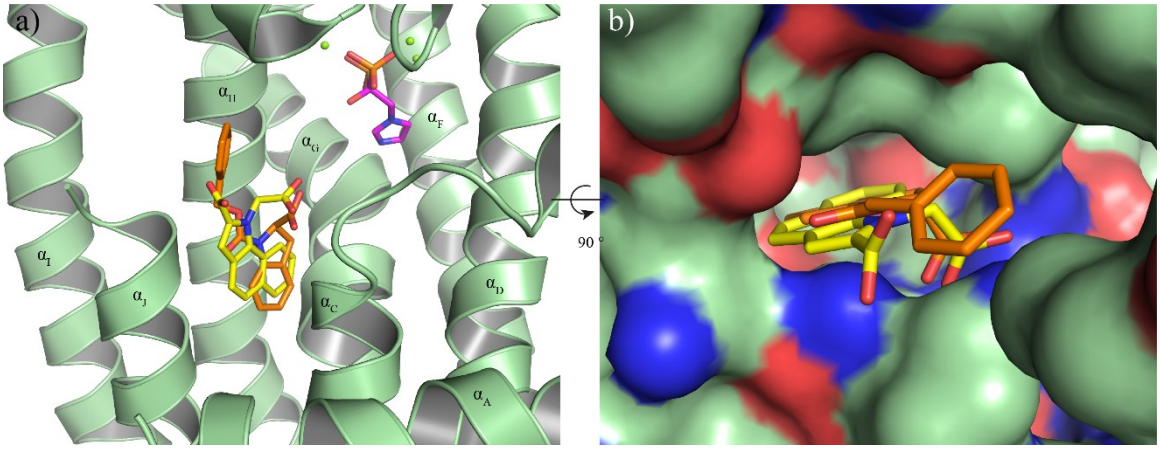


Figure 16: Allosteric site on human FPPS. (a) Close-up view on allosteric binders (PDB ID 3N46, 3N6K) with *T. brucei* FPPS protein backbone in green Cartoon representation (PDB ID 3N46) and helices around allosteric site labelled. Zoledronate (PDB ID 3N46) is shown in magenta. **(b)** Surface representation of *T. brucei* FPPS protein (PDB ID 3N46, green with nitrogens in blue and oxygen in red) and close-up top view on allosteric site with allosteric binders (PDB ID 3N46, 3N6K).

Initial benzothiopene and indole fragments were optimised with an NMR reporter assay[311] and further medicinal chemistry optimisation based on SAR studies using structural data together with results from a biochemical[312-313] and cellular assay[314] revealed two benzoindole derivatives (Figure 17**,** compounds **97** and **98**) with IC_{50} values of 200 nM and 80 nM.[310] Enzyme inhibition by allosteric ligands was found to be the result of a sterical hindrance to adopt the closed conformation required for IPP binding. Additionally, allosteric ligands repulsed electrostatically the IPP substrate and disrupted in some cases the enzyme dynamics necessary for catalysis.[231, 310] Most significantly, these compounds were devoid of measurable affinity in NMR binding assays with hydroxyapatite and bone powder.[315]

Based on this pioneering work, virtual screening for ligands to bind to the allosteric site, experimental follow-up by an inhibition assay and computational similarity search of an initial hit identified three bisamidines (Figure 17, compounds **99** – **101**) as most potent lead compounds with IC_{50} values of 1.8 µM, 1.9 µM and 2.5 µM, respectively.[309] The enzyme-ligand structure with compound **100** (Figure 17) revealed a new binding mode of this diamidine inhibitor with the central biphenyl unit remaining quite surface-exposed next to the C-terminal α_J helix. Furthermore, the imidazole moiety reached between the C-terminal α_J and α_C helix in a new site, which Liu *et al.* called S4.[300]

Furthermore, two new series of allosteric inhibitors for FPPS, which are based on quinoline, and salicylic acid derivatives were discovered by an *integrated lead finding* approach. This approach combined different screening techniques with SAR studies, biophysical methods for hit validation and the final hit-to-lead optimisation.[316] After co-crystallization studies of an initial NMR salicylic acid fragment hit, docking studies of a virtual library with 17 000 compounds guided the medicinal chemistry efforts to prepare several derivatives in parallel synthesis, which were further

tested in LC-MS-based biochemical assays.[316] A 4-acetamidophenyl derivate **102** (IC_{50}: 17 nM, Figure 17) and a 6-methoxy-napthyl derivate **103** (IC_{50}: 38 nM, Figure 17) were identified as lead compounds with an around 400-fold improved binding affinity to the parent compound. The quinoline lead series was based on an HTS hit with an indole-carboxylate/naphthyl scaffold identified in the allosteric site with a similar binding mode to the salicylic acid hit. In SAR studies guided by crystal structures and in LC-MS biochemical assays, two most potent quinoline derivates **104** and **105** (Figure 17) were synthesised with an IC_{50} of 24 nM and 37 nM. However, for both lead series a cellular plasma membrane translocation assay[314] revealed low cellular activity presumably as a result of poor cell permeability due to the carboxylic acid moiety and therefore, additional medicinal chemistry optimisations are required.[316] Furthermore, Lui *et al.* also investigated another quinoline derivate series with compounds **106** and **107** (Figure 17) showing IC_{50} values of 3.46 µM and 3.51 µM. The compounds were initially suggested by molecular docking studies to bind to an allosteric binding site.[317]

Interestingly, replacing the carboxylic acid by a monophosphonate functionality on the benzoindole, salicylic acid and quinoline series described previously, these compounds still bound to the allosteric site and the vinyl phosphonate quinoline derivate **108** (Figure 17) showed an IC_{50} value of 40 nM. This strategy demonstrated a way to fine-tune the bone affinity of potential drug candidates and has therefore the general potential to reduce side effects.[318] Furthermore, in the search for new targets for neurodegenerative diseases, Schutter *et al.* discovered the new chemotype of thienopyrimidine derivates and identified a thienopyrimidine monophosphonate **109** (Figure 17) binding to the allosteric site with an IC_{50} of 4.5 µM.[280] Additional thienopyrimidine bisphosphonates were found to exhibit a dual binding mode in the active site and allosteric site similarly to new indole-based bisphosphonate analogues.[319]

Remarkably, the role of this allosteric pocket was not elucidated until recently, when an X-ray structure in complex with FPP showed FPP binding to the allosteric site, which leaves the enzyme in an unreactive state.[320] Further kinetic analyses demonstrated FPPS inhibition by its own product binding, and therefore strongly suggest that FPP has an allosteric regulatory role. This product inhibition further provides a negative feedback mechanism for enzyme regulation with implications on the complete sterol biosynthesis.[320] In this thesis, only this binding site is called allosteric site although other identified binding sites could also be allosteric sites. However, all other sites are not further investigated and are called additional binding sites in this work.

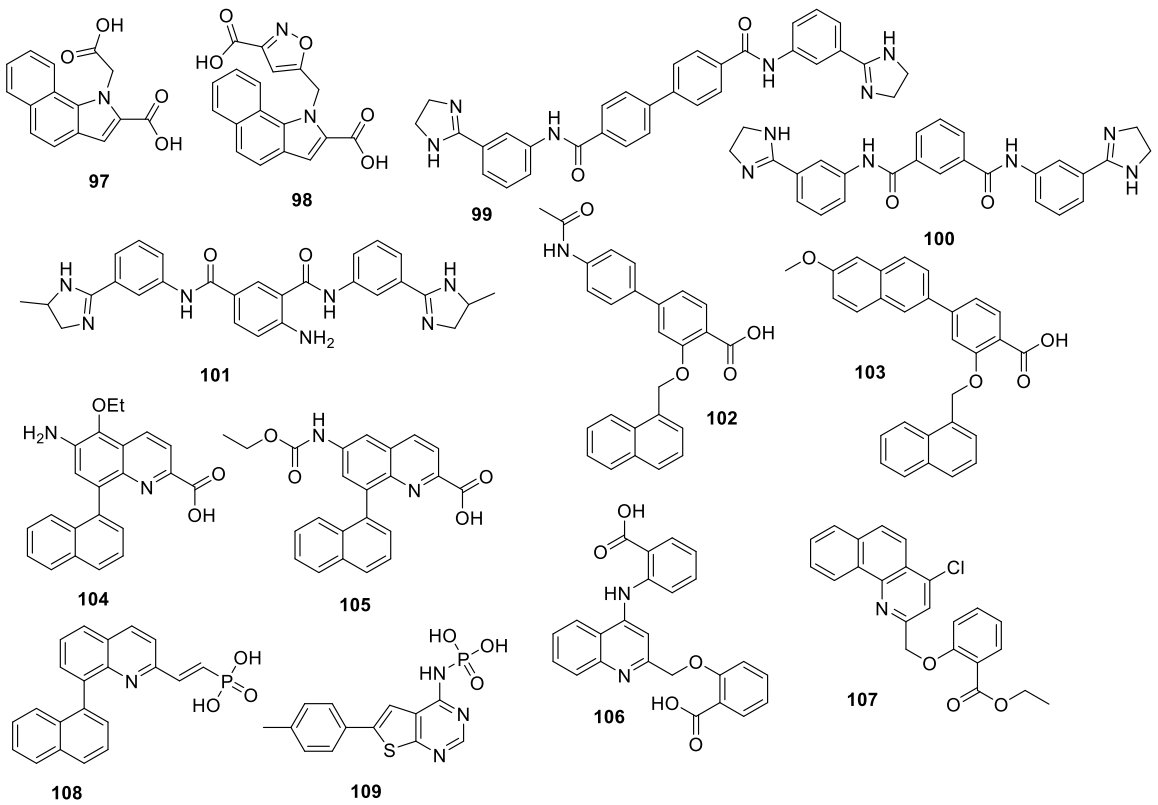
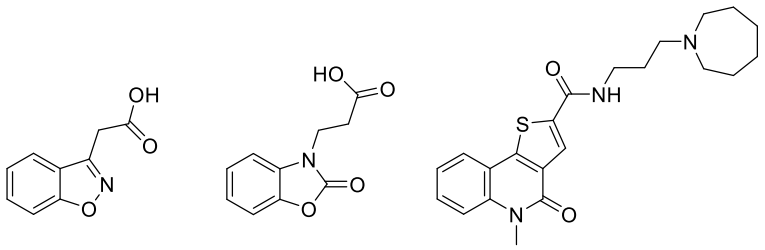


Figure 17: Allosteric inhibitors on human FPPS. Compounds **97-109** were identified in a previously unknown allosteric binding site on human FPPS with IC_{50} values ranging from 3.51 µM to 17 nM. Compounds **97 – 98**: benzoindole-derivatives. **99 – 101**: bisamidines. **102**: 4-acetamidophenyl derivate. **103**: 6-methoxy-naphtyl derivate. **104 – 108**: quinoline derivates. **109**: thienopyrimidine monophosphonate.

1.2.8 Compound binding in allosteric sites on FPPS in other organisms

In a fragment screen by differential scanning fluorimetry (DSF) on *Pseudomonas aeruginosa* FPPS, KM10833 and SPB02969 (compounds **110** and **111**, Figure 18) resulted in significant thermal shifts of 4.2 and 2.8 °C.[228] Crystallization studies of the binary FPPS-fragment and ternary FPPS-GPP-fragment complex revealed binding to the open conformation of the enzyme in two binding locations of an allosteric site adjacent to the catalytic site. Interestingly, this site has the same location as the allosteric site identified on human FPPS with a 30 % amino acid identity and an overall higher hydrophobic character.[228]



110 KM10833 **111** SPB02969 **112** MMV019313

Figure 18: ***Pseudomonas aeruginosa*** **FPPS and** ***Plasmodium vivax*** **FPPS/GGPPS allosteric binders.** KM10833 and SPB02969 (compounds **110** and **111**) were identified by DSF on *Pseudomonas aeruginosa* FPPS and MMV019313 (compound **112**) was identified by quantitative high throughput screen on *Plasmodium vivax* FPPS/GGPPS.

Recently, Gisselberg *et al.* published an interesting study for targeting a new site on the bifunctional FPPS/GGPPS enzyme in *Plasmodium* species.[298, 321] In a quantitative high throughput screen (qHTS) of the Malaria Box with 400 structurally diverse compounds, MMV019313 (compound **112**, Figure 18), a non-bisphosphonate inhibitor, was identified and described to bind to a non-substrate binding site of *P. vivax* FPPS/GGPPS.[321] Unfortunately, crystal structures were not obtained and the binding site could not be characterised. Additional compound optimisation of potency and pharmacokinetic properties is still needed before *in vivo* studies with *Plasmodium* infected mice can be performed. Nevertheless, these results show great potential for the development of a non-bisphosphonate drug with FPPS/GGPPS inhibitory effects.[321]

Both studies identified for the first time the binding of a non-bisphosphonate compound on *Pseudomonas aeruginosa* FPPS and *P. vivax* FPPS/GGPPS in a new binding site. These compounds could be exploited for the development of new lead candidates without inherited bone affinities.

1.3 Fragment-based drug discovery

1.3.1 Development of Fragment-based drug discovery

Fragment-based drug discovery (FBDD) has emerged as a widely applied and well-validated concept to explore small molecules binding to a target molecule, such as proteins or nucleic acids, based on various fragment-based screening methods as described in a plethora of reviews and books.[322-325] The earliest principle that supports the FBDD concept is that the Gibbs free energy of binding (ΔG^0) of a multivalent protein-ligand binding is the result of the contributions of the free energies of each interaction and the translational and rotational entropy.[326] Apart from that, computational methods inspired the initial ideas of what has developed into the FBDD approach. The GRID[327] method uses molecular probes to identify potential binding sites and the molecular dynamics approach MCSS[328] (Multiple Copy Simultaneous Search) maps functional group sites on the protein surface.[329] Many commonly known modelling programs to identify putative extension

points for functional groups to be used for fragment linking, such as CAVEAT[330], LUDI[331] and HOOK[332] have also been developed in the late 1990s.[325]

In 1996, Fesik's group at Abbott published their seminal article "SAR by NMR" in which a protein-observed nuclear magnetic resonance (NMR) approach on FK506 Binding Protein (FKBP) is described.[333] In this work, two identified fragments with 2 μM and 800 μM affinities in two different binding sites were successfully linked to boost potency and generate a 19 nM highly potent binder.[333] The concept stimulated great interest and was further extended. Abbott and Astex also used X-ray crystallography for fragment identification.[334-335] Additionally, Hann's publication on the interaction of receptors and ligands and the observation that with increasing complexity of ligands the number of valuable interactions decreases further supported the fragment-based screening concept.[336] Initially, mostly small companies, such as Sunesis and Vernalis explored the opportunities with success in the mid-2000s and only years later big pharma companies extended their portfolio with fragment-based platforms.[329, 337-338] Over the years, FBDD has evolved with the setup of improved fragment libraries, the establishment of new fragment screening techniques and fragment optimisation methods and the success is reflected in several fragment-to-lead optimisation programs, in two fragment-based molecules in clinics and 40 drugs in the clinical phase.[339-342]

1.3.2 Fragment library design

Fragments are molecules of low complexity generally following the "rule of 3" (Ro3) that was proposed in 2003[335], evolving from the Lipinski Rule of 5.[343] Lipinski's rule was formulated in 1997 to determine the drug-like properties and the likelihood of a chemical compound to make an orally available drug.[343] This rule was further developed for the application of fragment libraries to describe fragments with satisfactory properties for lead compound generation. A fragment hit generally has a molecular weight <300 Da, a clogP ≤3, less than three hydrogen acceptors and donors and less than three rotatable bonds.[335] However, the design principles of a fragment library have evolved over the last years with the implementation of new concepts to maximise the potential in terms of size and complexity as well as physicochemical properties of these libraries.[325, 344-347] The Ro3 is still widely accepted but is more considered as a rough guideline. The typical fragment size has shifted to smaller fragments ranging from 140 – 230 Da with up to 18 heavy atoms.[345, 347-348] The small size is very useful for probing new binding pockets on a target molecule but fragments need to have a detectable affinity and specificity and therefore, a minimal fragment size containing 5 – 10 heavy atoms is desired as a compromise to fulfil both requirements.[345, 348] Fragments typically show binding affinities in the mM to μM range and it was widely shown that specificity and affinity is highly improved by directional polar interactions whereas increasing lipophilicity to improve affinity rather leads to nonspecific binding with target promiscuity.[346, 349-350]

Fragment libraries should also cover a relevant portion of chemical space comprising fragments of appropriate complexity and pharmacophores for target interaction.[345] The complexity model of Hann *et al.* that initially supported the development of the FBDD framework was further extended and validated and still finds broad application.[336, 351-352] It describes a higher probability of binding and detection of a molecule with less complexity and therefore, less ligand-receptor interactions that have to be correct.[336] Therefore, the simplicity of fragments allows setting up fragment libraries with a few thousand fragments or even less to explore a large chemical space and the probability of identifying hits is sufficiently high. Fragment libraries in practical use mostly contain 1000-5000 fragments.[353-354] By contrast, high-throughput screening (HTS) libraries contain hundreds of thousands of compounds and the identification of a binding event is drastically decreased due to the higher complexity of the library compounds.[345] Not surprisingly, FBDD has even shown to identify fragment hits on challenging and unprecedented targets, such as the disruption of protein-protein interaction, when HTS campaigns already failed.[353, 355]

In a recent review by Keseru *et al*., the hit rates against multiple targets obtained at different pharma companies are reported and vary dramatically from 2 – 26 %. The hit rate is also largely dependent on the applied method and the defined cut-off value for hits, but in any case a minimum of 0.2 % hit rate is observed.[356] In this aspect, the overall shape of fragments is highly discussed. It is speculated that a higher number of three-dimensional fragments reduces the promiscuity of binding events but similarly, the higher complexity of the fragments also leads to reduced hit identification.[357] Nevertheless, three-dimensional fragments extend the covered chemical space and can show potential for challenging targets.[358] Other researchers argue for target-specific libraries and have shown their success.[345, 359] However, preconceived bias can limit the chance for serendipitous discovery of new binding pockets and scaffolds.[345, 359]

A major consideration is that fragments have easily accessible functional groups that enable a rapid optimisation from fragment-to-lead but do not contain undesired decorations.[360] Therefore, fragment libraries have continuously been improved by e.g. excluding fragments with reactive functional groups, chelators, aggregators and so-called pan-assay interference compounds (PAINS), which potentially interfere with the assay or even attach covalently to the target.[83] For example, this is particularly crucial for fragment cocktail screening commonly performed by NMR. If in fragment mixtures of eight compounds one in eight compounds distorts the screening result, the full screening of the entire library is not trustworthy.[361] Therefore, it is also crucial to have a library with fragments of acceptable purity desirably above 95 % and a continuous quality control of fragment stocks is essential.[362] Most commonly, fragments are stored as DMSO stocks and a study of 7200 compounds at room temperature over a period of more than a year showed that after one year 48 % of compounds had degraded.[362]

Over the years, chemical suppliers[363] have developed a portfolio of fragment libraries and also pharma companies and academic institutes, such as Novartis[360], Vernalis[346], Astex[357],

AstraZeneca[364-365], Pfizer[366] and the University of Leeds[367-368] have designed their own libraries for routine fragment screening.

1.3.3 Fragment screening methods

The identification of fragments requires a robust and sensitive screening method allowing to measure fragments at high concentrations without interfering with the assay. Fragment screening campaigns are dominated by biophysical techniques that can identify fragment molecules with affinities up to the low mM range.[329] However, biochemical methods such as high concentration biochemical assays or computational methods have also been successfully applied.[369-371]

Since the first key publication of Fesik *et al.* using protein-observed NMR techniques, several other screening methods, such as X-ray crystallography[372], surface plasmon resonance (SPR)[373-374], microscale thermophoresis (MST)[375] and thermal shift analysis (TSA)[376] are commonly applied in pharmaceutical companies and academic institutes. It is difficult to argue which technique is the "best" for fragment screening but at the end of the debate, it is inevitable that each technique has its own strengths and limitations and the most suitable screening technique needs to be selected depending on the target and available information.[329, 348, 353] In fact, a poll on the Practical Fragments blog in 2016 has shown that FBDD practitioners on average use 4.1 techniques for primary fragment screening and identified hits are further validated with one or more orthogonal methods.[377] In Figure 19, an overview of various screening methods is given with their individual measurable affinity ranges, advantages and disadvantages.

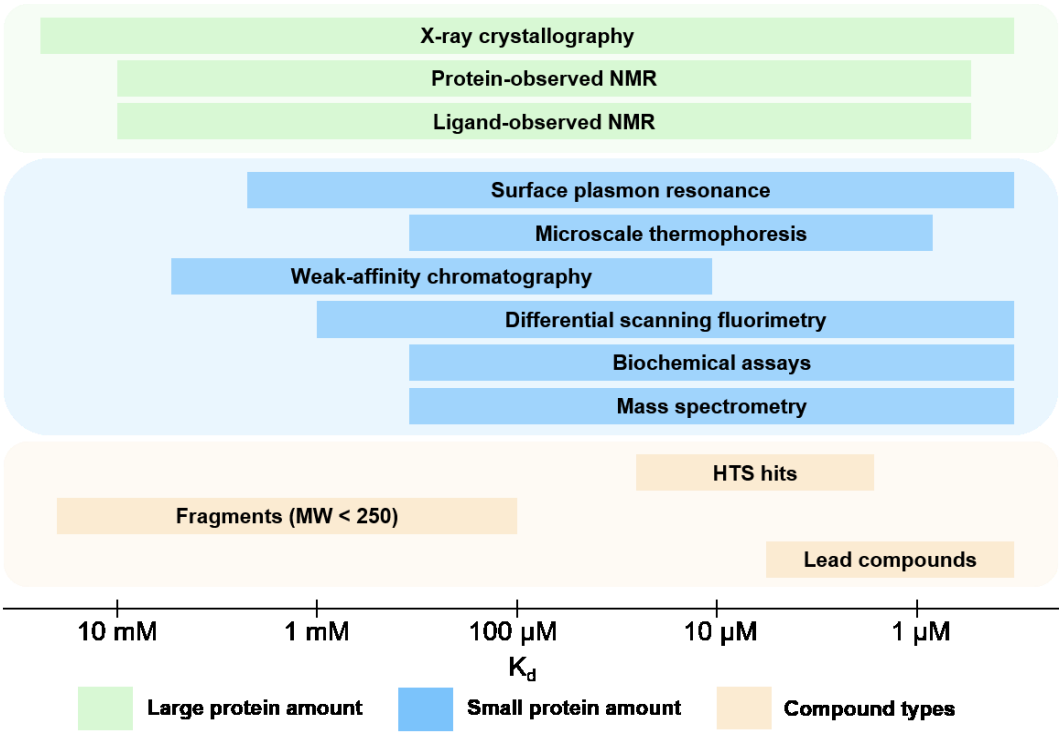


Figure 19: Sensitivity of different screening techniques Various fragment-screening techniques are shown with their associated sensitivity. Techniques are grouped by protein amount needed in large (green) and small (blue). As a reference, affinity of different compound types is depicted in orange. Adapted from [378-379].

1.3.3.1 Ligand-observed NMR

Ligand-observed NMR based on relaxation effects has evolved as one of the most predominant primary fragment screening methods.[329] The technique exhibits several advantages as a fragment screening method: the target molecule, generally a protein, does not require any previous modification and a low protein amount is needed. Furthermore, the binding interaction is measured in solution and the protein stability and solubility can be evaluated from the NMR spectra.[380] However, the method is restricted to targets above 20 kDa. Protein solubility at 10 µM and stability during the NMR measurement at room temperature is required.[329, 380] Additionally, due to their slow-exchange, ligand-protein interactions of high-affinity ligands are not detectable and appear to suggest false negatives. Furthermore, no information about the binding location can be obtained, as would be the case in protein-observed heteronuclear single quantum correlation spectroscopy (HSQC) based experiments.[329, 348, 380-381]

The method is typically considered as a medium-throughput method but the screening of fragment mixtures with generally eight compounds drastically decreases measurement time and allows high-throughput detection in fragment screening.[380] However, the screening of fragment mixtures also needs a precise selection of compounds for a mixture to avoid chemical reactions between fragments. Furthermore, the analysis of compound spectra to create fragment mixtures with compounds that give distinct NMR signals to allow unambiguous discrimination of binding events.[382] For this purpose, visual inspection is required but also different software was developed,

such as a computer-aided design method by Arroyo *et al*. and NMRmix to assist in creating NMR compound mixtures.[383-384]

The most popular NMR ligand-based screening experiments are saturation transfer difference spectroscopy (STD)[385] and water-ligand-observed via gradient spectroscopy (waterLOGSY)[386-387], which are both NOE-based experiments.[388] For either method, the ligand is in about 20-fold excess to the sub-stoichiometric concentration of the target molecule and fragment binders with affinities ranging from µM to mM can be detected.[388] WaterLOGSY experiments[387] are based on bulk water magnetisation transfer via spin diffusion and intermolecular NOE.[386] Bulk water magnetisation from water residues at the fragment-binding interface is transferred to the binding fragment via the slowly tumbling protein-fragment complex exhibiting an intermolecular NOE on the fragment with the same sign as the original water magnetisation. Then again, a small negative signal yields from the migration of water magnetisation to the non-binding fragment due to fast relaxation. [348, 386-387, 389]

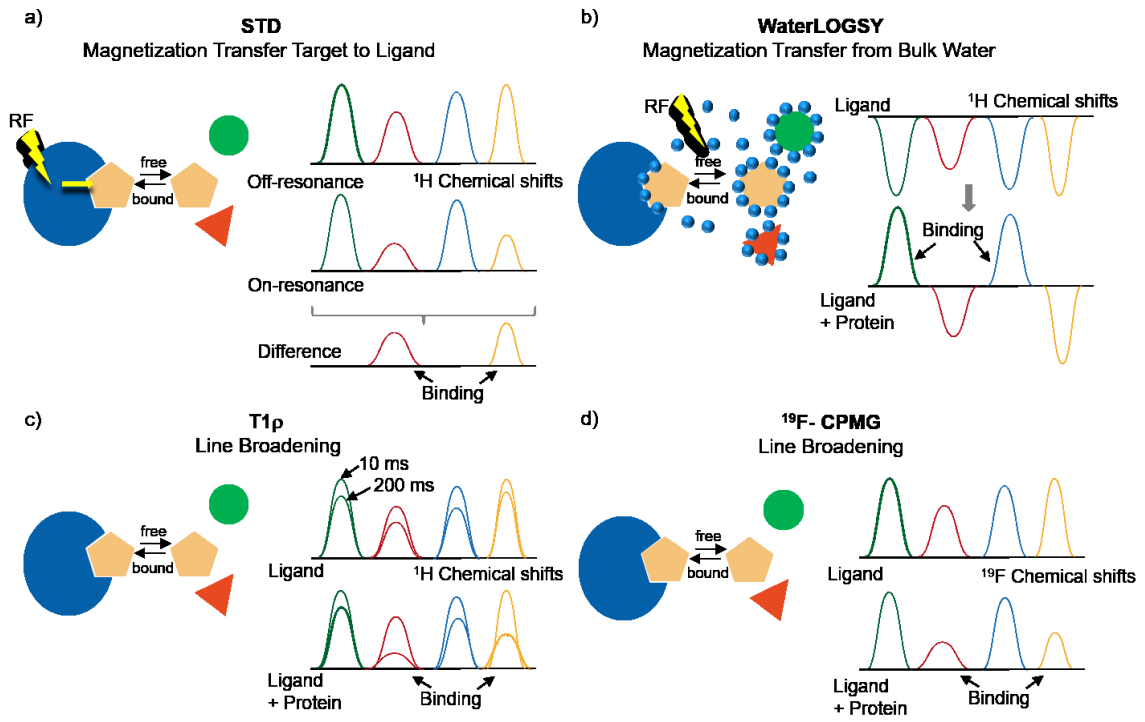


Figure 20: Schematic representation of different ligand-observed NMR screening methods. The blue circle represents the target protein and the pentagon, triangle and circle indicate different compounds. **(a)** Saturation transfer difference (STD) experiment showing an on-resonance, off-resonance and difference spectrum. The yellow lightning indicates radio frequency (RF) irradiation for selective target protein saturation. **(b)** Water-ligand obversed via gradient spectroscopy (WaterLOGSY). The blue dots represent water molecules the yellow lightning indicates RF irradiation for selective saturation of water signal. **(c)** ^{1}H-T1ρ relaxation-filtered experiment (T1ρ). Only ligand spectrum and ligand and protein spectrum with overlay of 10 ms and 200 ms spin lock durations. **(d)** ^{19}F- Carr-Purcell Meibom Gill (CPMG) experiment. ^{19}F chemical shifts spectra of only ligand and ligand with protein. Adapted from [363].

STD experiment[385] is based on recording an on-resonance spectrum with selective protein saturation and an off-resonance spectrum without protein saturation. The on-resonance spectrum is

acquired by selective saturation of a subset of the protein protons through selective inversion pulses, which leads to saturation of the bound ligand via spin diffusion through the Nuclear Overhauser effect. In contrast, free ligands in solution do not receive any saturation transfer and attenuation of the signal resulting in a high-intensity signal. The off-resonance spectrum is acquired with selective saturation pulses off the protein and ligand resonances and high-intensity signals will be recorded for bound and unbound ligands. In the difference spectrum subtracting the "off" from the "on" resonance spectrum only protons that interact with the protein remain and thus, binding fragments are distinguished from non-binders.[348, 380, 390]

Alternatively, a ^{1}H-T1ρ relaxation-filtered experiment[391] can be acquired using the differential relaxation rates of large slowly tumbling molecules, such as proteins and short relaxation times of fragments. The ^{1}H-T1ρ experiment uses spin lock sequences to suppress signals from slowly tumbling molecules. Two different experiments with short 10 ms and long 200 ms spin-lock durations are acquired and overlaid. For binding fragments, long spin lock sequences result in line broadening and smaller signals due to fast relaxation of the protein-fragment complex compared to free fragments in solution. Therefore, binders can be distinguished from non-binders by monitoring the line width and intensity changes.[348, 380, 392]

Generally, a combination of different ligand-observed experiments are recorded on the same sample as it is relatively quick and binding fragments identified in several NMR-binding experiments have a higher probability to be validated also in other biophysical methods. [348, 381]

In addition to the fragment mixture screens described above, in recent years, the potential of fluorine libraries containing CF_3, CF_2 and CF groups has been increasingly recognized leading to the specific setup of fluorine libraries.[393-394] Fluorinated fragment libraries are generally screened by a T_2 pulse sequence[395-396] assessing fragment binding by line width broadening. ^{19}F-NMR spectroscopy gives simplified spectra due to a large chemical shift range based on a strong chemical shift anisotropy and the observation of generally only a single peak for each fragment.[380, 397] The significant line width difference between free fragment in solution and fragment in protein-bound state enables the use of five times lower protein concentrations than in ^{1}H-NMR experiments and even very weakly binding fragments can be detected.[380] Therefore, fragment mixtures of up to 32 fragments can be screened highly increasing the throughput of the method.[380] Initially, concerns were put forward that with solely fluorinated fragments only a limited chemical space is screened and therefore lower hit rates would be observed. However, Jordan *et al.* showed in their publication on the practical use of ^{19}F-NMR spectroscopy in fragment-based screening (FBS) that the impact on the success of a screen is negligible.[398]

In addition to the screening methods described, if a high-affinity binder for the protein of interest is available, competition experiments by ligand-observed methods can be performed to explore the binding site of hits.[399] Competition experiments are especially of high interest if for

instance, an orthosteric binder is available and the identification of fragment binders only in new binding sites is desired. The addition of a known binder blocks the specific binding site and therefore, identified fragment hits are bound in other sites. Additionally, the limitations of NMR screening to detect high-affinity binders can be overcome by an NMR reporter screen with a medium affinity binder (K_d > 10 μM). In contrast to the detection of binders in new sites, in this case, high-affinity binders in the orthosteric site can be detected through the recognition of the displacement of the reporter by the test compound.[400] The competition approach was also further extended to the use of a fluorinated so-called spy molecule with the advantage of higher sensitivity and therefore also the feasibility of screening with a weak-affinity spy molecule and high compound mixtures as described by the labelled fluorine chemical shift anisotropy and exchange for screening (FAXS) method.[401-402]

The processing and analysis of the large amount of NMR fragment screening data can be tedious but the MNova Screen software recently developed by MestreLab Research in collaboration with NMR experts at Novartis simplifies and accelerates the process.[403] Hits identified in fragment mixture experiments need to be validated in single compound NMR experiments to exclude binding events observed due to mixture effects. Since any method comes with its own limitations as thoroughly summarized by Davis *et al.*[404] after confirmation of a hit in ligand-observed NMR fragment hits further need to be validated in an orthogonal method.

1.3.3.2 Protein-observed NMR

Protein-observed NMR is an information-rich NMR method that was used in the original FBDD approach and is still a widely applied approach for fragment screening.[348, 405] The method not only provides information on fragment binding but also allows binding site mapping, revealing which specific residues are part of the binding event.[329, 348] For this information, a backbone amide resonance assignment is required, obtained from triple-resonance NMR experiments.[406-407] However, even if a backbone assignment is not available or possible, clusters of perturbed peaks provoked by different fragment binders can allow the identification of specific or at least overlapping binding sites.[348] The requirement of large amounts of isotopically labelled protein for fragment screening can pose challenges for protein-observed NMR but nowadays expression of labelled protein is not limited to *E. coli* anymore.[380] Labelling protocols are available for insect cells, yeast and mammalian cells, among others, enabling a broader range of isotopically labelled proteins.[408-409] Generally, ^{15}N-labelled protein with an upper size limit of 30 kDa is used for fragment screening but ^{13}C-labelling can extend the application to targets up to 100 kDa.[380] For much larger proteins, ^{2}H-labelling and selective amino acid labelling can be applied and recently selective ^{13}C labelling at the methyl groups of amino acids by the ILV labelling approach was introduced for the NMR based screening approaches.[380, 410]

The predominantly applied 2D screening technique is [^{15}N,^{1}H]-HSQC[409-411] in which for each backbone amide one signal is observed as a result of magnetisation transfer between ^{1}H and ^{15}N nuclei.[348] The main limitation of the method is that it is time-consuming and especially proteins larger than 30 kDa need long acquisition times or high amounts of protein.[380] Alternatively, [^{13}C,^{1}H]-HSQC or [^{15}N,^{1}H]-transverse relaxation-optimised spectroscopy (TROSY) experiments can yield significantly improved sensitivity for large proteins and fast pulsing experiments, such as SOFAST-HMQC[411] can speed up the acquisition.[348, 361]

For protein observed approaches, just as for ligand-observed approaches, compound mixtures with 8 – 12 compounds are measured, which could increase the throughput but require follow-up deconvolution to identify which ligand has bound.[380, 388] A 2D-HSQC-NMR spectrum in presence and absence of fragments is performed. If a fragment binds, the local chemical environment of a residue changes, observed by chemical shift perturbations (CSP).[412-413]

In contrast to ligand-observed NMR methods, protein-observed approaches do not show limitations on binding kinetics.[361, 413] The dissociation constant K_d of a reversibly binding ligand to a protein is given by P + L ↔ PL and K_d further equals for a simple 1:1 reaction to k_{off}/k_{on} (back reaction rate constant divided by the forward reaction constant) at the equilibrium. The fragment can be titrated to the protein and depending on the binding regime slow, intermediate and fast exchange binding is distinguished.[413] For slow exchange rates, over a titration experiment, the free signal disappears and the bound signal appears with the intensities representing the respective concentration of free and bound protein. In contrast, for fast exchange rates, the shift of the signal from the free position to the bound position is followed and one single peak is observed at the weighted average shift of the free and bound state signals. In an intermediate exchange rate, the signal broadens and shifts. By analysis of the chemical shifts of the peaks, the K_d can be determined for binders in the fast exchange regime.[413] Compound titration experiments by NMR even allow K_d determinations in the mM range if the compound solubility allows it.[348, 413] Here, it is important to mention that especially but not exclusively at high concentrations of fragments, CSP need to be confirmed to originate from fragment binding and not due to high concentrations of DMSO or pH changes in the sample.[348]

1.3.3.3 X-ray crystallography

Ringe *et al.* who emphasised the application of organic solvents to identify energetically relevant hot spots already described X-ray crystallography with low molecular weight ligands.[414-415] Coincidentally, one of the earliest X-ray fragment soaking experiments was on the *T. brucei* triose-phosphate isomerase described in a book chapter in 1997.[416] In the early 2000s, increasing interest in the approach led to the setup of X-ray screening facilities in various pharmaceutical laboratories. It is now routinely integrated in the FBDD workflow to obtain three-

dimensional structural details of protein-fragment interactions.[334, 417-418] In fact, Schiebel *et al.* showed in a comprehensive study comparing six biophysical methods with X-ray crystallography the high value of primary X-ray screen for fragment identification.[419] Fourty-four percent of X-ray hits on endothiapepsin would have been missed with any other technique.[419]

Two of the main strengths of X-ray crystallography are its high sensitivity with fragment detection up mM range, which would exceed the limits of other techniques by far, and obtaining detailed structural information about the fragment's binding mode.[348] Therefore, it was even possible to develop a method using compounds smaller than fragments, so called Minifrags for X-ray experiments.[420] If another method than X-ray screening is performed still 34 % of researchers responded in a poll that they would not proceed with a fragment without a crystal structure confirming its binding and spatial location.[421]

Nevertheless, X-ray crystallography only provides a crystallographic model of the protein in the crystalline state, which can result in the identification of fragment binding at crystal contacts that are not biologically relevant.[422 372] Additionally, crystal contacts can block binding sites or the required conformational change disrupts the crystal lattice and electron density from crystallization condition components can be misinterpreted as the ligand.[372, 404] Furthermore, the technique does not provide any binding affinity information, which needs to be obtained by other experimental techniques for which some of the very weak shifters can reach the detection limit.[372, 404] However, most importantly highly diffracting crystals need to be obtained ideally beyond a 2.5 Å diffraction limit. This process can be extremely time-consuming as discussed below and does not always lead to success.[423]

X-ray fragment screening is most commonly performed by X-ray fragment soaking experiments, which includes the growth of a crystal and the soaking of the crystal in a solution containing a single fragment or a fragment cocktail as first described by Hartshom *et al.*[418] For that, highly reproducible protein crystals for rapid and efficient fragment soaking experiments have to be developed. However, crystallization is empirically founded and we do not have a comprehensive understanding of the requirements for crystal growth. Thus, it is fundamentally a trial-and-error approach and requires patience, persistence, consistency and a bit of luck. Nevertheless, since the first accidental discovery of a protein crystal more than 150 years ago[424] and the first published protein structure at 6 Å of myoglobin[425], the scientific community has gained extensive experience, introducing rational and systematic approaches on the basis of understanding the fundamentals of the system to most efficiently screen the various parameters influencing protein crystallization.[426-427] For that purpose, various crystallization screens are available to explore efficiently a diverse set of crystallization conditions.[428-429] Most commonly, crystallization plates for the sitting drop or hanging drop vapour diffusion method are employed. Moreover, crystallization experiments are performed in a highly controlled environment including temperature and humidity control among others. A broad range of excellent books and reviews is available

describing various optimisation strategies, such as variation of buffer pH, temperature, precipitant concentration and type, using macroseeding technique[430-431] or additive screens[432-434] containing alcohols, ions and sugars.[426, 435-437]

Crystals are built up by the periodic array of molecules and the smallest entity of the crystal that needs to be translated in all three dimensions to form the crystal is defined as the unit cell. A unit cell is described by six parameters, the three axes a, b and c and the three interaxial angles α, β and γ.[438-440] A crystal is defined by the infinite repetition of a unit cell along three principal directions and the shape allows to group into seven different crystal systems (triclinic, monoclinic, trigonal, hexagonal, orthorhombic, tetragonal and cubic).[440] These are further extended to 14 Bravais lattices that are obtained from the combination of the seven lattices with an additional centering component (primitive (P), face (F), body (B), base (A, B or C) and rhombohedral (R)). Further including symmetry operators which have to be applied to the so-called asymmetric unit of the unit cell (motif by which the entire crystal packing can be formed) results in the description of 230 space groups. However, only 65 are enantiomorphic and allowed for proteins.[438-439, 441]

The regular pattern of a crystalline structure can also be described as a series of planes. Consequently, an X-ray beam of a fixed wavelength striking on the crystal will be diffracted at a specific angle on the different planes. Constructive interference will take place when diffracted beams are in phase based on a different path equal to an integer number of the wavelength.[440] The geometric relationship between the wavelength of the X-ray beam, the distance between two planes and the angle of reflection is described in Bragg's law as $n\lambda = 2d \sin\theta$ and was first recognized by William Henry Bragg and William Lawrence Bragg in 1912.[442-443] In the equation n is the integer number, λ is the X-ray wavelength of the beam, d is the interplanar space and θ is the angle of incidence.[440] Diffraction maxima, also described as diffraction spots, are observed during crystal rotation on a dimensional detector and each diffraction spot is defined by three so-called Miller indices h, k, l indicating the position of the diffracting lattice plane in the crystal.[438] Each diffraction spot is the result of a wave scattered from the crystal plane described by an amplitude and a phase and each wave is further mathematically represented by a structure factor $F_{(hkl)}$. Then again, structure factors are the Fourier transform of the distribution of the atoms in a unit cell.[438-440] The intensity of the diffraction spot can be measured; from there, the amplitude can be calculated, as it is proportional to the square root of the observed intensity. However, to calculate an electron density map to obtain a model of the molecular structure of the protein the phases of the reflections hkl, are further needed. The phase angles cannot be experimentally determined and this problem is well-known as the phase problem.[438, 444] Thus, indirect methods were developed to obtain and reconstruct the required phase information. For proteins most commonly multiple isomorphous replacement (MIR)[425, 445-446], multiwavelength anomalous dispersion (MAD)[447-449] and molecular replacement[450-451] are used as summarized in many excellent reviews.[452-453]

Macromolecular crystallography has undergone a series of major advances in the last 30 years. Especially with the development of charged coupled device (CCD) detectors and Dectris pixel array detectors[454-455], the increased power and focused X-rays at synchrotron beamlines, and the automation at the beamline site, the throughput of X-ray data collection has increased, reaching more than 20 crystals per hour.[456] Most notably, at the Diamond light source a sample storage dewar with liquid nitrogen monitoring and a capacity of up to 37 pucks (592 samples) including the implementation of a BART sample changer enables an unattended automatic data collection for up to 600 crystals.[457] Nevertheless, also for beamline users in Grenoble, a high quality and rapid automated data collection service is provided since the first upgraded MASSIF beamline in 2014.[458]

Additionally, the throughput of X-ray fragment screening efforts reached the next level with the implementation of specific fragment screening facilities offered by the lab of Marquez at the EMBL, Grenoble and Frank van Delft through the XChem facility at the Diamond in Oxford. Both facilities have implemented a streamlined process for high-throughput crystal soaking to data analysis. Marquez's lab has developed a CrystalDirect™ automatic crystal harvester to accelerate the time-consuming and delicate crystal-harvesting step and to improve the reproducibility and reliability.[459-460] Apart from that, a Crystallization Information Management System (CRIMS) was implemented, which allows remote tracking of the complete streamlined crystallization process from the crystal plate setup to the analysed data in real-time. On the other hand, the XChem workflow is built around TeXRank[461], which is a tool for crystal ranking, a Shifter as a supportive tool for crystal harvesting, a XChemExplorer[462] as an interface for user-friendly data processing and a PanDDA method[463-464] for the identification of weak fragment binders.

1.3.4 Fragment hit validation and characterization

No matter which primary fragment screening approach is used, hit validation is a key step in the FBDD approach. Fragment identification pushes the limits of assays and the use of high concentrations of fragment solutions can result in experimental artefacts. In principle, each technique is expected to identify the same fragment hits but due to different screening conditions, sensitivities, technique-specific artefacts various degrees of overlap were experienced depending on the protein target.[348] Therefore, identified fragment hits should be validated in orthogonal methods and the differences need to be evaluated to identify potential pitfalls and issues in the FBDD campaign.[404, 465]

In addition, the detailed characterisation of validated fragment hits including binding affinity and binding mode data are valuable information to improve the fragment-to-lead optimisation and determine the appropriate fragment evolution strategy.[348] In many cases this includes acquiring many high-resolution crystal structures; however, Jahnke *et al.* have recently published a review on the potential of NMR structures and computational approaches as alternatives.[466] At this early stage,

the chemical space of initial fragment hits should be explored by a "SAR by catalogue" approach to identify related compounds with improved binding affinities or alternative functional groups beneficial for fragment evolution.[348, 378]

1.3.5 Fragment-to-lead optimisation

Validated and characterised fragment hits have affinities up to the mM range and require a hit and target-specific optimisation strategy to evolve into lead compounds and eventually enter the clinics as a drug.[348] During this process, it is important to consider the affinity in combination with the MW and lipophilicity and therefore, as a guide for the fragment optimisation process, the ligand efficiency (LE) metric, introduced by Hopkins *et al.* has been widely applied.[467-468]

$$LE = -\frac{\Delta G}{HA} = RT\ x \ln\frac{K_d}{C^0} = -2.303RT\ x \log\frac{K_d}{C^0}$$

R = ideal gas constant 1.987 x 10^{-3} kcal/K/mol
T = absolute temperature in Kelvin, which is 300 K at standard conditions
C^0 = standard concentration that is generally set to 1 M
K_d = dissociation constant

LE is the free energy of the ligand binding ΔG in kilocalories per mole, which is divided by the number of non-hydrogen atoms (HA). The metric is considered to describe how efficiently a ligand binds to its target based on its atoms.[467, 469] In the literature it is claimed that fragment hits with a high LE > 0.3 kcal mol^{-1} per non-hydrogen atom should be chosen for optimisation as the LE generally decreases during hit optimisation.[467] However, a study of Johnson *et al.* on different fragments-to-lead programs reported no statistically significant difference in the hit-to-lead compounds, which fits with an examination of Astex researchers of different fragment optimisation programs as well.[470-472] In recent years, critical opinions on the usefulness of LE[470, 473-474] were published pointing out the limitations of the metric and for that reason, also other metrics, such as ligand-lipophilicity efficiency (LLE)[475], group efficiency (GE)[476], fit quality (FQ)[477] were introduced. Nevertheless, if its limitations are kept in mind these metrics are a simple tool and can be used as a crude guideline for ligand design in hit-to-lead optimisation programs.[467, 469, 478]

Three main strategies for the fragment-to-lead optimisation are distinguished: fragment linking, growing and merging (Figure 21).[348] One of the most challenging approaches is fragment linking based on the identification of two low-affinity fragments in appropriate distance to bridge the fragments without losing any important functional groups of either fragment and without significant changes in the binding orientation.[329, 348] The most difficult aspect of this strategy is to identify an appropriate linker and to preserve the binding mode of the two fragments. Nevertheless, this strategy can rapidly elaborate weak affinity binders to high potency leads.[329, 348] Until now, the approach was successfully applied on several targets, such as the pyruvate kinase M2[479], the serine

protease thrombin[480], the aspartic protease endothiapepsin[481], the CK2 kinase[482] and also retrospective analysis for a better understanding of the linking process have been described.[348, 483-485]

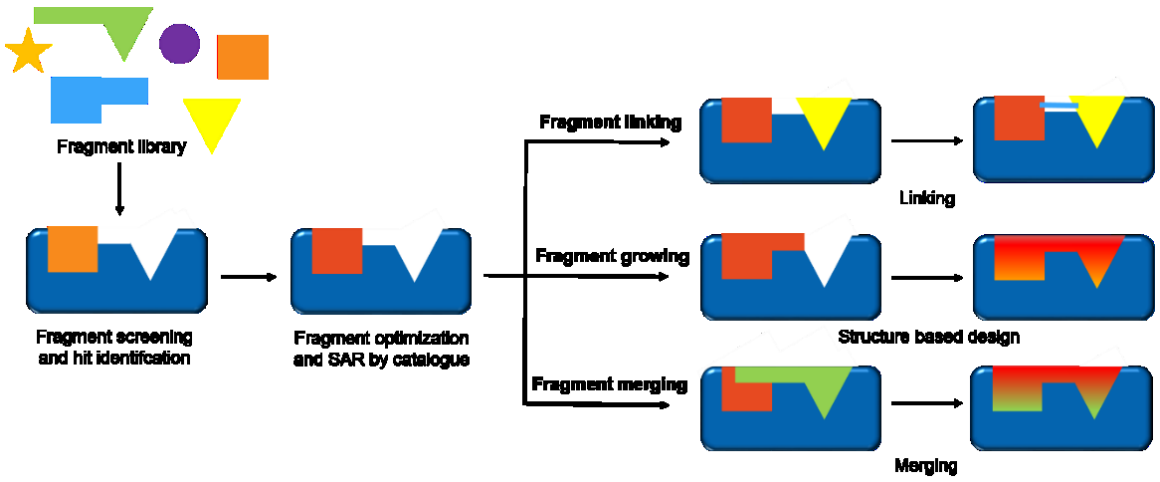


Figure 21: Strategies for fragment-to-lead optimisation. The blue rectangular shape represents the protein and the various other shapes indicate fragments. The three different main fragment evolution strategies are depicted. Fragment linking: two fragments (red and yellow) are binding in adjacent binding sites and a linker (light blue) links these fragments together. Fragment growing: A fragment binding in one binding site is optimised by structure-based design to fill the complete binding site. Fragment merging: Two fragments (red and green) overlap in the binding site and are merged together to fill the complete site. Adapted from [329].

Alternatively, fragments can be optimised following the fragment merging approach if different fragments that overlap by at least one atom are identified. In this case, multiple crystal structures are necessary and substructure components of different fragment hits are combined in a structure-based design.[329, 348] Apart from that, the approach can be useful to combine already patented structures with the information about identified fragment hits to develop alternative chemical scaffolds.[329, 348] In two recent examples on mycobacterial tuberculosis EthR and cytochrome P450 (CYP121), two fragments for each of the targets were merged together resulting in a 100-fold and 15- to 60-fold increased potency, respectively.[363, 486-487] Although fragment merging does not require the thorough identification of a fragment linker it still poses critical challenges and initial fragment merging optimisation on mycobacterium tuberculosis CYP121 did not lead to potent binders.[348, 487] However, the approach has proven successful on targets such as chymotrypsin-like serine protease[488] and tankyrase enzymes 1 and 2[489].

Last but not least, the fragment-growing approach is most commonly applied and describes the optimisation of a single fragment hit by introducing further groups and improving the target interaction in iterative medicinal chemistry cycles.[363] Similarly to the initial fragment hit characterisation, in a variation of this approach, SAR by catalogue can be performed with the advantage of an accelerated discovery process, as described for the development of an inhibitor of Hsp90.[329] Of all possible hit-to-lead approaches, this strategy is best feasible also in the absence of crystal structures. Then, alternative approaches, such as NMR and docking, for structural information can be used.[348, 466, 490] However, in a recent review summarising published

fragment-to-lead programs of 2017, it is shown that only three out of 30 optimisation campaigns were followed without obtaining any X-ray structure during the optimisation process.[340]

Overall, in the last couple of years lead molecules derived from a fragment-based approach were developed for a range of target classes, including targets previously considered as undruggable, such as protein-protein interactions (PPIs), and several of those molecules have already entered the clinic.[340]

1.4 Aim of the thesis

Currently available drug treatment options for Human African Trypanosomiasis (HAT) caused by *Trypanosma brucei* (*T. brucei*) parasites suffer from severe side effects and low efficacy.[42] Hence, new drugs with a better safety and efficacy profile are urgently needed. The bone disease drugs, nitrogen-containing bisphosphonates have been shown to block *T. brucei* parasitic growth by inhibiting farnesyl pyrophosphate synthase (FPPS), an enzyme part of the sterol biosynthesis, but they exhibit undesired pharmacokinetic properties for non-bone indications. Fragment-based drug discovery has shown broad success for the identification of new binding pockets and in a reported fragment-based approach on human FPPS, several fragment binders were detected in a previously unknown allosteric site.[310] Human FPPS shares an 64 % overall similarity[300] with *T. brucei* FPPS but only limited research on *T. brucei* FPPS has been reported and non-bisphosphonate inhibitors are lacking. [115-116, 123, 299-300]

The goal of this thesis was the identification of non-bisphosphonate binders in a fragment screen and the optimisation of fragments to potent tool compounds. For this purpose, *T. brucei* FPPS was required in sufficient yields and purities and was expressed and purified from *E. coli* cells. In addition, the identification of a robust crystal system that could enable high-throughput crystallization of both free and ligand-bound *T. brucei* FPPS was required. Published fragments discovered in the previously identified allosteric pocket in human FPPS were tested on *T. brucei* FPPS to investigate the allosteric site of human FPPS in *T. brucei* FPPS. Additionally, *T. brucei* FPPS was subjected to a biophysical fragment screen based on NMR and X-ray with the aim to identify new fragment binders and to follow-up identified fragment hits in the respective other orthogonal method.

Fragment screening by X-ray and NMR on *T. brucei* FPPS afforded binders in various binding sites including the active and allosteric site. Thus, structure-activity relationship by archive and catalogue was performed to identify more potent analogues and starting points for structure-based fragment optimisation by medicinal chemistry. Furthermore, analogues were synthesised by medicinal chemistry to optimise binding affinities. The overarching goal beyond the scope of this thesis was the optimisation of identified fragment hits towards a tool compound.

2. Material and Methods

2.1 Cell lines

Competent BL21 (DE3) phage T1 resistant *Escherichia coli* cells (NEW ENGLAND BIOLABS)

Genotype : *fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS Δ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*

Aliquots: 0.05 mL/tube

Competent BL21 (DE3) phage T1 resistant *Escherichia coli* cells (NEW ENGLAND BIOLABS)

Transformed with plasmid containing *BirA* gene under the control of an Ara promoter and made chemically competent again were kindly provided by Cecile Delmas and used for *in-vivo* avi-tagged biotinylated FPPS expression.

Aliquots: 0.1 mL/tube

Competent DH5α cells (THERMO SCIENTIFIC FISHER)

Genotype: F- φ80lacZΔM15 Δ(lacZYA-argF)U169 recA1 endA1 hsdR17(rk - , mk +) phoA supE44 thi-1 gyrA96 relA1 λ

Aliquots: 50 µL

2.2 Plasmids

***T. brucei* FPPS**

N – His_6 – HRV-3C protease cleavage site – $TbruFPPS_{1-367}$ – C

Trypanosoma brucei FPPS (UniProt: Q86C09, ec_opt), M1-K367, in pACE0

T7 promotor, lac operon, Kan^R, 5611 bp

Avi-tagged *T. brucei* FPPS

N – His_6 – linker(3xGGGS) – HRV-3C protease cleavage site – Avi-$TbruFPPS_{1-367}$ – C

Trypanosoma brucei FPPS (UniProt: Q86C09, ec_opt), M1-K367, in pACE-GP9

T7 promotor, lac operon, Kan^R, 5698 bp

Human FPPS

N – His_6 – HRV-3C protease cleavage site – $hFPPS_{67-419}$ – C

Human FPPS (UniProt: P14324, ec_opt), M67-K419, in pACE0

T7 promotor, lac operon, Kan^R, 5569 bp

Avi-tagged human FPPS

N – His_6 – linker(3xGGGS) – HRV-3C protease cleavage site – Avi-$hFPPS_{67-419}$ – C

Human FPPS (UniProt: P14324, ec_opt), M67-K419, in pACE-GP9

T7 promotor, lac operon, Kan^R, 5656 bp

Avi-tagged *T. cruzi* FPPS

N – His_6 – linker(3xGGGS) – HRV-3C protease cleavage site – Avi-$TcruFPPS_{64-425}$ – C

Trypanosoma cruzi FPPS (UniProt: Q8WS26, ec_opt) M64-K425 in pACE-GP9

T7 promotor, lac operon, Kan^R, 5683 bp

Human rhinovirus 3C protease

N - MBP - His_6 - Thrombin cleavage site - HRV - 3C $protease_{1-182}$ – C

Human rhinovirus 3C protease (UniProt: P03303), G1-Q182

T7 promotor, lac operon, Kan^R, 6722 bp

***E. coli* bifunctional ligase/repressor ligase A (BirA)**

N - His_6- $BirA_{1-321}$ - C

E. coli birA (UniProt: P06709) M1-K321, in pACYC184
araBAD promoter, araC operon, Cam^R

Felix Freuler, NIBR, Basel, Switzerland kindly did codon optimisation for *E. coli* expression of all plasmids used in this thesis. *T. brucei* FPPS and human FPPS plasmids were ordered from GeneArt and diluted in TE buffer solution at 0.1 µg/L plasmid concentration. Cloning of avi-tagged *T. brucei* FPPS, avi-tagged *T. cruzi* FPPS and avi-tagged human FPPS plasmid vectors was done as part of this thesis in collaborative work together with Simon Hänni, NIBR, Basel, Switzerland. Plasmids were diluted to a final concentration of 0.1 µg/L in elution buffer of Plasmid DNA purification kit. Avi-tagged *T. cruzi* FPPS plasmids were provided to Joy Petrick, NIBR, Basel, Switzerland.

2.3 Proteins

Proteins listed in Table 1 were overexpressed in *E. coli* cells. Protein expression of human FPPS was performed in collaboration with Joy Petrick, NIBR, Basel, Switzerland.

Table 1: Proteins listed with their respective MW and Absorbance correction (A280) for 1 mg/mL.

Proteins	MW	corr. A_{280} for 1 mg/mL
T. brucei FPPS	42125.3	1.289
Avi-tagged *T. brucei* FPPS	43936.3	1.362
Human FPPS	40686.6	1.346
Avi-tagged human FPPS	42497.6	1.418
HRV 3C protease	62854.5	1.150

Pierce bovine serum albumin (BSA) standard, 1.5 mg/mL, THERMO FISHER SCIENTIFIC

Pierce bovine serum albumin (BSA) standard, 2 mg/mL, THERMO FISHER SCIENTIFIC

2.4 Chromatography resins and columns

Protein purification

Ni-NTA Superflow 5 mL, QIAGEN

HisTrap™ HP crude, 5 mL, GE HEALTHCARE

Size-Exclusion column, Highload Superdex™ 16/60 S200, 120 mL, GE HEALTHCARE

Buffer exchange and desalting

HiPrep™ 26/10, desalting column, GE HEALTHCARE

PD-10, desalting columns, pre-packed columns containing Sephadex™ B-25 Medium, GE HEALTHCARE

Medicinal chemistry

12 g silica gel prepracked columns, RediSep® Rf, TELEDYN ISCO

2.5 Buffers and Solutions

2.5.1 Protein expression and purification buffers

Table 2: Protein expression and purification buffers and their components.

Buffer	Components
Lysis buffer	50 mM Tris-HCl, 300 mM NaCl, 5 mM imidazole, 10 % v/v glycerol, 2 mM TCEP-HCl, pH 8
Ni^{2+} affinity chromatography	
Buffer A	50 mM Tris-HCl, 300 mM NaCl, 5 mM imidazole, 2 mM TCEP-HCl, 10 % v/v glycerol, pH 8
Buffer B	50 mM Tris-HCl, 150 mM NaCl, 200 mM imidazole, 2 mM TCEP-HCl, 10 % v/v glycerol, pH 8
SEC buffer	
Tris buffer	50 mM Tris-HCl, 200 mM NaCl, 2 mM TCEP-HCl, pH 8
Bis-Tris buffer	25 mM Bis-Tris, 50 mM NaCl, 2 mM TCEP-HCl, pH 8

2.5.2 NMR buffers

Table 3: NMR buffers and their components.

Buffer	Components
NMR core fragment library screening buffer	10 mM d_6-Tris, 150 mM NaCl, 1 mM d_{12}-TCEP-HCl, pH 8
Bis-Tris NMR buffer	25 mM Bis-Tris, 50 mM NaCl, 2 mM TCEP-HCl, pH 6.5
Solubility buffer	25 mM d_{18}-HEPES, 150 mM NaCl, 1 mM TCEP-HCl, 2 mM $MgCl_2$, 150 µM DSS, 0.01 % v/v Tween 20, 10 % v/v D_2O, pH 7.5

2.5.3 SPR buffer

Table 4: SPR buffers and their components.

Buffer	Components
SPR buffer	50 mM HEPES, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM TCEP-HCl, 0.01 % v/v Tween 20, pH 7.4
SPR buffer with DMSO	50 mM HEPES, 150 mM NaCl, 2 mM $MgCl_2$, 2 mM TCEP-HCl, 0.01 % v/v Tween 20, 0.9 % v/v DMSO, pH 7.4

2.5.4 Other buffers

Table 5: SDS, DSF, HPLC, UPLC-MS and Plasmid buffers with their components.

Buffer	Components
SDS buffer	
SDS PAGE running buffer	25 mM Tris pH 8.3, 192 mM glycine, 0.1 % w/v SDS
SDS sample buffer (4 x):	4 × Laemmli loading buffer, VWR Life Science Amresco
HPLC buffer	
Buffer A	MilliQ + 1 % v/v TFA
Buffer B	90 % v/v acetonitrile, 10 % v/v MilliQ + 1 % v/v TFA
UPLC-MS	
Buffer A	water + 0.05 % v/v TFA
Buffer B	acetonitrile + 0.04 % v/v TFA
Plasmid buffer	
TE buffer	10 mM Tris-HCl, 0.1 mM EDTA, pH 8
Elution buffer	5mM Tri-HCl pH 8.5

2.5.5 Other solutions

Solubility and Stability Screen II, Hampton Research
Guanidine cleaning solution containing 20 mM TRIS, pH 8.0, 6 M guanidine, 100 mM NaCl, and 10 mM DTT
30 mM DSS stock solution in milliQ, prepared in-house

2.5.5.1 Crystallization screens

Table 6: List of Crystallization screens used with supplier.

Crystallization screen	Supplier
NeXtal DWBlock $AmSO_4$ Suite	QIAGEN
NeXtal DWBlock Cryos Suite	QIAGEN
NeXtal DWBlock JCSG+ Suite	QIAGEN
NeXtal DWBlock MbClass II Suite	QIAGEN
Index HT	HAMPTON RESEARCH
PEGRx HT	HAMPTON RESEARCH
SaltRX HT	HAMPTON RESEARCH
NeXtal DWBlock MbClass Suite	QIAGEN
NeXtal DWBlock PACT Suite	QIAGEN
NeXtal DWBlock PEGsII Suite	QIAGEN
NeXtal DWBlock pHClear II Suite	QIAGEN
NeXtal DWBlock PEGs Suite	QIAGEN
PEG/Ion HT	HAMPTON RESEARCH

2.5.5.2 Solutions for crystallization plates

Table 7: List of stock solutions with suppliers used for preparation of crystallization conditions.

Solution	Supplier
1 M caesium chloride	HAMPTON RESEARCH
50 % w/v PEG 3350	HAMPTON RESEARCH
1.5 M sodium tartrate dibasic dihydrate	HAMPTON RESEARCH
2.5 M potassium citrate tribasic monohydrate	HAMPTON RESEARCH
1 M HEPES pH 6.6	FLUKA, solution prepared
1 M HEPES pH 8.5	FLUKA, solution prepared
50 % w/v PEG 8000	HAMPTON RESEARCH
100 % v/v ethylene glycol	FLUKA, solution prepared
1 M sodium cacodylate pH 5.0	FLUKA, solution prepared
1 M sodium cacodylate pH 7.4	FLUKA, solution prepared
1 M magnesium acetate tetrahydrate	HAMPTON RESEARCH
50 % w/v PEG MME 2000	HAMPTON RESEARCH
50 % w/v PEG MME 5000	HAMPTON RESEARCH
50 % w/v PEG 6000	HAMPTON RESEARCH

Table 7 continued.

50 % w/v PEG 10000	HAMPTON RESEARCH
30 % w/v PEG 20000	FLUKA, solution prepared
100 % v/v tert-butanol	MERCK, solution prepared
50 % v/v 1,4-dioxane	HAMPTON RESEARCH
0.5 M barium chloride	MERCK, solution prepared
1 M proline	HAMPTON RESEARCH
0.1 M EDTA	FLUKA, prepared
75 % PEG 400	HAMPTON RESEARCH
4 M magnesium chloride	FLUKA, solution prepared
30 % v/v xylitol	SIGMA ALDRICH, solution prepared
50 % v/v PE 15/4 EO/OH	HAMPTON RESEARCH
2 M NDSB-211	HAMPTON RESEARCH, solution prepared
100 % v/v tacsimate pH 4.0	HAMPTON RESEARCH
100 % v/v tacsimate pH 8.0	HAMPTON RESEARCH
3.4 M sodium malonate pH 4.0	HAMPTON RESEARCH
3.4 M sodium malonate pH 8.0	HAMPTON RESEARCH

2.5.6 Media and other reagents

Table 8: List of solutions for protein expression in *E. coli* cells and their components.

Solution	Components
1000 × trace metals mix	50 mM $FeCl_3$, 20 mM $CaCl_2$, 10 mM $MnCl_2 \times 4H_2O$, 10 mM $ZnSO_4 \times 7H_2O$, 2 mM $CoCl_2$-$6H_2O$, 2 mM $CuCl_2 \times 2H_2O$, 2 mM $NiCl_2 \times 6H_2O$, 2 mM $Na_2MoO_4 \times 5H_2O$, 2 mM $Na_2SeO_3 \times 5H_2O$, 2 mM H_3BO_3
100 × vitamin mix	50 g/L thiamine hydrochloride, 10 g/L D-panthothenic acid, 10 g/L biotine, 10 g/L pyridoxal hydrochloride, 10 g/L folic acid, 10 g/L nicotinamide, 1 g/L riboflavin
1000 × kanamycin stock	50 mg/mL kanamycin
1000 × IPTG stock	0.5 M IPTG

Table 9: List of media and their components for protein expression in *E. coli* cells.

Medium	Components
LB medium	10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract
SOC medium, Fluka	20 g/L tryptone, 5 g/L yeast extract, 4.8 g/L $MgSO_4$, 3.603 g/L dextrose, 0.5 g/L NaCl, 0.186 g/L KCl
MDG medium	25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.2 × metals, 0.5 % w/v glucose, 0.25 % w/v aspartate
TB 2 ×	26.6 g/L tryptone, 53.4 g/L yeast extract, 4.4 mL glycerol
ModAl medium	25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 1 × metals, 2.5 % w/v tryptone, 5 % w/v yeast, 1 % w/v glycerol, 0.1 % w/v glucose, 0.4 % w/v α- lactose
^{13}C, ^{15}N uniform labelling medium	50 mM KH_2PO_4, 50 mM Na_2HPO_4, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.4 % w/v ^{13}C-D-Glucose, 0.25 % w/v $^{15}NH_4Cl$ 0.25 × trace metals mix, 1 × vitamin mix

2.5.7 Kits and reagents for plasmid cloning

Table 10: List of kits and reagents with their supplier used for plasmid cloning.

Reagent	Supplier
NucleoSpin, Gel and PCR clean up - kit	Macherey-Nagel
Topvision Agarose tablets	Thermo Scientific
Mastermix Q5, Hot Start High-Fidelity 2 × Master mix	New England Biolabs
GeneJET Plasmid Miniprep Kit	Thermo Scientific

2.5.8 Fragment libraries

Novartis core fragment library

The 6th generation of the Novartis core fragment library consists of 1336 compounds based on the commonly used fragment definition "rule of 3" with the exception of allowing ≤ 5 H-binding acceptors and clogP < 3. Furthermore, structures with known undesired substructures for follow-up studies were excluded and each compound had to be available in more than 30 mg and with 20 or more fragment analogues in the Novartis archive. The library is setup in 176 mixtures of eight compounds at a total mixture concentration of 50 mM in d_6-DMSO in a 96 well plate layout readily to order from the Novartis compound archive.

Novartis fluorine fragment library

The Novartis fluorine fragment library consisted of 19 compound mixtures with 470 CF_3-containing fragments and 12 CF_2-containing fragments dissolved in d_6-DMSO. The library was originally setup to contain 30 compounds in each mixture with a total compound concentration of 50 mM. Therefore, each compound was at 1.66 mM concentration. However, over the time some compounds were found to decompose, to be toxic or the stock solution was emptied and in this cases the compound mixture did not contain 30 compounds anymore instead the compound was replaced by d_6-DMSO to keep the same dilution of each compound in the individual mixture. The library was available ready to order in a 96 well plate layout from the Novartis compound archive.

Enamine Golden Library

The Enamine Golden Fragment Library is a commercially available fragment library with 1794 compounds. It was designed by Enamine to contain the most diverse set of fragments based on the following features: novelty, unique chemotypes, diversity, all compounds fulfilling the "rule of 3", measured solubility of minimum 1 mM in PBS buffer, possibility of easily growing the fragments, excluding PAINS, a purity greater than 90 % determined by NMR and LC(GC)-MS data and availability.[491] Part of the Enamine Golden Fragment Library was screened by X-ray crystallography at the HTX lab in Grenoble. The library was available in a plate-layout with single compounds dissolved at 100 mM and 500 mM in DMSO.

DSPL Library

The 1st generation of Diamond and SGC Poised Fragment Library contains (DSPL) contains 407 fragments as single compounds in a 500 mM concentration in d_6-DMSO and 407 were used for X-ray fragment soaking experiments. The library was setup on 384 well Echo compatible plates.

Edelris Library

The Edelris Keymical Fragments library contains 280 fragments as singles in 250 mM compound concentration stock solution in d_6-DMSO and 278 fragments were used for X-ray fragment soaking experiments. The library is set up on a 384 well Echo compatible source plate.

2.5.9 Equipment and Devices

Table 11: List of equipment/devices and manufacturer.

Equipment / Device	Manufacturer
Polytron® PT 1200 E	POLYTRON
Centrifuge 5810R, Rotor A-4-62	EPPENDORF
Centrifuge, SORVAL RC3BP, Rotor H600A	THERMO SCIENTIFIC
Centrifuge Avanti J30 I, Rotor JLA-16.250	BECKMAN COULTER
Chromatography system, automated purification, Äktaxpress™	AMERSHAN BIOSCIENCE
Chromatography system, ÄKTA Avant 25	GE HEALTHCARE
CrystalCap™ SPINE HT 0.1 – 0.2 mm CryoLoops	HAMPTON RESEARCH
CrystalCap™ SPINE HT 0.05 – 0.1 mm CryoLoop	HAMPTON RESEARCH
Dual-Thickness MicroLoops LD™ 35 um	MITEGEN, LLC
Dual-Thickness MicroLoops LD™ 50 um	MITEGEN, LLC
Dual-Thickness MicroLoops LD™ 70 um	MITEGEN, LLC
VDX micro plate, 24-well, 18 mm, greased	HAMPTON RESEARCH
Swissci 2-drop, 96 well	HAMPTON RESEARCH
Swissci 2-drop, 96 well	HAMPTON RESEARCH
CrystalDirect™, 96 well plate	MITEGEN, LLC
Rock Imager 1000 – Dual Imager	FORMULATRIX INC
Cuvettes PS semi-micro	VWR
Deepwell block, Masterblock 2 mL sterile 96 well v-shape	GREINER BIO-ONE
DynaPro Plate Reader	WYATT TECHNOLOGY
Electrophoresis chamber, Mini-Protean Tetra-System	BIORAD
Fermentor, Labfors 5	INFORS HT
Fermentor cleaning system, LabCIP	INFORS HT
Centrifugal filter with a 30 kDa MWCO, Amicon Ultra-15, cut-off 30 kDa	MERCK MILLIPORE
Filter unit, 0.45 µm, Millipore® Stericup™, 0.45 µm, PVDF, 1000 mL	MERCK MILLIPORE
Filter unit, 0.22 µm, Millipore® Stericup™, 0.22 µm, PVDF, 1000 mL	MERCK MILLIPORE
Echo 550	LABCYTE, INC.
Eppendorf sterile filter, Ultrafree-MC, Durapore 0.45 µm	MILLIPORE
Eppendorf sterile filter, Ultrafree-MC, Durapore 0.22 µm	MILLIPORE
French press, EmulsiFlex-C50	AVESTIN
Gel Doc™ EZ Imager	BIORAD
Glas fibre prefilter, non- sterile	MERCK
HPLC, 1290 Infinity II LC System	AGILENT TECHNOLOGIES
Microbiological Incubator, INCU-line	VWR
Micro centrifuge, Centrifuge 5415 R	EPPENDORF
Microwave 1700	KOENIG

Table 11 continued.

NanoDrop™ OneC	THERMO SCIENTIFIC
SampleJet	BRUKER
NMR spectrometer, 400 MHz, Bruker DPX 400 MHz	BRUKER
NMR spectrometer, Avance™ III HD 600 MHz	BRUKER
NMR spectrometer, Bruker Avance III 600 MHz, quadruple	BRUKER
NMR spectrometer, Bruker Avance III HD 800 MHz	BRUKER
Bruker Avance 1 TXI 600 MHz	BRUKER
NMR tubes, 3 mm SampleJet	BRUKER BIOSPIN AG
Sample Jet Rack, 96 × 3 mm	BRUKER BIOSPIN AG
PCR detection system, CFX384 real-time	BIORAD
pH-Meter, Seven Compact	METTLER TOLEDO
Photometer, BioPhotometer UV/VIS	EPPENDORF
Mosquito robot with humidity chamber	TTP LABTECH
CyBi-well robot	CYBIO
Tecan, Freedom evo	TECAN
Plate sealant, Adhesive PCR Sealing Foil sheets, aluminium	THERMO SCIENTIFIC
Plate sealant crystallization, Crystal clear sealing film	HAMPTON RESEARCH
Plate sealant, DSF, Microseal® B Adhesive sealing film	BIORAD
Plate sealant, SPR, microplate foil	GE HEALTHCARE
Plate shaker, MixMate®	EPPENDORF
Plate, DSF, 384-well, Hard-shell 384 microplate	BIORAD
Plate, SPR/NMR, 96 well Greinert plate	GREINERT
Plate, DLS, 384w Corning 3540 non treated; black; small volume; clear bottom	CORNING
Power source for Electrophoresis, PowerPAC™ Basic	BIORAD
Precast gels, Mini-PROTEAN® TGX™ Precast gels, 4-20%, 15-well, 15 µL	BIORAD
Harvesting bags, Whirl-Pak® Stand-up bag	NASCO
Thermomixer, comfort 2 mL	EPPENDORF
Shaking incubator, Shaker X	KUHNER
Rock Imager 1000 in Oxford and Grenoble	FORMULATRIX INC.
UPLC-MS, UPLC-ESI-Q-TOF-MS, Xevo-G2-S QTof, Zspray™ source, ESI, modular, Lockspray™ interface, Acquity™ UPLC system	WATERS
Polypropylene cell spreader, sterile	VWR
Water bath, TW12	JULABO
Cover slide, Siliconized circle cover slides, 18 mm × 0.22 mm	HAMPTON RESEARCH
Thermocycler, GeneAmp, PCR system 9700	THERMO SCIENTIFIC
Gel imager	WITEC AG
Multichannel pipette	THERMO SCIENTIFIC

Table 11 continued.

Multidispense pipette	METTLER TOLEDO
SPR vial, Polypropylene 4 mL Shell vial	WHEATON
SPR vial, Polypropylene 7 mm vial	GE HEALTHCARE
Rubber cap, type 5	GE HEALTHCARE
Rubber cap, type 3	GE HEALTHCARE
Pipettes, Research Plus®, 2.5 µl, 10 µl, 200 µl, 1000 µl	EPPENDORF
VDXm plates, 24 wells, 128 mm x 85 mm microplate footprint	HAMPTON RESEARCH

2.5.10 Software

Table 12: List of used software under Windows and Linux.

Name and version	Source / Reference
1290 Infinity II Walkup Software	AGILENT TECHNOLOGIES
ÄKTA XPRESS SOFTWARE: UNICORN™ 5.31	GE HEALTHCARE LIFE SCIENCE
ChemBioDraw® Ultra, 14.0	PERKINELMER
Dynamics 7.1.8	WYATT TECHNOLOGY
fitKD	NOVARTIS, IN-HOUSE SCRIPT BY ARMIN WIDNER
gedit	PYTHON TEXT EDITOR
Glide, Release 2018-1	SCHRÖDINGER, LLC
IconNMR	BRUKER BIOSPIN
IconNMR	BRUKER BIOSPIN
Image Lab 6.0	BIORAD
MassHunter Walkup Software	AGILENT TECHNOLOGIES
MNova screen	MESTRELAB RESEARCH
MS Office 2016	MICROSOFT
OpenLAB CDS Workstation Software Plus	AGILENT TECHNOLOGIES
ProtParam tool	EXPASY WEB SERVER
PyMOL Molecular Graphics System. Version 1.8.2.2.	SCHROEDINGER, LLC
RockMaker, up to version 3.14	FORMULATRIX INC.
TopSpin, fragment-based screening (FBS) tool, test version	BRUKER
TopSpin, up to version 3.2	BRUKER BIOSPIN

2.5.11 Chemicals

Table 13: List of chemicals and supplier.

List of chemicals	Supplier
1-(2-Hdyroxyethyl)imidazole	SIGMA ALDRICH
1-(3-bromopropoxy)-4-chlorobenzene	VWR INTERNATIONAL AG
1-(3-chlorobenzyl)piperazine	SIGMA ALDRICH
^{13}C-D-glucose	SIGMA ALDRICH
$^{15}NH_4Cl$	CAMBRIDGE ISOTOPE LABORATORIES INC.
1*H*-pyrrolo[3,2-b]pyridine-3-carbaldehyde	ACCEL PHARMTECH
1*H*-pyrrolo[3,2-c]pyridine-3-carbaldehyde	EMOLECULES
3-chlorobenzylchlorid	SIGMA ALDRICH
3*H*-indazol-3-one,5-chloro-1,2-dihydro	EMOLECULES
4-(Boc-amino)-piperidine	ASTATECH, INC.
5-(trifluoromethyl)isoindoline	EMOLECULES
5-bromoisoindoline hydrochloride	ASW MEDCHEM, INC.
5-chloro-1*H*-indazole	ARK PARM, INC.
5-chloroisoindoline	ABCR
5-methyl-2,3-dihydro-1*H*-isoindole hydrochloride	EMOLECULES
acetic acid	SIGMA ALDRICH
acetonitrile, ≥99.9 %	FLUKA
acetonitrile	BRENNTAG SCHWEIZERHALLE AG
aspartic acid	FLUKA
biotin	SIGMA-ALDRICH
BisTris	FLUKA
$CaCl_2$	MERCK
chloramphenicol	APPLICHEM
$CoCl_2 \times 6H_2O$	RIEDEL-DE HAËN
cOmplete™	ROCHE DIAGNOSTICS
$CuCl_2 \times 2H_2O$	MERCK
D-(+)glucose	FLUKA
d_{12}-TCEP·HCl	CAMBRIDGE ISOTOPE LABORATORIES INC.
d_{12}-Tris(2-carboxyethyl)phosphine hydrochloride	
D_2O	EURISOTOP
d_6-DMSO (d_6-dimethylsulfoxide) for chemistry	SIGMA ALDRICH
d_6-DMSO (d_6-dimethylsulfoxide) for NMR	EURISOTOP
d_6-TRIS (tris(hydroxymethyl)aminomethane)	CAMBRIDGE ISOTOPE LABORATORIES INC.
DCM	BRENNTAG SCHWEIZERHALLE AG
DCM, anhydrous, 99.8 %	SIGMA ALDRICH
DMF, anhydrous, 99.8 %	SIGMA ALDRICH

Table 13 continued.

D-Panthothenic acid (Vitamin B_5)	SIGMA-ALDRICH
DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid)	CAMBRIDGE ISOTOPE LABORATORIES INC.
DTT, dithiothreitol solution, 1M in H_2O	SIGMA
ethylacetate 99 – 100 %	BRENNTAG SCHWEIZERHALLE AG
$FeCl_3 \times 6H_2O$	SIGMA-ALDRICH
folic acid (Vitamin B_9)	SIGMA-ALDRICH
glycerol	SIGMA-ALDRICH
guanidine-HCl	SIGMA
guanidine-HCl solution, 8 M in H_2O	SIGMA
H_3BO_3	FLUKA
heptane isomerengemisch	BRENNTAG SCHWEIZERHALLE AG
imidazole	SIGMA-ALDRICH
IPTG, Isopropyl-β-D-thiogalactopyranosid	SIGMA LIFE SCIENCE
K_2CO_3	SIGMA ALDRICH
kanamycin sulfate	SIGMA LIFE SCIENCE
KH_2PO_4	FLUKA
$MgSO_4$	SIGMA
$MnCl_2 \times 4H_2O$	MERCK
Na_2HPO_4	MERCK
$Na_2MoO_4 \times 2H_2O$	SIGMA-ALDRICH
$Na_2SeO_3 \times 5H_2O$	SIGMA-ALDRICH
Na_2SO_4	SIGMA ALDRICH
NaCl	SIGMA-ALDRICH
NaH, 60 % dispersion in mineral oil	SIGMA ALDRICH
NH_4Cl	SIGMA LIFE SCIENCE
$NiCl_2 \times 6H_2O$	RIEDEL-DE HAËN
nicotinamide	FLUKA
PPG antifoam, polypropylene glycol 2000	VWR
pyridoxal hydrochloride	SIGMA
riboflavin (Vitamin B_2)	SIGMA
SOC medium	FLUKA
sodium triacetoxy borohydride	SIGMA ALDRICH
SYPRO® Orange	SIGMA
TCEP·HCl, Tris(2-carboxyethyl)phosphine hydrochloride	SIGMA
TEA, triethylamine	SIGMA ALDRICH
tert-butyl 3-formyl-1H-indole-1-carboxylate	ABCR
TFA, trifluoroacetic acid	SIGMA ALDRICH
THF	SIGMA ALDRICH

Table 13 continued.

thiamine hydrochloride (Vitamin B_1)	FLUKA
thionyl chlorid	SIGMA ALDRICH
TRIS, TRIZMA® Base	SIGMA LIFE SCIENCE
TRIS-HCl, TRIZMA® hydrochloride	SIGMA LIFE SCIENCE
tryptone	SIGMA-ALDRICH
water, CHROMASOLV® Plus, for HPLC	SIGMA-ALDRICH
yeast extract	SIGMA-ALDRICH
$ZnSO_4 \times 7H_2O$	SIGMA
α-lactose × H_2O	SIGMA
thionyl chlorid	SIGMA ALDRICH

2.6 Cloning, Transformation, Protein expression and Purification

2.6.1 Cloning

Cloning of plasmid vectors for *T. brucei* FPPS and human FPPS protein expression was kindly done by Simon Hänni, NIBR, Basel, Switzerland. Plasmid vectors for HRV-3C protease were already available at NIBR, Basel, Switzerland. Cloning of plasmid vectors for avi-*T.brucei* FPPS, avi-*T.cruzi* FPPS and human FPPS was done as described.

First, a PCR mix was prepared for the amplification of each template nucleotide sequence. In a PCR tube, 25 µL Mastermix Q5, 12.5 µL forward primer (2 µM) and 12.5 µL reverse primer (2 µM) and 0.5 µL template (0.1 µg/µL, plasmid vector) were mixed and the template was amplified in a thermocycler using the following program:

Start:	98 °C	30 s
30 cycles of:	98 °C	10 s
	58 °C	10 s
	72 °C	40 s
Final:	72 °C	10 min

Then the system was cooled down to 4 °C until the samples were taken out. The agarose gels were prepared from two tablets of 0.5 g Topvision Agarose in a 250 mL Erlenmeyer flask. The tablets were dissolved in 105 g 1 × TAE buffer + 0.1 mg/L ethidium bromide and the solution was heated for 2 min in the microwave. Afterwards the gel solution was poured in the gel chamber and an 8-slot comb was placed in the gel. After cooling, the gel was transferred in the gel frame containing 1 × TAE buffer + 0.1 mg/L ethidium bromide and the comb was removed. To each PCR mix 10 µL 6 × loading dye was added and 58 µL of the final sample were loaded in the well on the gel. Then 10 µL gel marker were also loaded and the gel chamber was connected to the power device and the gel run at 120 V for 1 hour. After the run was finished, the PCR products

were cut out of the agarose gel. The gel was transferred still in the chamber in the gel imager to take a gel image. Then the UV desk was covered with a plastic foil and the gel was slided onto the desk. After personal face protection with a UV shield, the UV was switched on to 70 – 100 % to visualise the gel bands. The bands were cut out with a scalpel and the gel piece of each PCR product was transferred in an Eppendorf tube.

The PCR product was extracted from the gel with a NucleoSpin Gel and PCR Clean-up kit following the provided protocol except for two variations: Instead of using the centrifuge for washing and spinning, the tubes were placed on an in-house vacuum suction plate to remove the excess liquid and only for the final removal of all liquids the tubes were centrifuged. Furthermore, the final DNA was eluted in a new Eppendorf tube.

The washed DNA was inserted in a plasmid vector by T2S ligation. The following components were mixed in an Eppendorf tube:

2 µL	Fermentas buffer B
1 µL	ATP (20 mM)
1 µL	DTT (20 mM)
2 µL	Acceptor vector (0.1 µg/µL)
2 µL	PCR product
1 µL	LguI
1 µL	T4 DNA ligase

The tube was filled up to 20 µL with sterile ddH_2O and the tubes were added on top of the thermocycler. After 2 ½ hours, 5 µL solution was taken out for the transformation. The tubes were left on the thermocycler overnight and then used for *E. coli* transformation as described in the next chapter with the following deviations: 5 µL DNA were added to DH5α cells and the 250 µL Super Optimal broth with Catabolite repression (SOC) medium were added of which all of the cell solution was spread on Luria-Bertani (LB) plates containing kanamycin. After overnight incubation at 37 °C, two to three colonies were picked and transferred into a snap cap tube containing 3 mL LB medium with 50 µg/mL kanamycin. Plasmid DNA was purified from *E. coli* cells following the protocol of the Plasmid DNA purification kit with the following deviation: Instead of 50 µL AE Elution buffer, 110 µL buffer was used. Eluted plasmid DNA samples were stored at 4 °C.

Forward and reverse primer for sequencing were ordered from GeneArt. For each plasmid DNA, three forward and three reverse primers were used after around each 400 bp. In a 96 well plate 8 µL plasmid DNA was mixed with 5 µL primer and sequencing was done by colleagues at NIBR, Basel, Switzerland and results were kindly analysed by Simon Hänni, NIBR, Basel, Switzerland.

2.6.2 Transformation of *E. coli*

For the preparation of LB agar plates 200 mL autoclaved LB medium were heated for 2 min in the microwave. Then, the medium was cooled down to 60 °C to add 50 µg/mL kanamycin and poured into 100 × 15 mm plates. The plates were piled to reduce condensation during agar solidification, stored at 4 °C, and used the next day.

For each transformation, one vial of chemically competent *E. coli* cells BL21 (DE3) was thawed on ice and 1 µL of plasmid was added to the cells. The tube containing *E. coli* cells and plasmid was gently mixed by tapping and incubated on ice for 30 min. Subsequently, the cells were heat shocked by incubation in a 42 °C water bath for one minute after which the tube was quickly transferred on ice for 5 min. Afterwards 400 µL SOC medium was added and the tube was incubated at 37 °C for 1 hour at 500 rpm in a thermomixer. Then 150 µL of transformation reaction solution was spread with a polypropylene cell spreader onto LB agar plates with 50 µg/mL kanamycin. The plates were incubated at 37 °C in a microbiological incubator overnight. The following morning, plates were checked visually for single colonies and single colonies were picked immediately for protein expression or plates were stored at 4 °C for up to two weeks until further use. When no *E. coli* cell colonies grew overnight on the agar, the protocol was repeated and only 150 µL SOC medium were added to the cells and the whole solution was spread on the agar plate.

E. coli cells containing a plasmid with BirA gene under the control of an Ara promoter and with a chloramphenicol resistance were further transformed with avi-tagged *T. brucei* FPPS and avi-tagged human FPPS as described here.

2.6.3 Protein expression of *T. brucei* FPPS and human FPPS

Transformation of *T. brucei* FPPS and human FPPS was performed as described in Chapter 2.6.2. Two to three colonies were inoculated in 2 mL LB with 50 µg/mL kanamycin and incubated at 37 °C for 6 hours at 200 rpm in a shaking incubator. Subsequently, this preculture I was inoculated in 40 mL MDG medium with 50 µg/mL kanamycin and incubated at 200 rpm and 37 °C in shaking incubator overnight. On the next day the OD_{600} of preculture II was at around 7 and it was inoculated in 1.5 L modAl medium with 50 µg/mL kanamycin and 1.5 mL PPG antifoam in a fermentor (37 °C, 900 rpm, pH 7.0, pO_2 80 %) until an OD_{600} of around 10 was reached. Then the temperature was decreased to 18 °C for overnight growth. On the next day, the OD_{600} reached around 60 – 70 and the cells were harvested in harvesting bags and centrifuged for 30 min at 4 °C and 3500 rpm (SORVAL RC3BP, H600A). The supernatant was discarded and the cell pellet stored at - 80 °C until further use. The fermenter was cleaned using an automatic LabCIP program.

2.6.4 Cell lysis

Cell pellets (15 – 70 g) were thawed and resuspended in lysis buffer (250 – 500 mL, approximately 100 mL per 15 g cell pellet) containing one tablet of cOmplete EDTA-free protease inhibitor cocktail per 100 mL lysis buffer. Cells were homogenised with a polytron and further lysed by french press in four cycles. Cell lysate was cleared by centrifugation for 1 hour at 4 °C and 10 000 rpm (Avanti J30 I, JLA-16.250). The supernatant containing the soluble protein was filtered through a Stericup filter unit, 0.45 µm with an additional glass fibre prefilter. The filtered lysate was further purified by Ni^{2+} affinity chromatography and size exclusion chromatography (SEC).

2.6.5 Protein purification

Cell lysates (for preparation, see Chapter 2.6.4) containing the target protein were purified by Ni^{2+} affinity chromatography. Protein purification was carried out on an automated multi-step purification system (ÄKTAxpress device). The ÄKTAxpress system was equipped with one 5 mL nickel-nitrilotriacetic acid (Ni-NTA) Superflow column to separate the His_6-tagged protein with high affinity for the Ni^{2+} charged resin from proteins lacking a tag. First, the system was equilibrated with Ni^{2+} affinity buffer A and lysate was loaded on the column. Then bound protein was eluted with an imidazole gradient from 100 % Ni-NTA buffer A to 100 % Ni-NTA buffer B. The Ni^{2+} column was followed by a 120 mL HiPrep 26/10 desalting column for buffer exchange to buffer A. The flow through was collected in a falcon tube containing 0.25 mg HRV 3C protease in 1 mL buffer A. The protein was incubated for at least 16 hours and successful cleavage was confirmed by UPLC-MS. Then the purification mixture was loaded on a second Ni^{2+} column to remove uncleaved target protein and His_6-tagged HRV 3C protease. Fractions were collected in a 2 mL Deepwell block and collected samples were analysed by UPLC-MS and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Fractions with pure protein were pooled and concentrated in a 15 mL centrifugal filter with a 30 kDa MWCO to a maximum concentration of 30 mg/mL.

The target protein was further separated from high molecular aggregates by purification on a 120 mL size exclusion column connected to an ÄKTA Avant 25 system. The system was pre-equilibrated with SEC Tris buffer and 2 – 3 mL protein solution was injected on the column. Fractions were collected in a 2 mL Deepwell block and the maximum protein peak absorbance was read on the chromatogram. Fractions with a 10 % absorbance of the maximum protein peak absorbance before and after the protein peak were separated. Pure fractions were combined and protein concentration was determined by measuring the absorbance at 280 nm wavelength on a NanoDrop. Protein solution was either further concentrated in a 15 mL centrifugal filter with a

30 kDa MWCO or immediately analysed by SDS-PAGE, HPLC, and UPLC-MS before being flash-frozen in liquid nitrogen and stored in aliquots at - 80 °C until further use.

2.6.6 $^{13}C^{15}N$ *T. brucei and* $^{13}C^{15}N$ *human* FPPS expression and purification

Protein expression of $^{13}C^{15}N$ *T. brucei* and $^{13}C^{15}N$ human FPPS was performed as described in Chapter 2.6.3 with the following deviations: Preculture II and protein expression was done in $^{13}C,^{15}N$ uniform labelling solution. Therefore, 50 mL 20 × P-labelling solution, 2 mL 1 M $MgSO_4$, 4 g ^{13}C-D - glucose, 2.50 g $^{15}NH_4Cl$ and 1 mL of 1000 × kanamycin stock was dissolved in 938 mL distilled water and filtered through a Stericup 0.22 µm filter unit. While filtering 10 mL 100 × vitamins mix and 0.2 mL 1000 × trace metals were added to the solution. For preculture II 100 mL of the freshly prepared $^{13}C,^{15}N$ Uniform labelling solution were inoculated with preculture I and incubated in a shaking incubator overnight (200 rpm, 37 °C). The OD_{600} reached around 6 overnight and preculture II was inoculated in 900 mL $^{13}C,^{15}N$ Uniform labelling solution with 1 mL PPG antifoam in a fermenter (37 °C, 900 rpm, pH 7.0, pO_2 80 %). After around 2 – 3 hours, OD_{600} of around 3 was reached and the temperature was reduced to 18 °C. Then 6 g ^{13}C-D – glucose dissolved in 20 mL milliQ were added and 5 min later protein expression was induced with 0.5 mM IPTG (500 µL 1000 × IPTG stock). After overnight growth, OD_{600} around 15 was reached and cells were harvested in harvesting bags and centrifuged for 30 min at 4 °C and 3500 rpm (SORVAL RC3BP, H600A).

The labelling rate was calculated based on the measured mass in UPLC-MS and calculated mass for a 100 % labelling of the respective protein. The carbon and nitrogen count of the respective proteins was determined using ExPASy ProtParam tool.[492]

Cell lysis and protein purification was performed according to the protocols described in Chapter 2.6.4 and 2.6.5 with the following deviations: SEC purification runs were performed in SEC Tris buffer for the first $^{13}C^{15}N$ labelled *T. brucei* FPPS protein batch and for a $^{13}C^{15}N$ *T. brucei* FPPS resupply SEC purification was performed in SEC BisTris buffer to allow NMR experiments without an additional buffer exchange.

2.6.7 Avi-tagged in-vivo biotinylated *T. brucei* and human FPPS expression and purification

Protein expression was performed as described in Chapter 2.6.3 with the following deviations: Two single colonies were picked from a LB agar plate and directly inoculated in 40 mL MDG medium with 50 µg/mL kanamycin and 36 µg/mL chloramphenicol for preculture II and incubated at 37 °C and 200 rpm in a shaking incubator overnight. On the next morning, preculture II was inoculated in 1.5 L modAl medium with 50 µg/mL kanamycin and 36 µg/mL

chloramphenicol and 1.5 mL PPG antifoam in a fermenter (37 °C, 900 rpm, pH 7.0, pO_2 80 %). After around 2 hours the OD_{600} reached 3 – 4 and the expression of BirA was induced with 4 g/L of L-arabinose (freshly prepared, 100 × solution in 15 mL milliQ H_2O) and 200 µM biotin (20 mM stock solution prepared from milliQ H_2O and NaOH) was added shortly after. The temperature was decreased at OD_{600} of around 10 to 18 °C and for overnight growth.

Cell lysis and protein purification was performed following the protocols described in Chapter 2.6.4 and 2.6.5. Biotinylation rate was determined from the detected mass in UPLC-MS.

2.6.8 Human Rhinovirus (HRV) 3C protease expression and purification

Protein expression and cell lysis was performed as described in Chapter 2.6.3 and 2.6.4. Protein purification was carried out on an ÄKTA Avant 25 system equipped with a 5 mL HisTrap™ HP crude column. First, the system was equilibrated with Ni-NTA buffer A and lysate was loaded on the column. Bound protein was eluted with an imidazole gradient from 100 % Ni-NTA buffer A to 100 % Ni-NTA buffer B. Fractions were collected in a 2 mL Deepwell block and fractions containing the His-tagged target protein according to SDS-PAGE and UPLC-MS were combined in a falcon tube. Buffer of the protein sample was changed to SEC Tris buffer in a 15 mL centrifugal filter with a 30 kDa MWCO. Then HRV 3C protease was purified by SEC following the protocol described in Chapter 2.6.5.

2.6.9 Protein characterisation and concentration determination

Protein concentration during purification processes and of final protein samples was determined by HPLC and by measuring the absorbance at 280 nm on NanoDrop. Proteins were further characterised by SDS-PAGE, UPLC-MS and in some cases by DLS measurements.

2.6.9.1 NanoDrop

Absorbance at 280 nm (detecting tryptophanes and other aromatic groups) was measured on a NanoDrop. A new experiment was opened and the method for determining the absorbance at 280 nm was chosen. The measurement type 1 Abs = 1 mg/mL was further selected and baseline correction at 340 nm default setting was left. The pedestals were cleaned with milliQ H_2O and 2 µL of buffer or the protein sample were pipetted on the detection surface for a blank measurement. Afterwards, the protein sample was measured in a triplicate and the average mg/mL concentration was calculated. The correction factor for the absorbance at 280 nm was computed on ExPasy ProtParam[492]. The calculated average mg/mL concentration of the triplicate was divided by the absorbance correction factor and the final concentration was determined.

2.6.9.2 HPLC

HPLC runs were performed on an AGILENT TECHNOLOGIES 1290 Infinity II LC System with a 100 × 2 mm column packed with POROS R1, 10 µm (Dr. Maisch GmbH) using the AGILENT TECHNOLOGIES OpenLAB CDS Workstation Software Plus, MassHunter Walkup Software and 1290 Infinity II Walkup Software. The HPLC system is calibrated with BSA as a reference for protein concentration determinations and runs on standby in 80 % buffer A (milliQ H_2O and 1 % v/v TFA) and 20 % buffer B (90 % v/v acetonitrile and 0.1 % v/v TFA). Around 3 µg protein sample can be loaded on the column and 30 µL were injected. Therefore, the protein concentration was first measured via NanoDrop to determine the sample dilution factor and then protein samples were diluted to around 100 µg/mL with milliQ H_2O. Triplicates were prepared for each sample and a 10 min elution method from 20 % buffer B to 100 % buffer B was run after sample application on the column. Before and after each run, the system was washed with 30 µL guanidine cleaning solution. The average value of the duplicate or triplicate was calculated and compared to the results of the NanoDrop. Data were processed automatically and manually re-integrated if necessary.

2.6.9.3 SDS-Polyacrylamide Gel electrophoresis (PAGE)

SDS-PAGE[493-494] was performed to assess purity and amount of protein in different fractions of the purification step and for final protein samples. Purification fractions or final protein samples were mixed 1:4 with 4 × SDS sample buffer and incubated for 5 min at 95 °C in a thermomixer. For comparison of sizes of separated proteins, Precision Plus Protein Kaleidoscope Protein Standard (BioRad) was loaded in the first and last line. Furthermore, 3 – 4 µL of each sample were loaded in the remaining wells of a 4 – 20% Precast Protein Gel with 15-wells. The gel was mounted in the Mini-Protean Tetra-System and the chamber was filled with 1 × SDS running buffer. Subsequently, the chamber was connected to the electrophoresis power source and the proteins migrated through the gel for 30 – 40 min at 200 V until the dye front reached the bottom of the gel. The gel was stained in Instant Blue™ Coomassie staining solution overnight to visualise protein bands. Then, the gel was incubated with water twice for at least 10 min to reduce background staining. Finally, the gel was imaged with a GelDoc™ EZ Imager for documentation.

2.6.9.4 Ultrahigh pressure liquid chromatography – Mass spectrometry (UPLC-MS)

Mass spectrometry was performed on an UPLC-MS ESI system to assess the quality and purity of expressed proteins. The protein sample (1 – 2 µL of 3 – 15 mg/mL protein) was applied onto an Acquity UPLC BEH C_4, 2.1 × 100 mm column, 1.7 µm at 80 °C. The eluents consisted of buffer A (water + 0.05 % TFA) and buffer B (acetonitrile + 0.04 % TFA). A 10 min protein elution

method was run with a 5 % to 60 % B gradient in 8 min with a flow rate of 0.5 mL/min. The UPLC chromatography was followed by positive-ion electrospray ionization (ESI) and application onto a Xevo-G2-S QT Quadrupole Time of flight mass spectrometer. Mass spectra were acquired over a mass range from 700 m/z to 3000 m/z with a deconvolution range from 10 – 150 kDa using Maximum Entropy (MaxEnt).

2.7 MDMix hot spot identification

The pyMDMix software was readily available in Xavier Barril's lab at the University of Barcelona and a full description was published by Alvarez-Garcia *et al.* in 2014.[495] All following steps were performed in collaboration with PhD students Serena Piticchio, Moira Rachman and Maciej Majewski and here, a short description is given. The mixed solvent molecular dynamic simulation was performed on a protein structure (PDB ID 2I19, *T. brucei* FPPS in complex with bisphosphonate) downloaded from the PDB and a previously solved in-house apo *T. brucei* FPPS structure. First, the protein was protonated and the termini were capped using the molecular operating environment (MOE) software. Then the system was prepared by generating an assisted model building with energy refinement (AMBER) object file with Leap and parameterized using AMBER force field 99SB. The protein was placed in an octahedral box with a margin of 13 Å away from the protein in all directions with a pre-equilibrated solvent mixture of 20 % ethanol in water. The system was minimised for 3000 steps and the initial velocities were chosen to obtain a distribution at 100 K. The protein and solvent were further equilibrated by heating up to 300 K and then pressure correction where the system reaches 1 atm of pressure was performed. Three independent simulation replicas of 50 ns production runs were carried out with the atomic coordinates being stored every picosecond. Restraints were applied on heavy atoms as soft harmonic potentials with a restraining force of 0.01 kcal/mol.A^2.

The three replicas were aligned excluding the caps and using the following sequences: 2 – 62; 70 – 98; 105 – 160; 166 – 330 and 3 – 63; 76 – 110; 119 – 230; 238 – 362; 372 – 429; 444 – 476; 496 – 548; 570 – 732 for the in-house *T. brucei* FPPS apo protein and PDB ID: 2I19 structure, respectively. Density maps of replicas were generated. The atomic coordinates stored during the simulation were overlaid on the atoms of the protein structure and a static quadratic grid of 0.5 Å size in each direction was built for CH_3 (hydrophobic), OH (hydrogen-bond donor and acceptor) and water. The density grid was constructed based on the number of times the respective atom is found in each grid and the binding free energy of the atoms was computed by the inverse Boltzmann equation. The energy grids were inspected in PyMOL.

2.8 Nuclear Magnetic Resonance (NMR)

2.8.1 Instrumentation

Spectra were recorded on a Bruker Avance III HD 600 MHz NMR spectrometer equipped with a 5 mm triple resonance inverse (TCI) cryoprobe ^{1}H, ^{13}C, ^{15}N with a ^{2}H lock and z-gradient, operating at a frequency of 600.23 MHz for ^{1}H or on a Bruker Avance III HD 800 MHz NMR spectrometer equipped with a 5 mm TCI cryoprobe ^{1}H, ^{13}C, ^{15}N with a ^{2}H lock and z-gradient, operating at a frequency of 800.19 MHz for ^{1}H or on a Bruker Avance III 600 MHz quadruple equipped with a 5 mm QCI cryoprobe ^{1}H, ^{19}F, ^{13}C, ^{15}N with a ^{2}H lock and z-gradient operating at a frequency of 600.13 MHz for ^{1}H. The cryoprobe is cooled at a few Kelvin, thus reducing the electronic noise considerably and increasing the sensitivity of the magnet. All spectrometers were equipped with a SampleJet™ sample changer with five cooled positions for a 96 SampleJet Rack with 3 mm Bruker® SampleJet NMR tubes.

2.8.2 General procedure

Nuclear magnetic resonance was performed in a fragment screening approach using ligand-observed experiments for the Novartis core and fluorine fragment library and further protein-observed NMR was used as a fragment hit validation method.

Each NMR sample contained DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) at either 100 µM or 150 µM concentration as an internal control and 10 % v/v D_2O. Compounds used for NMR experiments were dissolved in 90 % v/v d_6-DMSO and 10 % v/v D_2O and stored as 100 mM or 500 mM compound stocks. NMR spectra were acquired at 296 K if not otherwise stated.

Before each set of experiments, the first NMR sample was transferred to the magnet and the magnetic field was adjusted using the field-frequency lock on D_2O to compensate the magnetic field drifting and to hold the magnetic field homogeneity stable. Afterwards, spatial inhomogeneity of the main magnetic field was cancelled out by shim coils to obtain the required homogenous field across the whole sample. Moreover, the shimming process was followed by “tuning and matching” of the probe to achieve an optimal-signal-to-noise ratio. Tuning and matching was performed on ^{1}H, ^{13}C, ^{15}N nuclei to match the tuned circuit with the Larmor frequency and to optimise the power transfer from the probe to the transmitter and receiver. Furthermore, the 90 ° pulse was calibrated to flip all protons exactly 90° from the Z-axis into the XY plane. Additionally, if necessary, the soft pulse for water suppression[496] was calibrated.

For data acquisition, processing and analysis TopSpin software was used and measurement of numerous samples was performed using a SampleJet in combination with IconNMR.

2.8.3 Fragment screen of Novartis core fragment library by ligand-observed NMR

The Novartis core fragment library was screened in mixtures of eight compounds with each compound at 20-fold excess against 10 µM *T. brucei* FPPS on an Avance III 600 MHz. For this, 5.7 µL of each compound mixture in d_6-DMSO of the Novartis core fragment library were ordered in two 96 well plates. *T. brucei* FPPS protein was buffer exchanged to NMR core fragment library screening buffer in a 15 mL centrifugal filter with a 30 kDa molecular weight cut-off (MWCO) and the protein concentration was determined by measuring the absorbance at 280 nm on a NanoDrop. Then, an NMR master mix containing NMR core fragment library screening buffer with 10 % v/v D_2O, 150 µM DSS and 10 µM *T. brucei* FPPS protein was prepared in a 50 mL Falcon tube. To each plate well with 5.7 µL compound mixture 174.3 µL of NMR protein screening solution were added and one control sample with 5.7 µL d_6-DMSO was prepared. The sample plate was mixed on a plate shaker (1250 rpm, 5 min, 4 °C) and then 175 µL of each mixture were transferred to a 3 mm NMR tube with the TECAN robot. The SampleJet Rack was loaded in the SampleJet and a 1D ^{1}H-zgesgp (128 scans), a T1ρ 10 ms and 200 ms (each 128 scans) and a wLOGSY[386] (256 scans) experiment was performed on an Avance UltraShield 600 MHz to obtain binding information. Acquisition times were approximately 4 min, 6 min, 6 min and 18 min, respectively, resulting in a full experiment acquisition time of 4.3 days. Reference spectra of compound mixtures without protein were kindly shared by Chrystelle Henry, NIBR, Basel, Switzerland.

Data were analysed in TopSpin and ASSOC files for each wLOGSY and T1ρ experiments were generated. For analysis of wLOGSY experiments, the spectra with protein were compared with reference spectra and if a compound showed a positive wLOGSY signal that was not already positive in the reference spectrum the compound fulfilled hit criteria for the wLOGSY method. For analysis of T1ρ experiments, the T1ρ 10 ms and T1ρ 200 ms spectra were overlaid and faster relaxation of a compound bound to the protein is obtained resulting in broadened and smaller peaks in T1ρ 200 ms compared to T1ρ 10 ms. The effect was noted and a hit for the T1ρ method was defined as showing a ≥ 20 % effect compared to the reference spectra. A true fragment hit was defined as a wLOGSY and T1ρ hit. Fragment hits were further followed-up by NMR single compound hit validation to exclude mixture effects.

For single compound experiments, compounds were again ordered from the Novartis compound storage and each compound was tested at 200 µM against 10 µM *T. brucei* FPPS on an Avance™ III HD 600 MHz as described for fragment mixtures. This time, an NMR master mix without protein was prepared and a 1D ^{1}H-zgesgp (128 scans), a T1ρ 10 ms and 200 ms (each 128 scans) and a wLOGSY[386] (256 scans) reference spectrum was first recorded and then protein was added to each sample with a TECAN robot for protein spectra. Identified fragment hits were further validated in protein-observed NMR experiments.

2.8.4 Fragment screen of Novartis fluorine library by ligand-observed NMR

The Novartis fluorine library was screened in mixtures of 30 compounds at 18 µM against 3.7 µM *T. brucei* FPPS protein on an Avance™ III HD 600 MHz, quadruple. Therefore, 2 µL of each mixture were ordered in a 96 well plate from the Novartis compound storage. First, an NMR sample buffer solution by addition of 10 % v/v D_2O and 100 µM DSS to BisTris NMR buffer was prepared. To each well containing 2 µL of fragment mixture 178 µL NMR sample buffer solution was added. Additionally, one sample with 2 µL BisTris buffer and one sample with 2 µL DMSO were prepared in the same way as control samples. The solutions were transferred to 3 mm SampleJet NMR tubes. A 1D ^{1}H-zgesgp (128 scans), a ^{19}F-CPMG 80 ms and 400 ms (each 512 scans) experiment was recorded for each NMR sample as reference spectra. Acquisition time was approximately, 4.8 min, 18.3 min, and 21.2 min, respectively, resulting in a 15 hours of overall acquisition time.

T. brucei FPPS protein stock was filtered through a 0.45 µm Eppendorf filter and pre-diluted in BisTris NMR buffer and the protein concentration was determined by measuring the absorbance at 280 nm wavelength on a NanoDrop. Then, 4.4 µL of pre-diluted protein solution were added to each NMR tube with the TECAN robot. This time, a 1D ^{1}H-zgesgp (128 scans), a ^{19}F-CPMG 80 ms and 400 ms (each 512 scans) experiment was recorded for each sample with protein.

For data analysis, a test version of the fragment-based screening (FBS) tool from Bruker was used. A peak list, reference spectra of the compound mixtures and spectra of samples with protein were loaded in the tool and an overlay of ^{19}F-CPMG 80 ms and 400 ms spectra was generated. Binding compounds were presented with a broadened and smaller peak in the ^{19}F-CPMG 400 ms compared to the ^{19}F-CPMG 80 ms spectrum. The peak ratios of the 400 ms to 80 ms spectrum in the samples with protein were compared to the peak ratios of the 400 ms and 80 ms spectrum of the reference. Fragment hits showing a $\geq$ 20 % ^{19}F-CPMG effect in the spectrum with protein compared to the reference were further validated in single compound experiments and in protein-observed [^{13}C,^{1}H]-SOFAST-HMQC experiments.

Single compound experiments were performed in the same way as described for compound mixtures. Compounds were ordered from the Novartis compound storage and reference spectra and protein spectra were recorded on an Avance III HD 600 MHz, quadruple. For data analysis, ASSOC files were generated as described in Chapter 2.8.3 and validated fragment hits were followed-up by protein-observed NMR.

2.8.5 Protein-observed NMR

Fragments identified in ligand-observed NMR fragment screens of the Novartis core and fluorine fragment library, an X-ray fragment screen at Diamond, Oxford and an X-ray fragment

screen at EMBL/ESRF, Grenoble as well as available human allosteric binders, fragment analogues of the Novartis compound storage and externally ordered from Enamine and synthesised fragments as part of the PhD thesis of Joy Petrick, NIBR, Basel, Switzerland were tested in protein-observed [^{13}C, ^{1}H]-SOFAST-HMQC experiments.

For fragments of all sources except for fragments of the fluorine library, protein-observed [^{13}C, ^{1}H]-SOFAST-HMQC experiments (32 scans, 42 min acquisition time) were recorded with 1 mM compound against 30 µM protein on an Avance III HD 800 MHz NMR spectrometer at 310 K. However, some fragments, especially synthesised fragments showed solubility issues and the soluble fraction of the compound in the sample was lower than the target concentration. Fragment hits of the Novartis fluorine library were measured against 30 µM protein and at 700 µM fragment concentration. Before and after each [^{13}C, ^{1}H]-SOFAST-HMQC experiment a 1D ^{1}H-zgesgp (512 scans, 10 min acquisition time) was set up for controlling protein precipitation, aggregation, compound concentration and integrity over the course of the 2D-NMR experiments.

First, ^{13}C^{15}N-labelled *T. brucei* FPPS was buffer exchanged to NMR BisTris buffer in a 15 mL centrifugal filter with a 30 kDa MWCO and protein concentration was determined by NanoDrop. Then, a NMR master mix containing 30 µM protein in NMR BisTris buffer with the addition of 10 % D_2O, 150 µM DSS was prepared. Compounds from 100 mM and 500 mM compound stock solutions in 90 % v/v d_6-DMSO and 10 % v/v D_2O were added to the NMR master mix to give a final compound concentration of 1 mM and a final sample volume of 170 µL. Samples were vortexed, spun down and transferred to NMR sample tubes. In addition, an apo-protein and DMSO control sample were prepared in the same way as reference spectra.

Data were processed and analysed in TopSpin and protein spectra were overlaid with reference spectra to observe chemical shift perturbations. Peak shifts observed due to DMSO effects were not considered for the determination of compound binding effects. Binding compounds were categorised as very weak (1 to 3 shifts, below 0.005 ppm), weak (3 to 5 shifts below 0.005 ppm), medium (1 to 7 shifts below 0.005 ppm and up to 3 shifts between 0.006 and 0.02) and strong shifters (1 to 7 shifts below 0.005 ppm and more than 3 shifts with at least 0.02 ppm) according to number of shifted peaks and extend of chemical shift difference.

2.8.6 K_d by NMR

K_ds were estimated by recording [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans, 42 min acquisition time) spectra with various compound concentrations against 30 µM ^{13}C^{15}N-labelled *T. brucei* FPPS protein on an Avance III HD 800 MHz NMR spectrometer. A serial compound dilution of the compound stock in 90% v/v d_6-DMSO and 10 % v/v D_2O solution was prepared to add the same d_6-DMSO concentration to each individual NMR sample. Additionally, an apo protein sample without d_6-DMSO and a d_6-DMSO control sample with the same amount of d_6-DMSO as added for

the compound samples was prepared. NMR samples were prepared as described in Chapter 2.8.5. Again before and after each [^{13}C, ^{1}H]-SOFAST-HMQC experiment a 1D ^{1}H-zgesgp (512 scans, 10 min acquisition time) was additionally recorded.

Data were processed in TopSpin and an in-house Linux script named fitKD written by Armin Widmer prepared K_d plots. Therefore, a simple in-put file with file locations of the spectra of the d_6-DMSO blank and corresponding compound concentrations was written. Curves were generated by plotting the chemical shift versus the ligand concentration. Curve fits and K_ds were displayed automatically by fitKD. K_d values were calculated of several peaks to compare if derived values fall into the same range. The final K_d values are an average of the K_d values of between 2 to 8 NMR signals of the ^{1}H dimension considered in the averaging.

2.8.7 Solubility measurements of compounds

For SPR experiments prior determination of solubility of compounds in SPR buffer was crucial to avoid destroying the SPR chips by compound precipitation. Therefore, 1D ^{1}H-zgesgp spectra (128 scans, 4 min acquisition time) were recorded at a 1 mM compound concentration in SPR buffer on an Avance UltraShield 600 MHz to determine the solubility threshold based on DSS as an internal standard.

NMR samples were prepared by addition of 1.7 µL compound to 170 µL solubility buffer. Signals of aromatic protons were counted and normalised to the DSS peak at 150 µM concentration and compared with compound structure for agreement of multiplets and intensities and to estimate solubility.

2.9 Crystallization at Novartis, Basel

2.9.1 General crystallization protocol

Crystallization experiments were performed with *T. brucei* FPPS protein at various concentrations in low and high salt buffer in a temperature controlled room at 20 °C (Table 14). Buffer exchange was performed in 15 mL centrifugal filter units with a 30 kDa MWCO in three cycles of fresh buffer addition. Protein aliquots stored at - 80 °C were quickly thawed in hand and stored on ice until plate preparation.

Table 14: List of protein stock solutions with buffer compositions.

Protein concentration [mg/ml]	Buffer composition	Salt conc.
4.88 mg/mL	50 mM Tris-HCl, 200 mM NaCl, 2 mM TCEP-HCl, pH 8	high
5.06 mg/mL	10 mM Tris-HCl, 10 mM NaCl, 2 mM TCEP-HCl, pH 8	low
11.50 mg/mL	10 mM Tris-HCl, 10 mM NaCl, 2 mM TCEP-HCl, pH 8	low
12.24 mg/mL	50 mM Tris-HCl, 200 mM NaCl, 2 mM TCEP-HCl, pH 8	high
13.81 mg/mL	10 mM Tris-HCl, 10 mM NaCl, 2 mM TCEP-HCl, pH 8	low

Crystallization experiments were set up on Swissci 2-drop sitting drop vapour diffusion plates or on 24-well VDXm hanging drop vapour diffusion plates. Commercially available crystallization screens were stored in the fridge and 80 µL per well were transferred to Swissci 2-drop plates with a liquidator. Optimised crystallization conditions were pipetted as reservoir solution on Swissci 2-drop plates (80 µL solution per well) or 24-well VDXm plates (500 µL solution per well) with a Formulator using commercially available or in-house prepared stock solutions (see Chapter 2.5.5). Volatile solutions, such as DMSO were pipetted manually to the reservoir solution and solutions were mixed by pipetting up and down with a multichannel pipette. Crystallization drops on Swissci 2-drop plates were pipetted with a Mosquito robot in various protein solution, reservoir and microseed stock ratios (Table 15).

Table 15: Different protein, reservoir, microseed ratios pipetted on Swissci 2-drop plates.

Protein solution [nL]	Reservoir [nL]	Microseed stock [nL]
300	200	-
200	200	100
300	200	100

Crystallization drops for 24-well VDXm plates were pipetted manually on 18 mm cover slides in various protein solution, reservoir and microseed stock ratios (Table 16). Then, the cover slide was put upside down on top of the well to equilibrate the drop against the reservoir solution.

Table 16: Different protein, reservoir, microseed ratios pipetted on VDXm plates.

Protein solution [µL]	Reservoir [µL]	Microseed stock [µL]
1.2	0.8	-
1	1	-
0.8	1.2	0.4
1	1	0.5

Microseeding stocks were prepared by crushing and combining fully-grown crystals in a seed bead tube. The tube was filled up to 50 – 1000 µL with the respective reservoir solution and the solution was vortexed 5 × 1 min and cooled on ice in between. If not otherwise stated, microseeding solutions of crystals grown in 6 wells of a VDXm plate from condition Cryos Suite C3 (1.6 M $NH_4H_2PO_4$, 0.08 M Tris HCl, pH 8.5, 20 % v/v glycerol) diluted to 1 mL in reservoir solution was used for the setup of crystallization plates. Several new microseed stock solutions were prepared over the course of the work and for each microseed stock solution the optimal dilution ranging between 1:100 (v/v) and 1:1000 (v/v) in reservoir solution was determined. These microseeding stock solutions were stored at RT up to 4 weeks. For longer storage, the solution was flash frozen in liquid N_2 and stored at - 80 °C. Storage at 4 °C leads to salt precipitation and was therefore avoided.

Crystal plates were incubated and imaged in the Rock Imager 1000 Dual Imager at 20 °C. UV images of crystals were taken immediately after putting plates, after 12 hours, 1 day 12 hours, 2 days 12 hours, 3 days 12 hours, and periodically up to three months. Images were inspected in Rockmaker. Crystals were harvested without the addition of a cryoprotectant if not otherwise stated. Crystals were mounted on a 0.1 – 0.2 mm or 0.05 – 1 mm CryoLoop, flash-frozen in liquid N_2 and stored until data collection.

2.9.2 Reproducing published crystallization conditions

Published[229] crystallization conditions of *T. brucei* FPPS protein were attempted to be reproduced on a VDXm plate (Table 17). For each condition 4 wells with crystallization drops of 1 µL protein at 4.88 mg/mL in high salt buffer and 1 µL reservoir solution and 4 wells with 1 µL protein and 0.5 µL reservoir solution were pipetted.

Table 17: Composition of published crystallization conditions.

Condition	Composition
Crystallization condition 1	0.2 M potassium citrate tribasic monohydrate, 20 % w/v PEG 3350
Crystallization condition 2	0.1 M HEPES pH 7.5, 10 % w/v PEG 8000, 8 % v/v ethylene glycol
Crystallization condition 3	0.1 M sodium cacodylate pH 6.5, 0.2 M Mg acetate tetrahydrate, 20 % w/v PEG 8000

2.9.3 Crystallization screens and first crystal optimisations

Three rounds of commercially available crystallization screens were performed on Swissci 2-drop plates with crystal drops of 300 nL protein and 200 nL reservoir solution (Table 18). For suppliers of the different commercially available screens employed, see Chapter 2.5.5. In the

first and second screening round 4.88 mg/mL protein in high salt buffer and in the third screening round 5.06 mg/mL protein in low salt buffer was used.

Table 18: Crystallization screens of 1st – 3rd screening round.

1st screening round	2nd screening round	3rd screening round
$AmSO_4$ Suite	Index HT	PEG/Ion HT
Cryos Suite	MbClass Suite	PEGsII Suite
JCSG+ Suite	PACT Suite	PEGRx HT
MbClass II Suite	pHClear II Suite	PEGs Suite
Index HT	PEG/Ion HT	
PEGRx HT	PEGsII Suite	
SaltRx HT	PEGs Suite	

Cryos Suite condition C3 and JCSG+ Suite condition C1 of the first screening round were chosen for optimisation with crystal drops of 1.2 µL protein and 0.8 µL reservoir on VDXm plates. Buffer strength and pH, precipitant and salt concentration were varied as shown in Table 19.

Table 19: Optimisation of condition Cryos Suite C3 and JCSG+ Suite C1.

Condition	Components
Cryos Suite C3	0.08 M Tris-HCl pH 8.5, 1.6 M $NH_4H_2PO_4$, 20 % v/v glycerol
Variation	0.08 M Tris-HCl pH 7 – 8.5, 1 – 1.6 M $NH_4H_2PO_4$, 20 % v/v glycerol
JCSG+ Suite C1	0.2 M NaCl, 0.1 M phosphate-citrate pH 4.2, 20 % w/v PEG 8000
Variation	0.1-0.4 NaCl, 0.1 M phosphate-citrate pH 4.2, 15-25 % w/v PEG 8000

Furthermore, condition PEG/Ion E4 (0.2 M sodium malonate pH 5, 20 % w/v PEG 3350) and condition PEG/Ion HT E12 (8 % v/v tacsimate, pH 5, 20 % w/v PEG 3350) of the 2nd screening round were set up with crystal drops of 1.2 µL protein and 0.8 µL on VDXm plates. Crystallization condition components were varied (Table 20).

Table 20: Overview of sodium malonate and tacsimate crystallization condition variations.

Buffer conc.	pH	PEG and PEG conc.
0.2 M sodium malonate	5	20 % w/v PEG 3350
	4.5 – 5 – 5.5 – 6	20 % w/v PEG MME 2000 20 % w/v PEG MME 5000 15 % w/v PEG 3350 20 % w/v PEG 6000 20 % w/v PEG 10000
0.05 – 0.1 – 0.2 M sodium malonate	5	20 % w/v PEG 3350 10% w/v PEG 10000 10 % w/v PEG 20000 15 % w/v PEG MME 5000
8 % v/v tacsimate	5	20 % w/v PEG 3350
	4.5 – 5 – 5.5 – 6	20 % w/v PEG MME 2000 20 % w/v PEG MME 5000 15 % w/v PEG 3350 20 % w/v PEG 6000 20 % w/v PEG 10000

2.9.4 Crystal screens at high protein concentration and with microseed matrix technique

Crystals of the first screening rounds from the following conditions: 1) 0.05 M calcium chloride, 45 % v/v NaOAc, 0.1 M MES pH 6, 2) 12 % w/v PEG 3350, 4 % w/v tacsimate pH 5, 3) 10 % w/v PEG 10000, 0.1 M sodium malonate pH 5 and 4) 0.08 M Tris-HCl, pH 8.5, 1.6 M $NH_4H_2PO_4$, 20 % v/v glycerol were used for the preparation of microseeding stocks of crystals from one well in 100 µL of respective reservoir solution. In a fourth screening round crystallization drops of 200 nL protein at 5.06 mg/mL in low salt buffer, 200 nL reservoir and 100 nL microseeds were set up on Swissci 2-drop plates. A fifth screening round was performed with crystallization drops of 200 nL protein at 13.81 mg/mL in low salt buffer and 200 nL reservoir solution on Swissci 2-drop plates (Table 21).

Table 21: Crystallization screens of 4th – 6th screening round.

4th screening round	5th screening round	6th screening round
PEG/Ion HT	$AmSO_4$ Suite	Cryos Suite
PEGRx HT	Cryos Suite	Index HT
Cryos Suite	JCSG+ Suite	JCSG+ Suite
Index HT	PEG/Ion HT	PEG/Ion HT
JCSG+ Suite	Index HT	PEGs Suite
PEGs Suite	PEGs Suite	PEGRx HT
	PEGRx HT	

Crystals of condition Cryos Suite C3 (1.6 M $NH_4H_2PO_4$, 0.08 M Tris HCl pH 8.5, 20 % v/v glycerol) from the first screening round were reproducible but with undesired crystal morphology. Therefore, a new microseeding stock from 6 wells of crystals grown in the same condition on a VDXm plate was prepared in 1 mL of reservoir buffer. This microseeding stock was used for all crystallization plates using microseed matrix screening from now on. A sixth screening round was performed with 200 nL protein at 13.81 mg/mL in low salt buffer, 200 nL reservoir and 100 nL microseeds on Swissci 2-drop plates.

2.9.5 Crystallization optimisation and additive screens

Further optimisation trials for crystals that appeared on crystal plates of the sixth screening round were performed on VDXm plates at 20 °C (Table 22). Various crystallization drop ratios as listed in Table 16, reservoir compositions and microseed stock dilutions were tested.

Table 22: Crystallization conditions with crystals for optimisation.

Screen	Well	Components
Index HT	D1	25 % w/v PEG 1500
Index HT	D7	0.1 M Bis-Tris pH 6.5, 25 % w/v PEG3350
Index HT	G2	0.2 M lithium sulfate monohydrate, 0.1 M Bis-Tris, 25 % w/v PEG 3350
Index HT	G11	0.2 M magnesium chloride hexahydrate, 0.1 M Bis-Tris pH 6.5, 20 % w/v PEG 3350
PEG/Ion HT	A1	0.2 M sodium fluoride, 20 % w/v PEG 3350
PEG/Ion HT	A2	0.2 M potassium fluoride, 20 % w/v PEG 3350
PEG/Ion HT	C7	0.2 M lithium sulfate monohydrate, 20 % PEG 3350
PEG/Ion HT	C9	0.2 M sodium sulfate decahydate, 20 % w/v PEG 3350
PEG/Ion HT	E6	0.2 M sodium malonate pH 6, 20 % w/v PEG 3350

Table 22: continued.

PEG/Ion HT	H9	0.15 M caesium chloride, 15 % w/v PEG 3350
PEGs Suite	G12	0.2 M sodium tartrate, 20 % w/v PEG 3350
PEGs Suite	H1	0.2 M sodium/potassium tartrate tetrahydrate, 20 % w/v PEG 3350
PEGs Suite	H7	0.2 M $NH_4H_2PO_4$, 20 % w/v PEG 3550
JCSG+ Suite	E9	0.1 M MES pH 6.5, 1.6 M $MgSO_4$
JCSG+ Suite	D2	0.2 magnesium chloride, 0.1 M HEPES pH 7.5, 30 % w/v PEG 400
JCSG+ Suite	B4	0.1 M HEPES, 10 % w/v PEG 8000, 8 % v/v ethylene glycol
JSCG+ Suite	F5	0.2 magnesium chloride, 0.1 M Tris-HCl pH 8.5, 45 % v/v ethylene glycol
Cryos Suite	F9	0.04 M KH_2PO_4, 20 % v/v glycerol, 16 % PEG 8000
Cryos Suite	G9	20 % v/v glycerol, 24 % w/v PEG 1500
Cryos Suite	G8	8.5 % w/v PEG 1000, 8.5 % w/v PEG 8000, 15 % v/v glycerol
Cryos Suite	G12	0.07 sodium acetate pH 4.6, 30 % v/v glycerol, 5.6 % w/v PEG 4000

Crystals of condition PEGs Suite G12 (0.2 M sodium tartrate, 20 % w/v PEG 3350) were reproducible and an additive screen HT was performed on Swissci 2-drop plates. Therefore, 16 µL of additive screen were added to 64 µL reservoir (25 % w/v PEG 3350 and 0.25 M sodium tartrate) on a Swissci 2-drop plate. Crystal drops of 300 nL protein at 13.81 mg/mL in low salt buffer, 200 nL reservoir and 100 nL microseeds were pipetted. In this manner, the additives in the screen were diluted 1:15 in the final drop. Furthermore, additive screens with condition PEG/Ion HT H9 (0.15 M caesium chloride, 15 % PEG 3350) and JCSG+ Suite D2 (0.2 magnesium chloride, 0.1 M HEPES pH 7.5, 30 % w/v PEG 400) were performed by adding 16 µL of additive screen HT to 64 µL reservoir (PEG/Ion HT H9 and JCSG+ Suite D2), which results in a final concentration of reservoir components (0.12 M caesium chloride, 12 % PEG 3350 and 0.08 M HEPES pH 7.5 20 % PEG 400, 0.12 magnesium chloride). Crystal drops of 200 nL protein at 13.81 mg/mL in low salt buffer, 200 nL reservoir and 100 nL microseeds were pipetted. Crystallization conditions with various additives were set up on new Swissci 2-drop plates to reproduce crystals (Table 23). Single crystals were harvested and sent for X-ray diffraction. When protein stocks were running low, additional crystallization plates with conditions marked in green (Table 23) were set up with protein after buffer exchange to low salt buffer at 11.5 mg/mL protein concentration from 12.24 mg/mL protein stocks. Crystals were harvested and sent for X-ray diffraction.

Table 23: List of crystallization conditions and respective additives that were set up on new plates.

Conditions		
0.2 M sodium tartrate 20 % w/v PEG 3350	0.08 M HEPES, 0.12 M magnesium chloride, 20 % w/v PEG 400	0.12 M caesium chloride, 12 % w/v PEG 3350
Additives		
6 % dioxane	0.02 M barium chloride	0.02 M magnesium chloride
8 % tert-butanol	0.02 M proline	0.02 M proline
	0.02 M EDTA	6 % v/v DMSO
	10 % PEG 400	8 % PE 15/4 EO/OH
		6 % xylitol
		6 % ethylene glycol
		6 % glycerol
		0.4 M NDSB-211

Crystallization condition (0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % v/v DMSO) was further optimised by DMSO concentration variations from 6 – 30 % v/v. Crystallization drops of 200 nL protein at 11.5 mg/mL, 200 nL reservoir and 100 nL microseeds were set up on Swissci 2-drop plates. A solution of 0.12 mM caesium chloride, 15 % w/v PEG 3350, 40 – 45 % v/v glycerol or 15 % w/v PEG 3350 and 35 % v/v ethylene glycol was added to the drop 1:2 (v/v) for 30 s – 2 min immediately before harvesting. After several successful crystallization plates with protein at 11.5 mg/mL in low salt buffer, crystals were not reproducible anymore. Protein of the same batch at 12.24 mg/mL in high salt buffer was used instead for crystallization experiments.

2.9.6 X-ray fragment soaking

Crystal soaking experiments were performed with fragments identified in NMR fragment screens of the Novartis core and fluorine fragment library, available human allosteric binders, fragment analogues of the Novartis compound storage and externally ordered from Enamine, synthesised fragments as part of this thesis and fragments synthesised by Joy Petrick, NIBR, Basel, Switzerland.

Soaking experiments were performed with fully-grown crystals from condition 0.12 M caesium chloride, 12 % w/v PEG 3350 with one of the following additives: 6 % v/v xylitol, 6 % v/v ethylene glycol or 6 % v/v DMSO. However, when crystal experiments were irreproducible

with conditions containing xylitol or ethylene glycol, only soaking experiments with crystals from condition containing DMSO were performed from then on.

Crystal soaks were set up on Swissci 2-drop plates following one of the following three protocols. Compounds were pipetted from 100 mM or 500 mM compound stock solutions in 90 % v/v d_6-DMSO and 10 % v/v H_2O and final crystal soaking drops contained up to 15.9 % v/v DMSO. Compound soaking concentrations and soaking times were varied from 5 mM to 75 mM and 1 min to 2 days, depending on crystal integrity and quality in the presence of the individual compounds.

Protocol 1

Compound stocks at either 100 mM or 500 mM compound concentration were pre-diluted in reservoir solution to the desired concentration. For reservoir solutions containing DMSO, a solution without DMSO was prepared for the preparation of pre-dilutions. Then, the compound pre-dilution was mixed with protein buffer solution (high salt buffer or low salt buffer, see Table 14) and microseed solution without seeds 1:1:0.5 (v/v). Of this soaking mixture 1 µL to 2 µL were pipetted on the second drop position of a Swissci 2-drop plate. Crystals were transferred with a CryoLoop from the first crystal drop position to the soaking drop and were incubated at 20 °C. Immediately before harvesting a cryosolution (12 % w/v PEG 3350 and 30 % v/v ethylene glycol or 40 – 45 % v/v glycerol) with the addition of compound at the same concentration as in the soaking mixture was added to the soaking drop in a 1:1 ratio for 30 s – 2 min.

Protocol 2

The soaking method as described in Protocol 1 was performed with the following deviations: Compound pre-dilutions were mixed with cryosolution (12 % w/v PEG 3350 and 30 % v/v ethylene glycol or 40 – 45 % v/v glycerol) instead of protein buffer solution and seed buffer solution 1:1:0.5 (v/v). No further cryosolution was added after soaking.

Protocol 3

A drop of compound stock was directly pipetted on the edge of the crystallization drop on the Swissci 2-drop plate and the compound solution slowly diffused in the crystallization drop. Crystals were harvested without addition of cryosolution.

2.9.7 Co-crystallization

Co-crystallization experiments were performed with fragments identified in NMR fragment screens of the Novartis core fragment library, available human allosteric binders and synthesised fragments as part of this thesis. Different co-crystallization protocols were adapted:

Protocol 1

Protocol 1 was performed with condition PEGs Suite G12 (0.2 M sodium tartrate, 20 % w/v PEG 3350). First, a pre-dilution to 12.5 mM compound in reservoir solution was prepared

in an Eppendorf tube. Crystal drops of 1 µL protein at 13.81 mg/mL in low salt buffer, 1 µL pre-diluted compound in reservoir and 0.5 µL microseeds were pipetted on VDXm plates. Before harvesting 2 – 3 µL of a cryoprotectant solution (5 mM compound, 33.25 % w/v PEG 3350, 28.5 % v/v glycerol) were added to the crystallization drop and crystals were harvested after 30 s – 2 min.

Protocol 2

Protocol 2 was performed with a variation of condition PEG/Ion HT H9 (0.12 M caesium chloride, 12 % w/v PEG 3350, 12 % v/v DMSO). First, a pre-dilution of 25 mM compound in reservoir without DMSO was prepared and the pre-dilution was mixed 1:1 (v/v) with protein at 12.24 mg/mL in an Eppendorf tube. The solution was vortexed and incubated on ice for approx. 1 ½ hours. Then the mixture was centrifuged at 1300 rpm in a micro centrifuge for 1 min and crystal drops of 400 nL protein/compound mixture and 100 nL microseeds were pipetted.

Protocol 3

Compound stock solutions were mixed with protein at 12.24 mg/mL to 5 mM or 10 mM final compound concentration. Mixtures were incubated on ice for 3 hours and the solution was centrifuged as described in Protocol 2. Crystallization screens (PEGs Suite, Cryos Suite, PEG/Ion HT, PEGs II Suite) were performed with crystal drops of 200 nL protein/compound mixture, 200 nL reservoir and 100 nL microseeds.

2.9.8 Data collection at SLS Villigen

All diffraction data were collected with a PILATUS 6M (25 Hz) pixel detector on X10SA (PXII) at the Swiss light source (Villigen, Switzerland) using monochromatic radiation at λ = 0.99995 Å to 1.00000 Å on crystals at 100 K. Data sets of 720 images in 180 ° rotation with 0.25° oscillation steps were recorded.

2.9.9 Data processing and structure determination

Data sets were indexed and integrated using XDS[497], release 20180226 and scaled using AIMLESS[498], release 0.7.2. Furthermore, diffraction data processing was carried out using Global Phasing Pipedream automatic pipeline[499] including autoPROC[500], version 1.1.7, molecular replacement using PHASER[501], version 2.8.2 and BUSTER refinement[502], version 2.11.7. For the first structure solved as part of this thesis, chain A of structure PDB ID 4RYP[300] was used as a search model and for additional structures a previously solved high resolution apo protein structure (in-house) was used as search model. Complexes with fragments were solved in the same way by molecular replacement with PHASER and the PanDDA approach was used for fragment-protein structure determination. For this purpose, autoPROC files were used as input files for PanDDA

maps generation. PanDDA event maps were analysed. Ligands were fitted in identified additional electron density and alternative conformations around the bound fragment were modelled. Additional refinement was performed by several rounds of alternating model building in COOT[503-504] and PHENIX[505] refinement. R_{free} values[506] were generated from 5 % diffraction data excluded from the refinement.

For human allosteric binders **116**, **113** and **121** (see Chapter 10.2.1) and synthesised merged compound **LM10** (see Chapter 10.2.7) ligand fitting was also done without PanDDA analysis and refinement was performed by several rounds of alternating model building in COOT[503-504] and BUSTER refinement.[502] Translation/libration/screw (TLS)[507] refinement was used with automatically determined TLS groups. For occupancy refinement of ligands automatic occupancy refinement as part of BUSTER was performed using GELLY script.

2.9.10 Data deposition and accession codes

Coordinate and structure factor files in mmCIF format for PDB data deposition were generated using pdb extract, version 3.24[508]. The *T. brucei* FPPS apo protein was deposited in the Protein Data Bank (PDB) under the accession code 6R36. Protein-ligand structures with human allosteric binders **116**, **113** and **121** (see Chapter 10.2.1) and synthesised merged fragment **LM10** (see Chapter 10.2.7) were deposited under the accession codes 6R37, 6R38, 6R39, and 6SII (Ligands in PDB: **3N2**, **JQE**, **BFH**, and **LEZ**). Additionally, eight protein-ligand structures using PanDDA analysis were deposited in the PDB as a group deposition. The deposition of the ground state model contained the structure factores of all data sets used for the PanDDA analysis, the ground-state PanDDA map and the coordinate file of the reference model. The ground state model was deposited under the accession code 5QTC. For each protein-ligand structure the bound-state coordinate file was deposited together with a structure factor file containing the final refinement mtz file, the PanDDA input map, and the PanDDA event map. These structures with fragments of the Novartis core fragment library and one fragment analogue (Compounds **132**, **131**, **138**, **129**, **123**, **130**, **136**, and **153**) were deposited under accession codes 5QT5, 5QT6, 5QT7, 5QT4, 5QTA, 5QTB, 5QT9, and 5QT8 (Ligands in PDB: **PBV**, **JNE**, **PBY**, **P9Y**, **PJD**, **PJG**, **PJ4**, and **PJ1**). All PDB submissions were put on hold for publication.

2.10 X-ray fragment screen in Oxford

2.10.1 Crystallization in Oxford

At the beamline I04-1 of Frank von Delft at Diamond, Oxford an XChem fragment screening workflow is implemented for high-throughput X-ray screening experiments. As part of a

collaboration with PhD student Elliot Nelson in Frank von Delft's lab at the SGC/Diamond, Oxford a fragment screen was performed at the facility in a one-week on-site stay.
Therefore, all required materials including aliquots of 12.24 mg/mL *T. brucei* FPPS protein samples, diluted microseeding stock (1:1000) and a Deepwell block with 2 mL of the crystallization condition (0.12 M caesium chloride, 12 % w/v PEG 3350, 12 % v/v DMSO), microseed buffer (1.6 M $NH_4H_2PO_4$, 0.08 M Tris HCl pH 8.5, 20 % v/v glycerol), 75 % glycerol solution, milliQ H_2O, Swissci 2-drop plates, microseals and aluminium cover seals were shipped to the Diamond, Oxford.

Crystallization experiments were set up on Swissci 2-drop (80 µL reservoir) and Swissci 3-drop (20 µL reservoir) sitting drop plates using the vapour-diffusion method. Crystallization drops of 200 nL protein sample, 200 nL crystallization reservoir and 100 nL diluted microseeds were pipetted with a Mosquito robot and crystal plates were stored and imaged in a Rock Imager 1000 at 20 °C.

2.10.2 X-ray fragment screen in Oxford

An X-ray fragment screen was performed on the XChem fragment-screening platform at Diamond, Oxford. First, images of crystal drops were loaded in TexRank[461] to rank drops according to the presence and quality of crystals and to specify the target location on the edge of the drop for compound transfer. Then, compounds were ejected as multiple 2.5 nL acoustic droplets on the crystallization plate using an Echo.[509] Crystals were soaked with fragments of the DSPL and Edelris library for 3 – 4 hours at a final concentration of 37.5 mM and 18.75 mM, respectively at 7.5 % v/v DMSO concentration. After soaking 332.5 nL cryoprotectant (50 % v/v glycerol) was additionally transferred with an ECHO to give a final concentration of 20 % v/v glycerol in the drop. Immediately after, crystals were manually harvested in Dual-thickness micro loops using the Crystal Shifter and flash-frozen in liquid N_2.

2.10.3 Data collection

X-ray diffraction data were collected with a Pilatus 6M-F detector (25 Hz, DECTRIS) in an unattended X-ray centring mode (17 crystals/min) using monochromatic radiation at $\lambda = 0.92$ Å on beamline I04-1 at the Diamond Light Source (Oxford, UK).

2.10.4 Data processing, structure determination and refinement

Collected data were automatically processed using the following pipelines: Xia2[510], DIALS[511], XDS[497], POINTLESS[512], DIMPLE, REFMAC5[513] and CCP4[514-515] and data processing

results were combined and displayed in the ISPyB[516] data management system. For molecular replacement, a previously solved apo protein structure (in-house) was used. Data were further processed by PanDDA as an integrated part of the XChemExplorer[462] graphical interface. Ligand restraints were generated in AceDRG[517], Grade, Phenix.elbow.[518] Initial models were improved by several rounds of alternating model building in COOT[503-504] and REFMAC5[513] refinement.

2.11 X-ray fragment screen in Grenoble

2.11.1 Crystallization in Grenoble

A proposal (Targeting *Trypanosoma brucei* FPPS by fragment-based drug discovery, PID: 2848) under the iNEXT framework (Marie Curie Horizon 2020) was accepted for an X-ray fragment screen in the HTX lab supervised by José Marquez at the EMBL in Grenoble. The HTX lab of Marquez had implemented an X-ray fragment screening facility including an automatic crystal harvester.[459 519-520]

The CrystalDirect harvester is based on the sitting drop vapour-diffusion method and CrystalDirect (CD2) plates with 50 µL crystallization reservoir volume and a 25 µm thick cyclic olefin copolymer film as a plate bottom layer for up to three crystallization drops were used. Therefore, the identified crystallization condition on Swissci 2-drop plates as part of this thesis described in Chapter 2.9.5 was further optimised for CD2 plates. Crystal drops consisting of 200 nL protein (11.5 mg/mL in Tris buffer), 200 nL crystallization reservoir (0.12 mM caesium chloride, 12 % w/v PEG 3350, 12 % v/v DMSO) and 100 nL microseeding stock were set up on CD2 plates and thin needles grew. Furthermore, variations of PEG 3350 (12 – 18 % w/v) were tested as well as halved crystallization drop volumes to decrease protein consumption. Optimised crystal morphology was obtained in condition 0.12 mM caesium chloride, 15 % w/v PEG 3350 and 12 % v/v DMSO with the initial crystallization drop size. In any case, microseeding concentration was most crucial for crystallization and thick needles for harvesting were observed with a microseeding stock solution of 1:1000 in seeding buffer.

Protein stocks in aliquots of 70 µL for the setup of one CD2 crystallization plate with up to three crystallization drops per well, 50 µL diluted seeds (1:1000), 50 µL microseed stock and the optimised crystallization condition in a 2 mL Deepwell block were provided to the HTX lab at the EMBL Grenoble. The Deepwell block was shipped at - 20 °C and therefore the impact of the freeze-thaw cycle on the crystal morphology and diffraction quality was additionally tested. Furthermore, protein-ligand structures of active-site binders compound **131** and compound **132** (for chemical structures, see Chapter 5.2.1) identified in the NMR fragment screen and human allosteric binder compound **116** (for chemical structure, see Chapter 4.2.4) were solved previously and were provided as positive controls for the crystal soaking system.

At the HTX lab of EMBL Grenoble, crystallization experiments were performed with the provided material using the sitting-drop vapour diffusion method on CD2 plates. Drops were set up of 200 nL protein sample, 200 nL crystallization reservoir and 100 nL diluted microseeds with a Mosquito robot. Crystal plates were stored and imaged in a Rock Imager 1000 at 20 °C. Initial plates showed single needles in up to 95 % of drops. When crystals were fully grown, a cryoprotectant solution (15 % PEG 3350, 45 % glycerol) to a final concentration of 20 % glycerol and compounds to a final concentration of 100 mM or 500 mM were added to the drop with a Mosquito robot. All crystallization experiments are designed and implemented in the web-based crystallization information management system (CRIMS) and therefore, experiments were remotely accessed and the crystallization hit rate was determined. Throughout the crystallization process, crystallization experiments were not reproducible anymore and further adjustment of microseeding concentration was needed. Higher dilutions (1:500, v/v) of the microseed stock were prepared. When crystallization experiments at Grenoble were not reproducible at all anymore, crystal plates were set up at Novartis and taken to Grenoble.

2.11.2 iNEXT fragment screen in Grenoble

Soaking experiments with fragments of the Enamine Golden Fragment library were performed. Therefore, full-grown crystals were soaked by addition of compound solution on top of crystallization drops with a Mosquito robot. Plates were incubated in a Rock Imager 1000 system at 20 °C up to 4 – 6 hours.

Crystals were harvested using the automatic crystal harvester equipped with a Satsuma femtosecond laser (Amplitude Systems) regulated by a scanner (Sunny Technology), a motorized and mobile 96 well microplate support, a glue-application station and a robotic arm to transfer crystal pins from the plate to a liquid nitrogen storage dewar.[520] In a first step, the crystallization plate was inserted with inverted orientation in the automatic crystal harvester. In the CrystalDirect robot control software the location of the aperture for aspirating the crystallization solution, the location of the target crystal and the shape and position of the excising area of the crystallization film around the crystal were defined. The crystallization liquid was aspirated in two intervals for 500 ms and a glue was added on the tip of a crystallization pin, which was further transferred on the outside of the film as specified in the software. The laser excised the film around the crystal and the crystal was transferred by the robotic arm to a cryocooling jet and finally to a nitrogen storage dewar with SPINE-pucks. [520] Crystals were stored in liquid nitrogen until data collection.

2.11.3 Data collection and processing in Grenoble

X-ray diffraction data were collected with a PILATUS3 2M or PILATUS3 6M (25 Hz, DECTRIS) pixel detector at the fully automated MASSIF-1 (ID30A-1)[458, 521] or at ID30B-1[522] beamline at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) on crystals at 100 K. First, four images were collected and automatically analysed by the program EDNA[523] to calculate the respective optimised strategy for a full data set collection. For each data set, 3600 diffraction images in 360° rotation with 0.1° oscillation steps were recorded and the following pipelines were run: EDNA[523], GrenADeS[524], autoPROC[500], XDSAPP[525], xia2_DIALS[510-511] and results of each pipeline were summarised in the beamline database ISPyB[516] and further displayed in CRIMS, version 4. Additionally, data were processed by the Global Phasing Pipedream automatic pipeline[499] including autoPROC[500], molecular replacement using PHASER[501], BUSTER refinement[502] and Rhofit (Global Phasing Ltd) for ligand placement. All data were also transferred to Novartis and additional data analysis was performed using PanDDA as described in Chapter 2.10.4.

2.11.4 Structure determination and refinement at Novartis

Output files from autoPROC processing were transferred to run PanDDA software at Novartis. PanDDA software enabled the identification of less populated binders and therefore, data processing was continued on PanDDA generated ensemble models. Successive rounds of manual rebuilding in COOT[503-504] and automatic refinement using PHENIX[505] were performed.

2.11.5 Data deposition and accession codes

Seven protein-ligand structures using PanDDA analysis were deposited in the PDB as a group deposition. The deposition of the ground state model contained the structure factors of all data sets used for the PanDDA analysis, the ground-state PanDDA map and the coordinate file of the reference model. The ground state model was deposited under the accession code 5QTD. For each protein-ligand structure the bound-state coordinate file was deposited together with a structure factor file containing the final refinement mtz file, the PanDDA input map, and the PanDDA event map. These structures with fragments **185**, **186**, **187**, **188**, **189**, **190**, and **191** identified in the fragment screen by X-ray at EMBL Grenoble were deposited under accession codes 5QTK, 5QTJ, 5QTI, 5QTH, 5QTE, 5QTF, and 5QTG (Ligands in PDB: **PJP**, **PJS**, **PJV**, **PJY**, **PKO**, **PK7**, and **PK4**). The PDB submission was put on hold for publication.

2.12 Fragment analogue search

Fragment analogues were searched from external providers on the eMolecules website. Fragment structures were provided as search input and the list of similar structures was visually inspected to choose compounds with high diversity and to explore efficiently the chemical space around the fragment. Similarly, fragment analogues were also searched within the Novartis compound collection. Therefore, a chemical substructure search with 90 % and 80 % overall compound similarity was performed with an in-house tool and compounds were filtered for availability as powder. After visual inspection as described for compounds from external providers, analogues were ordered in batches of up to ten compounds and compound stocks at 100 mM or 500 mM compound concentration were prepared. Then, fragment analogues were tested in protein-observed NMR, X-ray crystallography and SPR.

2.13 Fragment hit optimisation by medicinal chemistry

2.13.1 General procedure

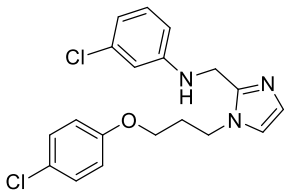
All chemicals and solvents were obtained from commercial sources and used without further purification. The reactions were monitored by UPLC-MS, acquired on a Waters MS system using a Waters Acquity UPLC with Waters SQ detector (ESI+/-). ^{1}H-NMR spectra of intermediates were recorded on a Bruker DPX 401 MHz spectrometer in d_6-DMSO. Crude products were either purified by flash chromatography using a Teledyne ISCO, CombiFlash Rf200 system with 12 g silica gel prepacked columns (RediSep® Rf) with a heptane:ethyl acetate gradient 1 – 100 % in 20 min or by AutoPurification mass-directed HPLC system by Waters eluting with an ammonium bicarbonate gradient (XBridge 30 × 150, C18, 50 mL/min, H_2O + NH_4HCO_3/ACN, 1 min 95/5, 11 min 20/80, 11.1 min 0/100, 14 min 0/100) or a trifluoroacetic acid (Sunfire 30 × 150, C18, 50 mL/min, H_2O + TFA/ACN, 1 min 95/5, 11 min 20/80, 11.1 min 0/100, 14 min 0/100). Preparative SFC was conducted on a Waters Preparative SFC-100-MS system with a Waters 2998 Photodiode Array Detector and a Waters MS Single Quadrupole Detection on a 250 × 30 mm Waters Viridis 2-EP 130A 5 µm column using a 7 – 15 % MeOH gradient in 10 min. ^{1}H-NMR and ^{13}C-NMR spectra of final products were recorded on a Bruker Avance 1 TXI 600 MHz. The purity of all compounds was determined by LC-MS and NMR and found to be >90 % unless otherwise stated. All chemical shifts were reported in ppm and multiplicity as follows: s, singlet; brs, broad singlet; d, doublet; dd, doublet of doublets; ddd, doublet of doublet of doublets; t, triplet; dt, doublet of triplet and m, multiplet.

Purity analysis and mass spectra were determined on Waters MS system using a Waters Acquity UPLC with Waters SQ detector (ESI+/-). High resolution mass spectrometry of final

products was performed by using electrospray ionization in positive ion modus after separation by liquid chromatography. The elemental composition was derived from the mass spectra acquired at high resolution of about 35'000 on a Q Exactive Plus mass spectrometer coupled to an Ultimate 3000 UHPLC. The high mass accuracy below 1.5 ppm was obtained by using a lock mass.

2.13.2 Experimental procedure for chemical synthesis

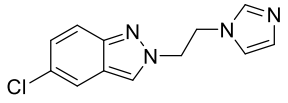
3-chloro-N-((1-(3-(4-chlorophenoxy)propyl)-1*H*-imidazol-2-yl)methyl)aniline (LM1)



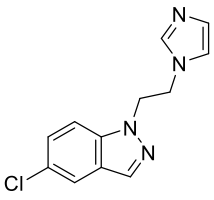
To a solution of N-((1*H*-imidazol-2-yl)methyl)-3-chloroaniline (0.241 mmol, 50 mg) in acetonitrile (2 mL) at 0°C under argon atmosphere was added potassium carbonate (1.5 eq, 0.361 mmol) and 1-(3-bromopropoxy)-4-chlorobenzene (1 eq, 0.241 mmol, 60.1 mg) and the mixture was stirred at room temperature for 1 hour. The mixture was further stirred at 70 °C for 3 days and put in the microwave for 1 hour at 100 °C two times. The mixture was transferred to a separation funnel and diluted with water (5 mL) and extracted with ethyl acetate (10 mL). The aqueous phase was back-extracted with ethyl acetate (2 × 10 mL) and dried over sodium sulfate. The solvent was evaporated under *vacuo* and dissolved in acetonitrile/water (1:1), filtered through a syringe filter and purified by prepLC-MS with a trifluoroacetic acid gradient. The final product was obtained as a beige powder.

Yield: 33 % (42.3 mg), Purity: 92 % (NMR), MS 376.10 $[M + H]^+$, calc. mass for $C_{19}H_{20}ON_3Cl_2$: 375.09. ^{1}H NMR (600 MHz, DMSO-d_6) δ 7.78 (d, J = 2.0 Hz, 1H), 7.61 (d, J = 2.0 Hz, 1H), 7.34 – 7.28 (m, 2H), 7.09 (t, J = 8.0 Hz, 1H), 6.93 – 6.87 (m, 2H), 6.77 (s, 1H), 6.71 – 6.63 (m, 2H), 6.57 – 6.49 (m, 1H), 4.67 (s, 2H), 4.36 (t, J = 7.1 Hz, 2H), 4.01 (t, J = 6.0 Hz, 2H), 2.27 (p, J = 6.5 Hz, 2H). ^{13}C NMR (151 MHz, DMSO) δ 156.97, 148.71, 144.84, 133.81, 130.60, 129.26, 124.49, 122.99, 119.01, 116.92, 116.13, 111.87, 111.19, 64.80, 44.34, 37.73, 28.64.

2-(2-(1*H*-imidazol-1-yl)ethyl)-5-chloro-2*H*-indazole (LM2) and 1-(2-(1*H*-imidazol-1-yl)ethyl)-5-chloro-1*H*-indazole (LM3)



LM2



LM3

5-chloro-1*H*-indazole (0.459 mmol, 70 mg) was dissolved in *N,N*-dimethylformamide (2 mL) under argon atmosphere and potassium carbonate (3 eq, 1.376 mmol, 190 mg) was added. The mixture was stirred for 3 hours and 1-(2-chloroethyl)-1*H*-imidazole (2 eq, 0.918 mmol, 120 mg) was added. The mixture was heated to 60 °C and stirred overnight. The reaction mixture was transferred to a separation funnel and diluted with dissolved sodium hydrogen carbonate (5 mL) and extracted with ethyl acetate (10 mL). The

aqueous phase was back-extracted with ethyl acetate (2 × 10 mL) and dried over sodium sulfate. The crude product was dissolved in acetonitrile: water (1:1), filtered through a syringe filter and purified by prepLC-MS with a trifluoroacetic acid gradient. A mixture of 1-(2-(1*H*-imidazol-1-yl)ethyl)-5-chloro-1*H*-indazole and 2-(2-(1*H*-imidazol-1-yl)ethyl)-5-chloro-2*H*-indazole was obtained that was separated by prepSCF purification. 1-(2-(1*H*-imidazol-1-yl)ethyl)-5-chloro-1*H*-indazole and 2-(2-(1*H*-imidazol-1-yl)ethyl)-5-chloro-2*H*-indazole were obtained as a white/yellowish solid.

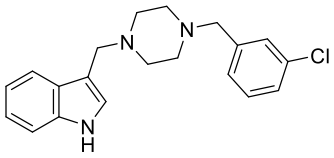
2-(2-(1H-imidazol-1-yl)ethyl)-5-chloro-2H-indazole (LM2):

Yield: 7.8 % (13.5 mg), Purity: 95 % (UPLC-MS), MS 247.07453 $[M + H]^+$, calc. mass for $C_{12}H_{12}N_4Cl$: 246.07. ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.14 (d, J = 1.0 Hz, 1H), 7.79 – 7.74 (m, 1H), 7.66 (dd, J = 9.1, 0.9 Hz, 1H), 7.42 (s, 1H), 7.22 (dd, J = 9.1, 2.1 Hz, 1H), 6.98 (s, 1H), 6.85 (s, 1H), 4.81 (dd, J = 6.7, 5.1 Hz, 2H), 4.59 (dd, J = 6.7, 5.1 Hz, 2H). ^{13}C NMR (151 MHz, DMSO) δ 146.58, 127.85, 126.56, 125.57, 124.81, 121.60, 119.56, 119.08, 53.32, 46.11.

1-(2-(1H-imidazol-1-yl)ethyl)-5-chloro-1H-indazole (LM3):

Yield: 15.8 % (28.7 mg), Purity: 91 % (UPLC-MS), MS 247.07445 $[M + H]^+$, calc. mass for $C_{12}H_{12}N_4Cl$: 246.07. ^{1}H NMR (600 MHz, DMSO-d_6) δ 8.09 (d, J = 1.0 Hz, 1H), 7.80 (d, J = 1.9 Hz, 1H), 7.47 (dt, J = 8.9, 0.9 Hz, 1H), 7.33 – 7.28 (m, 2H), 6.97 (d, J = 1.3 Hz, 1H), 6.74 (d, J = 1.3 Hz, 1H), 4.76 (dd, J = 6.5, 5.2 Hz, 2H), 4.46 (dd, J = 6.5, 5.2 Hz, 2H). ^{13}C NMR (151 MHz, DMSO) 138.35, 137.18, 133.08, 128.07, 126.33, 124.94, 124.07, 119.86, 119.35, 110.93, 48.90, 45.78.

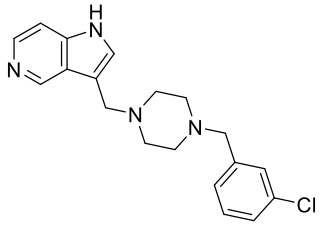
3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-indole (LM4)



tert-butyl 3-formyl-1*H*-indole-1-carboxylate (0.204 mmol, 50 mg) and 1-(3-chlorobenzyl)piperazine (1.1 eq, 0.224 mmol, 47.2 mg) were dissolved in dry DCM (2 mL) under argon atmosphere. A catalytic amount of acetic acid (3 drops) was added and the mixture was stirred for 3 hours. Sodium triacetoxyborohydride (1.5 eq, 0.306 mmol, 75 mg) was added and the mixture was stirred at room temperature overnight. The mixtures was transferred to a separation funnel, diluted with water (5 mL), and extracted with DCM (10 mL). The aqueous phase was back-extracted with DCM (2 × 10 mL) and the organic layer was dried over sodium sulfate. The solvent was evaporated *in vacuo*; the residue was dissolved in acetonitrile:water (1:1), filtered through a syringe filter and purified by prepLC-MS with an ammonium bicarbonate gradient. *Tert*-butyl 3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-indole-1-carboxylate was dried *in vacuo* and dissolved in DCM: trifluoroacetic acid (1:1). The mixture was stirred for 2 hours and the solvent was evaporated *in vacuo*. The final product was dissolved in acetonitrile:water (1:1), filtered through a syringe filter and purified by prepLC-MS with an ammonium bicarbonate gradient.

Yield: 46.3 % (32.2 mg), Purity: 100 % (UPLC-MS), MS 340.16 $[M + H]^+$, calc. mass for $C_{20}H_{23}N_3Cl$: 339.15. ^1H NMR (600 MHz, DMSO-d_6) δ 10.90 (d, J = 2.5 Hz, 1H), 7.61 (dd, J = 7.9, 1.1 Hz, 1H), 7.35 – 7.30 (m, 3H), 7.28 (dt, J = 8.2, 1.6 Hz, 1H), 7.23 (dt, J = 7.6, 1.4 Hz, 1H), 7.19 (d, J = 2.3 Hz, 1H), 7.05 (ddd, J = 8.1, 6.9, 1.2 Hz, 1H), 6.96 (ddd, J = 7.9, 7.0, 1.0 Hz, 1H), 3.61 (s, 2H), 3.44 (s, 2H), 2.36 (s, 8H). ^{13}C NMR (151 MHz, DMSO) δ 141.16, 136.29, 132.88, 130.01, 128.29, 127.59, 127.31, 126.80, 124.50, 120.87, 119.07, 118.35, 111.29, 110.79, 61.20, 53.17, 52.71, 52.53.

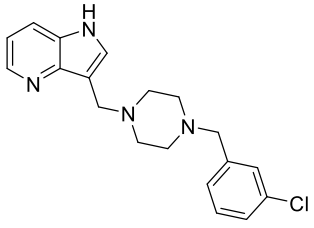
3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-pyrrolo[3,2-c]pyridine (LM5)



3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-pyrrolo[3,2-c]pyridine was prepared according to the procedure described for **LM4** without further treatment with DCM: trifluoroacetic acid (1:1) as the deprotected starting material was used. **LM5** was synthesised starting with 1*H*-pyrrolo[3,2-c]pyridine-3-carbaldehyde (0.342 mmol, 50 mg) and 1-(3-chlorobenzyl)piperazine (1.1 eq, 0.376 mmol, 79 mg) and the product was obtained as a white solid.

Yield: 22.89 % (37.5 mg), Purity: 95 % (UPLC-MS), MS 341.15 $[M + H]^+$, calc. mass for $C_{19}H_{22}N_4Cl_2$: 340.15. ^1H NMR (600 MHz, DMSO-d_6) δ 15.33 (s, 1H), 13.11 (s, 1H), 9.44 (s, 1H), 8.50 (d, J = 6.7 Hz, 1H), 8.07 (d, J = 2.3 Hz, 1H), 8.03 (d, J = 6.6 Hz, 1H), 7.49 (s, 1H), 7.44 – 7.41 (m, 2H), 7.35 (t, J = 4.2 Hz, 1H), 4.50 (s, 2H), 3.93 (s, 2H), 3.17 (s, 8H). ^{13}C NMR (151 MHz, DMSO) δ 142.43, 135.81, 133.24, 131.27, 130.45, 129.69, 128.69, 128.37, 123.96, 117.00, 115.06, 109.83, 58.96, 49.35, 49.15.

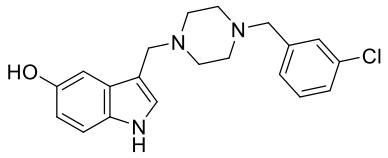
3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-pyrrolo[3,2-b]pyridine (LM6)



1*H*-pyrrolo[3,2-b]pyridine-3-carbaldehyde (0.345 mmol, 50 mg) and 1-(3-chlorobenzyl)piperazine (1.1 eq, 0.376 mmol, 79 mg) were dissolved in dry DCM (2 mL) under argon atmosphere. A catalytic amount of acetic acid (3 drops) was added and the mixture was stirred for 3 hours. Sodium triacetoxyborohydride (1.5 eq, 0.513 mmol, 109 mg) was added and the mixture was stirred at room temperature overnight. The mixture was transferred to a separation funnel, diluted with water (5 mL), and extracted with DCM (10 mL). The aqueous phase was back-extracted with DCM (2 × 10 mL) and the organic layer was dried over sodium sulfate. The solvent was evaporated *in vacuo*; the residue was dissolved in acetonitrile:water (1:1), filtered through a syringe filter and purified by prepLC-MS with a trifluoroacetic acid gradient. The product was obtained as a pinkish solid.

Yield: 45.3 % (73.4 mg), Purity: 96 % (UPLC-MS), MS 341.15 $[M + H]^+$, calc. mass for $C_{19}H_{22}N_4Cl_2$: 340.15. ^1H NMR (600 MHz, DMSO-d_6) δ 12.86 (s, 1H), 8.74 – 8.67 (m, 1H), 8.51 (d, J = 8.2 Hz, 1H), 8.24 (d, J = 2.9 Hz, 1H), 7.63 (dd, J = 8.2, 5.5 Hz, 1H), 7.49 (s, 1H), 7.43 (dd, J = 4.6, 2.0 Hz, 2H), 7.35 (q, J = 3.4, 2.6 Hz, 1H), 4.42 (s, 2H), 3.95 (s, 2H), 3.07 (d, J = 152.2 Hz, 8H). ^{13}C NMR (151 MHz, DMSO) δ 137.20, 136.55, 133.25, 131.80, 130.47, 129.74, 128.74, 128.41, 127.07, 117.19, 115.18, 58.88, 49.17, 48.96.

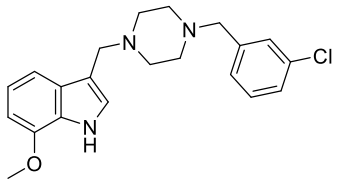
3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1H-indol-5-ol (LM7)



3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1H-indol-5-ol was prepared according to the procedure described for **LM4** without further treatment with DCM: trifluoroacetic acid (1:1) and starting with 5-hydroxy-1H-indole-3-carbaldehyde (0.310 mmol, 50 mg) and 1-(3-chlorobenzyl)piperazine (1.1 eq, 0.341 mmol, 71.9 mg). The final product was dissolved in acetonitrile/water (1:1), filtered through a syringe filter and purified by prepLC-MS with a ammonium bicarbonate gradient.

Yield: 6.20 % (7.6 mg), Purity: 88 % (NMR), MS 356.15 $[M + H]^+$, calc. mass for $C_{20}H_{23}ON_3Cl$: 355.15. ^1H NMR (600 MHz, DMSO-d_6) δ 10.57 (d, J = 2.5 Hz, 1H), 8.56 (s, 1H), 7.35 – 7.27 (m, 3H), 7.24 (dd, J = 7.5, 1.6 Hz, 1H), 7.11 (d, J = 8.6 Hz, 1H), 7.08 (d, J = 2.4 Hz, 1H), 6.92 (d, J = 2.3 Hz, 1H), 6.57 (dd, J = 8.6, 2.4 Hz, 1H), 3.51 (s, 2H), 3.44 (s, 2H), 2.48 – 2.16 (m, 8H). ^{13}C NMR (151 MHz, DMSO) δ 150.18, 141.17, 132.89, 130.86, 130.03, 128.32, 128.29, 127.31, 126.81, 125.01, 111.55, 111.24, 109.92, 103.06, 61.21, 53.43, 52.69, 52.61.

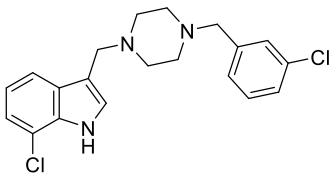
3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-7-methoxy-1*H*-indole (LM8)



3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-7-methoxy-1*H*-indole was prepared according to the procedure described for **LM4** without further treatment with DCM: trifluoroacetic acid (1:1) and starting with 7-methoxy-1*H*-indole-3-carbaldehyde (0.285 mmol, 50 mg) and 1-(3-chlorobenzyl)piperazine (1.1 eq, 0.314 mmol, 66.2 mg). The residue was purified by prepLC-MS with a trifluoroacetic acid gradient and the final product was obtained as a brown powder.

Yield: 46.1 % (51.8 mg), Purity: 94 % (UPLC-MS), MS 370.17 $[M + H]^+$, calc. mass for $C_{21}H_{25}ON_3Cl$: 369.16. ^1H NMR (600 MHz, DMSO-d_6) δ 10.99 (d, J = 2.5 Hz, 1H), 7.34 – 7.30 (m, 2H), 7.28 (ddd, J = 8.1, 2.2, 1.3 Hz, 1H), 7.25 – 7.18 (m, 2H), 7.08 (d, J = 2.4 Hz, 1H), 6.88 (t, J = 7.8 Hz, 1H), 6.62 (dd, J = 7.8, 0.8 Hz, 1H), 3.89 (s, 3H), 3.58 (s, 2H), 3.43 (s, 2H), 2.21 (s, 8H). ^{13}C NMR (151 MHz, DMSO) δ 146.03, 141.17, 132.87, 130.01, 129.13, 128.28, 127.30, 126.80, 126.31, 124.13, 118.91, 111.98, 111.29, 101.41, 61.19, 55.03, 53.18, 52.70, 52.48.

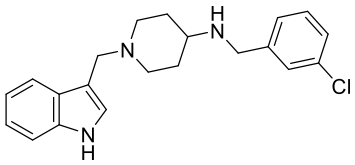
7-chloro-3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-indole (LM9)



7-chloro-3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1*H*-indole was prepared according to the procedure described for **LM4** without further treatment with DCM: trifluoroacetic acid (1:1) and starting with 7-chloro-1*H*-indole-3-carbaldehyde (0.278 mmol, 50 mg) and 1-(3-chlorobenzyl)piperazine (1.1 eq, 0.306 mmol, 64.5 mg). The residue was purified by flash chromatography and the final product was obtained as a white powder.

Yield: 27.1 % (37.6 mg), Purity: 98 % (UPLC-MS), MS 374.15 $[M + H]^+$, calc. mass for $C_{20}H_{22}N_3Cl_2$: 373.11. ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.30 (s, 1H), 7.61 (d, J = 7.9 Hz, 1H), 7.35 – 7.27 (m, 4H), 7.25 – 7.20 (m, 1H), 7.15 (d, J = 7.5 Hz, 1H), 6.99 (t, J = 7.7 Hz, 1H), 3.64 (s, 2H), 3.44 (s, 2H), 2.37 (s, 8H). ^{13}C NMR (151 MHz, DMSO) δ 141.08, 133.08, 132.89, 130.03, 129.56, 128.32, 127.34, 126.84, 120.51, 119.52, 118.26, 115.75, 61.12, 52.90, 52.40.

1-((1*H*-indol-3-yl)methyl)-N-(3-chlorobenzyl)piperidin-4-amine (LM9)



The first synthesis step is according to procedure described for **LM4** without further treatment with DCM: trifluoroacetic acid (1:1) and starting with *tert*-butyl 3-formyl-1*H*-indole-1-carboxylate (0.815 mmol, 200 mg) and tert-butyl piperidin-4-ylcarbamate (1.1 eq, 0.897 mmol, 180 mg). The residue was checked on TLC (heptane: ethyl acetate, 1:1), purified by flash chromatography (heptane: ethylacetate, 0 – 100 % in 20 min) and the solvent was evaporated *in vacuo*. *tert*-butyl 3-((4-((tert-butoxycarbonyl)amino)piperidin-1-yl)methyl)-1*H*-indole-1-carboxylate (0.656 mmol, 282 mg) was dissolved in *N, N*-dimethylformamide (5 mL) under argon atmosphere. The solution was cooled to 0 °C and after 10 min sodium hydride (3 eq, 1.969 mmol, 79 mg) was added under stirring. After 15 min, 1-chloro-3-(chloromethyl) benzene (1.1 eq, 0.722 mmol, 116 mg) was added and the mixture was stirred in an ice-bath overnight to allow a slow warm up to rt. The flask was again cooled to 0 °C and water was added (gas released). The mixture was extracted with ethyl acetate (15 mL) and brine, and the solvent was evaporated under *vacuo*. The product was checked on a TLC (heptane: ethyl acetate, 1:1) and purified by flash chromatography (heptane: ethyl acetate, 0 – 100 % in 25 min). *tert*-butyl 3-((4-((tert-butoxycarbonyl)(3-chlorobenzyl)amino)piperidin-1-yl)methyl)-*1H*-indole-1 carboxylate was dissolved in DCM: trifluoroacetic acid (1:1) and stirred overnight. The solvent was evaporated under *vacuo* and the product dissolved in DCM and purified by flash chromatography. The final product was obtained as a colourless powder.

Yield: 8.5 % (34.1 mg), Purity: 90 % (NMR), MS 354.17 $[M + H]^+$, calc. mass for $C_{21}H_{25}N_3Cl$: 353.17. ^{1}H NMR (600 MHz, DMSO-d_6) δ 11.57 (d, J = 2.7 Hz, 1H), 10.05 (s, 1H), 9.50 – 9.18 (m, 2H), 7.75 (d, J = 7.9 Hz, 1H), 7.62 (t, J = 1.4 Hz, 1H), 7.56 (d, J = 2.7 Hz, 1H),

7.51 – 7.41 (m, 4H), 7.20 – 7.07 (m, 2H), 4.46 (d, *J* = 2.9 Hz, 2H), 4.18 (d, *J* = 4.8 Hz, 2H), 3.57 (d, *J* = 12.7 Hz, 2H), 3.29 (s, 1H), 3.07 – 2.92 (m, 2H), 2.35 – 2.23 (m, 2H), 1.82 (qd, *J* = 13.5, 3.9 Hz, 2H).^{13}C NMR (151 MHz, DMSO) δ 136.02, 134.20, 133.20, 130.58, 129.82, 128.97, 128.72, 127.52, 121.76, 119.70, 118.50, 111.96, 102.11, 51.72, 50.37, 48.82, 46.77, 25.67.

2.14 SPR assay

An SPR assay was established for orthogonal SPR fragment screening of fragment hits identified by NMR and X-ray screening and for determination of K_d values of fragments in collaboration with Joy Petrick, NIBR, Basel, Switzerland.

Experiments were conducted with immobilized avi-tagged biotinylated proteins on a Series S Sensor Chip SA (GE Healthcare) on a Biacore T200 at 22 °C using the Biacore T200 Control Software. Each Sensor chip has four flow channels and therefore, avi-tagged biotinylated *T. brucei, T. cruzi* and human FPPS proteins were immobilised on one chip and the remaining channel was used as a reference channel without protein. Thus, in one experimental setup fragments were tested in parallel on the three different proteins. Samples were injected either from 4 mL SPR vials or from 7 mm SPR vials (GE Healthcare) covered with a Rubber cap, type 3 or type 5, respectively.

First, the chip was loaded and chip preconditioning was performed by three times injection of a solution containing 50 mM NaOH and 1M NaCl and two injections of milliQ H_2O on all four channels with a flow rate of 30 µL/min. The system was further primed with SPR buffer without DMSO and normalising was carried out with BIAnormalizing solution (GE Healthcare) containing 70 % glycerol. Afterwards, two to three injections of SPR buffer without DMSO were set up on all four flow channels with a flow rate of 30 µL/min to achieve a stable baseline. Then avi-tagged biotinylated *T. brucei*, *T. cruzi* and human FPPS proteins were filtered through a 0.45 µm Eppendorf sterile filter and 500 µL of protein pre-dilutions at 50 µg/mL in SPR buffer without DMSO were prepared. Proteins were immobilised on flow channels 2, 3 and 4 with a flow rate of 10 µg/mL. Flow channel 1 remained without protein as a control for nonspecific binding. Initially, a minimum of 1 min injection was set up and the response units (RUs) were read. Injections were repeated to achieve final chip densities between 2500 and 3500 RUs. Protein immobilisation was followed by several injections of SPR buffer without DMSO to check for baseline drifting.

Then, the device was primed twice with SPR buffer with DMSO and SPR buffer with DMSO was injected over all four flow channels at a flow rate of 50 µL/min with each injection for 60 s at least 20 times. An excluded volume correction (EVC) dilution series was prepared and included in the measurement to remove the DMSO effect on the observed responses. Thus, an eight point DMSO dilution series from 0.4 % v/v DMSO to 1.7 % v/v DMSO was prepared.

Compounds were either measured at the maximum soluble concentration but not higher than 500 µM as a single point measurement or an up to 12 point 2-fold dilution row was prepared. First,

compounds were pre-diluted in 90 % v/v d_6-DMSO and 10 % v/v D_2O solution to a concentration 100 times higher than the final highest concentration or single concentration to be measured in the SPR assay. A dilution plate with a 2-fold dilution series in 90 % d_6-DMSO and 10 % D_2O was prepared or the single compound dilution was transferred to a plate and an only d_6-DMSO/D_2O well was prepared for each compound. A final SPR source plate in a 96 well Greinert microplate with a 1:100 compound dilution in SPR buffer without DMSO was pipetted from the dilution plate and measured.

Additionally, two positive controls (compound **121**, compound **131**) with **121** at a compound concentration of 200 µM and **131** at a compound concentration of 50 µM were prepared. Furthermore, 2 mL of SPR buffer with DMSO referred to as "start-up" in the experiment software and a freshly pipetted mix of SPR buffer with 0.9 % v/v DMSO referred to as "buffer" were pipetted.

The experiment was designed in the Biacore T200 control software. The method was set up with an initial 20 times injection of the "start-up" solution followed by the injection of 24 fragment samples at a flow rate of 30 µL/min for 30 s followed by a dissociation of either 60 s or 180 s. The device run in automation for approx. 18 hours. Analysis, and curve fitting, K_d values were calculated in the Biacore T200 Evaluation Software. Base line drift, behaviour on reference channel was checked and curves were fitted in a 1:1 stoichiometry with an affinity analysis due to the fast k_{on} and k_{off} rates of fragments.

3. Setting the framework for *T. brucei* FPPS fragment-based drug discovery

3.1 Introduction

FBDD has received increasing attention since its first application in 1996.[333] It is now frequently applied both in academia and industry due to its ability to identify fragment binders even when HTS screens failed and due to its potential to identify ligands in new binding sites. As a first step, this approach requires sufficient high quality protein. In the following, protein production and characterisation results of *T. brucei* and human FPPS are presented. *T. brucei* FPPS and human FPPS were expressed unlabelled, $^{13}C^{15}N$-labelled, and biotinylated via avi-tag. The expression yielded up to 280 mg protein per litre *E. coli* cells and purification resulted in protein of >95 % purity. Purified unlabelled proteins for crystallization experiments and SPR experiments were also provided to Francesca Magari at the Philipps Universität Marburg, Germany, and Giulia Opassi at the Uppsala University, Sweden, respectively.

This thesis focuses on identifying *T. brucei* FPPS fragment binders, while human FPPS was mainly produced for comparison and counter screening. In addition, *T. brucei* FPPS was utilised for a large crystallization screen to identify a robust, reproducible, and well-diffracting crystal form that allows high-throughput structure determination, as is required for fragment-based drug discovery. *T. brucei* FPPS preferably crystallized in bundles of needles. In an extensive optimisation process of six screening rounds using microseed matrix screening and additives, a suitable crystallization condition was identified. X-ray diffraction data were collected at PXII (SLS Villigen, Switzerland) and an apo structure with a diffraction limit of 1.67 Å was determined and deposited in the PDB (PDB ID 6R36, on hold for publication).

3.2 Results

3.2.1 *T. brucei* and human FPPS protein production

Molecular cloning of expression plasmids encoding *T. brucei*, *T. cruzi* and human FPPS was performed as described in Chapter 2.6.1, and correct inserts were verified by DNA sequencing. Protocols for the expression of *T. brucei* FPPS and human FPPS are described in Chapters 2.6.3, 2.6.6 and 2.6.7. Briefly, protein expression in *E. coli* cells was performed in a temperature-, oxygen- and pH-controlled fermenter using auto-inducing medium.[526] Proteins were purified by IMAC followed by HRV 3C protease treatment to remove hexahistidine tags and subsequent size exclusion

chromatography, resulting in ~98 % pure and monodisperse protein as indicated by UPLC-MS (Table 24).

Table 24: Protein expression and purification summary.

Protein	Cell pellet weight[1] [g]	Protein amount[1] [mg]	Labelling rate [%]	Purity [%]	mass [Da]
Unlabelled *T. brucei* FPPS	128 – 142	78 – 280	-	97 – 99	42126.2
Unlabelled human FPPS	109 – 122	197 – 202	-	97 – 99	40687.5
$^{13}C^{15}N$ *T. brucei* FPPS	12 – 13	83 – 87	96.3 – 97.3	96 – 97	44397.3 – 4420.3
$^{13}C^{15}N$ human FPPS	17	62	96.9	100	42936.8
Avi-tagged biotinylated *T. brucei* FPPS	157	202	-	53	44162.8

[1] Cell pellet and protein amount obtained from 1 L *E. coli* cells

Overall yields were excellent ranging between 78 mg – 280 mg pure protein per litre of *E. coli* expression culture indicating that this protocol is highly suited to supply sample for fragment-based screens. Parts of the unlabelled protein purification were performed in collaboration with Francesca Magari within the framework of her three months secondment at the Novartis Institutes for Biomedical Research in Basel, Switzerland.

$^{13}C^{15}N$-labelled proteins were expressed using minimal labelling medium with the same protocol, and total protein pellet weight was reduced by approximately 10-fold for *T. brucei* FPPS and human FPPS, respectively. Final $^{13}C^{15}N$ *T. brucei* FPPS protein was obtained in yields up to approximately 3-fold reduced with a purity of 96 – 97 % and a stable isotope incorporation rate of 96.3 – 97.3 % as indicated by UPLC-MS. Final $^{13}C^{15}N$ human FPPS protein was obtained in yields reduced by approximately 2.5-fold with a purity of 100 % and a stable isotope incorporation rate of 96.9 % (Table 24). *T. brucei* FPPS was also expressed with an AviTag™ for *in vivo* enzymatic biotinylation utilising BirA as described in Chapter 2.6.7. Final avi-tagged *T. brucei* FPPS was 53 % biotinylated (Figure 22).

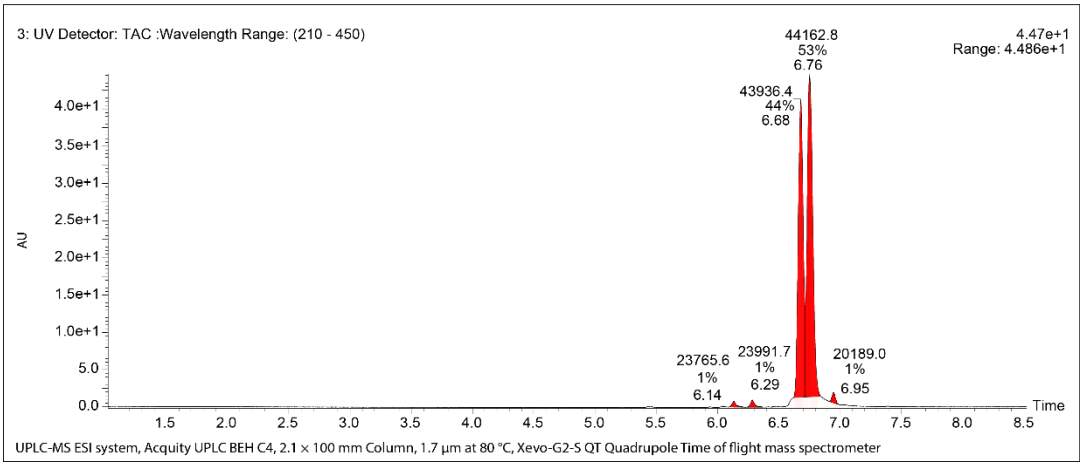


Figure 22: UPLC-MS chromatogram of avi-tagged biotinylated *T. brucei* FPPS. Two major products are observed, which correspond to the non-biotinylated (peak at 6.68 min, 44 %) and biotinylated (peak at 6.76 min, 53 %) avi-tagged *T. brucei* FPPS.

3.2.2 *T. brucei* and human FPPS protein characterization

Final unlabelled, $^{13}C^{15}N$-labelled and avi-tagged biotinylated *T. brucei* and human FPPS proteins were characterised by SDS-PAGE, UPLC-MS and HPLC and representative results of each method are shown in Figure 23.

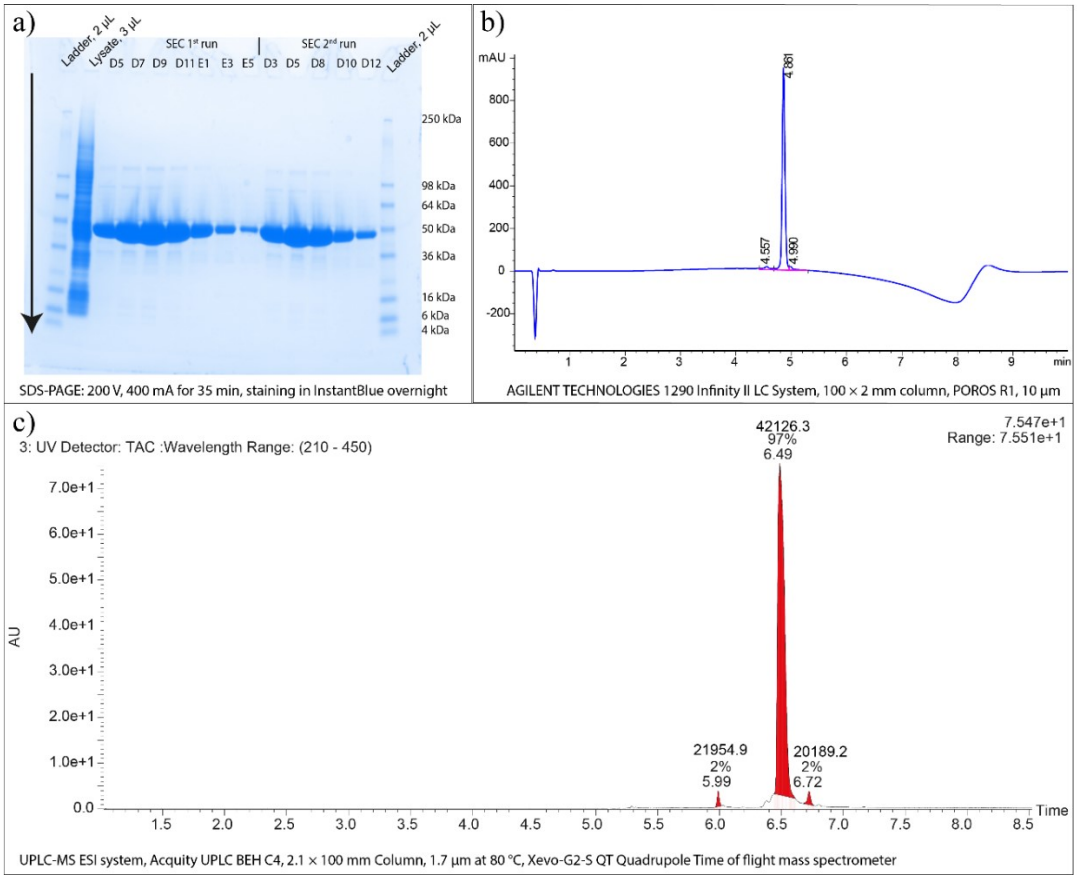


Figure 23: SDS-PAGE Gel, HPLC and UPLC-MS chromatogram of *T. brucei* FPPS. (a) SDS-PAGE gel of representative FPPS purification. **(b)** HPLC chromatogram of final protein. **(c)** UPLC-MS chromatogram of final protein.

3.2.3 Crystallization screening and optimisation for *T. brucei* FPPS crystal structures

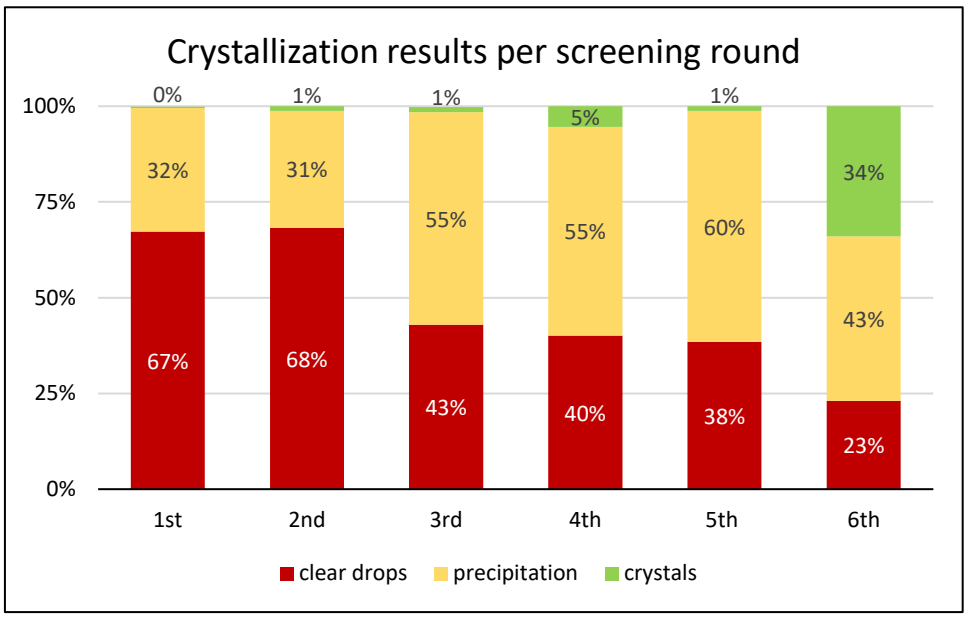
Reproduction of reported crystallization conditions[229] was unsuccessful and protein precipitation was observed in all crystallization drops. To identify a crystallization condition with reproducible crystal growth and defined crystal morphology for routine crystal harvesting, six iterative rounds of commercially available crystallization screens were performed with a screening of in total 1248 different crystallization conditions.

In the first screening round, 672 screened conditions yielded microcrystals in two crystal drops and 67.3 % of the drops remained clear. In the second screening round eight drops (1.2 % of total drops) with microcrystals and bundles of needles were observed (Figure 24). Four of these conditions (Cryos Suite C3, JCSG+ Suite C1, PEG/Ion HT E4 and E12, for crystal condition composition, see Chapter 2.9.3) with crystals in bundles of needles were chosen for optimisation on VDXm plates. Crystal growth was highly reproducible but crystals grew as clusters of needles and despite intensive optimisation no individual crystals were obtained.

To increase the search space, a third screening round was performed utilising *T. brucei* FPPS protein in a low salt buffer. However, this approach did not promote crystallization. For another screening process, the microseed matrix screening approach was performed and previously nucleated crystals in the microseeding stock were transferred from a high supersaturation condition to a new crystallization drop with a lower supersaturation level to favour a slow and ordered crystal growth process.[430-431] This method markedly improved crystallization of *T. brucei* FPPS and 5.4 % of crystal drops in six crystal screening plates with 576 conditions yielded crystals. Individual crystals were identified in condition PEG G12 (0.2 M sodium tartrate, 20 % w/v PEG 3350) and crystals were readily harvested and data were collected resulting in a 1.9 Å resolution structure. However, these crystals generally grew in thin needles up to 630 µm length but only 9 µm diameter and a different crystal morphology was desired for easier harvesting and X-ray data collection. In the fourth screening round using matrix microseeding approach still 40.1 % of the crystal drops remained clear, which indicated that the protein concentration was insufficient and the crystallization rate could be further increased with a higher protein concentration (Figure 24).

In the fifth screening round the protein concentration was increased to 13.96 mg/mL whereas in the sixth screening round also the microseed matrix screening approach was employed. Crystals appeared in remarkable 34 % of the crystal drops (Figure 24). Generally, a high crystallization rate was observed in conditions with PEGs Suite and *T. brucei* FPPS preferably crystallized in the presence of PEG 3350. Single crystals were obtained and optimised from condition PEG Suite G12, PEG/Ion HT A2, PEG/Ion HT H9, JCSG+ D2 with components listed in

Table 22 (Chapter 2.9.5,p. 78) and best X-ray data set was collected up to 2.6 Å and 2.7 Å of single needles harvested after 4 days.



1st screening round:
Low protein conc. in high salt buffer

2nd screening round:
Low protein conc. in high salt buffer

3rd screening round:
Low protein conc. in low salt buffer

4th screening round:
Low protein conc. in low salt buffer + microseeding

5th screening round:
High protein conc.in low salt buffer

6th screening round:
High protein conc. in low salt buffer + microseeding

Figure 24: Crystallization results per screening round. The chart shows the percentage of clear drops (red), drops with precipitation (yellow) and drops with crystals (green) per screening round inspected after 10 days of crystal setup. The result is visualised in a bar graph. Below the chart is a short description of the crystallization protocol used per screening round.

3.2.4 Additives in crystallization conditions – its decisive role

The improvement of nucleation, crystal growth and the quality of X-ray diffraction with the addition of additives is well-documented in the literature.[432-433] A Hampton additive screen on top of identified crystallization conditions improved crystal morphology and X-ray diffraction quality and for the first time in this work, data sets with a diffraction limit up to 2 Å were collected (Table 25).

Table 25: Crystallization conditions with small molecule additives.

Initial condition	Components of condition with additive	X-ray diffraction
PEGs Suite G12	0.20 M sodium tartrate, 20 % w/v PEG 3350, 6 % v/v dioxane	2.1 Å
PEG/Ion HT H9 var.	0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % v/v DMSO	2.0 Å
JSCG+ Suite D2 var.	0.12 $MgCl_2$, 0.08 M HEPES pH 7.5, 20 % w/v PEG 400, 0.02 M Pro	2.0 Å

In addition to conditions in Table 25, X-ray data sets of crystal soaks up to 2.5 Å diffraction limit were collected from crystals in PEG/ Ion HT H9 var. condition (0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % ethylene glycol and 0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % ethylene glycol). However, crystal growth was irreproducible under these conditions. Iterative rounds of microseeding ultimately resulted in highly reproducible crystals in the following condition: 0.12 M caesium chloride, 12 % v/v PEG 3350 and 12 % v/v DMSO. Final crystal drop composition is shown in Table 26. Unexpectedly, the change to a new protein batch required the use of high salt protein buffer for reproducible crystallization. Crystals appeared within 12 hours, grew to their full size in 30 hours and exhibited overall dimensions of 640 µm length and 28 µm diameter. As these crystals diffracted to 1.67 Å, the established crystallization system exhibited excellent properties for X-ray fragment screens and all subsequent crystals were grown and harvested under the same condition.

Table 26: Composition of final crystallization drop.

Volume [nL]	Component	Composition in crystal drop
200	Protein buffer	4.90 mg/mL *T. brucei* FPPS protein, 20 mM Tris-HCl, 80 mM NaCl, 0.8 mM TCEP-HCl
200	Reservoir	0.048 M caesium chloride, 4.8 % w/v PEG 3350, 4.8 % v/v DMSO
100	Microseed buffer	0.32 M $NH_4H_2PO_4$, 0.016 M Tris HCl, 4 % v/v glycerol

Over the course of this thesis, crystallization in single needles was the best crystal morphology observed. Other than that, *T. brucei* FPPS protein crystallized in sea-urchin like shape, microcrystals and bundles of needles (Figure 25). Crystal needles showed up to 15 % DMSO tolerance and allowed crystal soaking experiments up to 100 mM compound concentration. Depending on the final DMSO concentration and applied compound, compounds were soaked for up to three days. However, crystals aged over time. A small study of harvesting crystals after four days and after eleven days showed that the diffraction limit decreased from 1.8 Å to 3.3 Å.

Therefore, always fresh crystal plates were setup and harvested within 4 days to solve *T. brucei* FPPS crystal structures up to a diffraction limit of 1.67 Å.

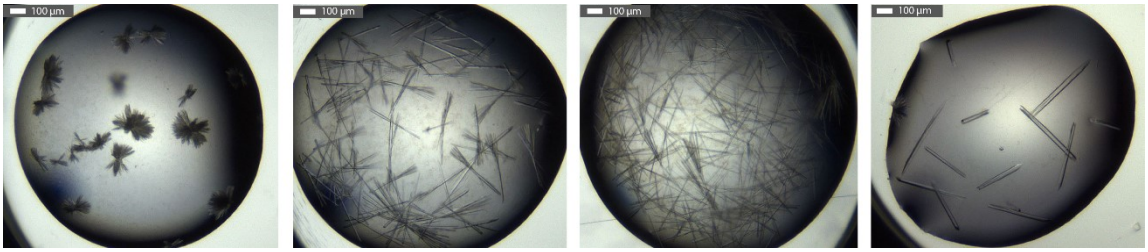


Figure 25: Various crystal morphology in different crystallization conditions. From left to right: 1) PEG/Ion HT E8: 0.2 M sodium malonate, 20 % w/v PEG 3350. 2) Index HT H2: 0.2 M potassium/sodium tartrate, 20 % w/v PEG 3350. 3) PEG/Ion HT A1: 20 % w/v PEG 3550. 4) PEG/Ion HT H9 var. + additive: 0.12 M caesium chloride, 12 % PEG 3350, 12 % DMSO.

3.2.5 Diffraction data collection and structure determination of *T. brucei* FPPS

Over the course of this thesis, 149 apo protein crystals were harvested and X-ray data sets were collected at the Swiss light source (Villigen, Switzerland) with a PILATUS 6M (25 Hz) pixel detector on X10SA (PXII). The best *T. brucei* FPPS apo protein data set had a diffraction limit of 1.67 Å and was stored in the PDB (Figure 26).

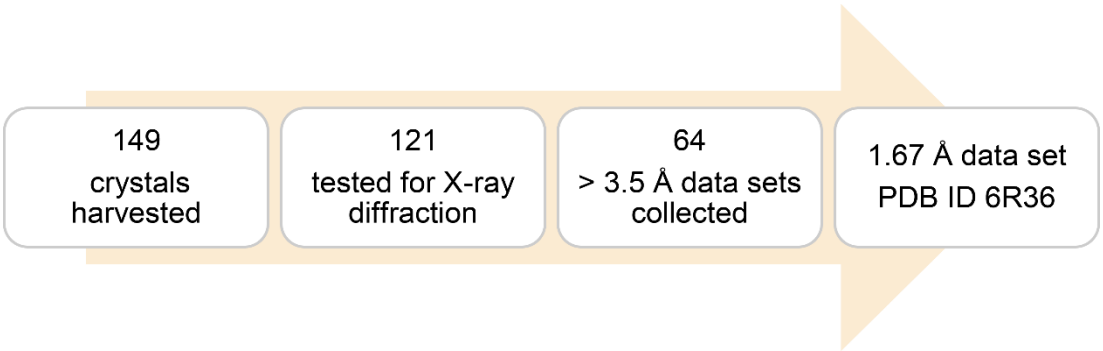


Figure 26: Overview of *T. brucei* FPPS apo protein harvesting to PDB submission. From left to right: the number of crystals harvested to the diffraction limit of the final data set submitted to the PDB.

X-ray diffraction data showed anisotropy and therefore, data processing was performed with STARANISO of the XDS software package. Crystals belong to the $P6_122$ space group and the apo crystal data set stored in the PDB had unit cell parameters of a = 60.5, b = 60.5 and c = 341. It is the first structure of *T. brucei* FPPS in this space group, and structures in the PDB belong to the C121 and $P3_121$ space group. The structure was solved by molecular replacement with PHASER. A Matthews coefficient of 2.14 $Å^3\,Da^{-1}$ was calculated and suggested one monomer of the homodimer in the asymmetric unit corresponding to a 42.6 % solvent content. Data collection and refinement statistics are found in the Appendix, Chapter 10.3. The *T. brucei* FPPS protein sequence contains 367 residues and the electron density map had some undefined areas of three different solvent exposed loop regions (residues 65 – 74, 108 – 121 and 184 – 194) and the

C-terminus (residues 364 – 367). Therefore, these residues were not modelled. The final model has relatively high R-values (R_{free}/R_{work} 0.2192/0.2429) due to anisotropy and weak electron densities close to the loop regions and the C-terminus as these regions show higher flexibility and are strongly disordered. Nevertheless, the structure shows a higher resolution compared to all *T. brucei* FPPS structures available in the PDB and therefore, especially with the respect to fragment soaking it enabled conclusive experiments which will be shown as part of this thesis.

3.3 Discussion

The target protein *T. brucei* FPPS was expressed and purified from *E. coli* cells. *E. coli* is a well-described and long used expression system to obtain high amounts of proteins in a quick, efficient and cheap way.[527] High protein yields of *T. brucei* FPPS were obtained that can be attributed to following aspects: the simplicity of the *T. brucei* FPPS protein sequence and the 3-dimensional protein structure, the *E. coli* BL21 (DE3)-pET expression system used for high efficiency protein expression of the target protein under the control of a T7 promoter, the TB auto inducing media that contains high levels of nutrient for high cell densities and the controlled cell growth environment in the fermenter. High amounts of protein with high purity were needed for in-house crystallization experiments and crystallization experiments by collaborators as well as for a FBDD campaign to identify *T. brucei* FPPS binders.

The first crystal structure of *T. brucei* FPPS was solved at 3.3 Å in 2004.[229] To date, 16 *T. brucei* FPPS apo protein and *T. brucei* FPPS-bisphosphonate complex structures are stored in the PDB with diffraction limits up to 1.94 Å. Reported[229] *T. brucei* FPPS crystallization conditions were not reproducible at Novartis. While several publications reported a *T. brucei* FPPS construct of 390 amino acids with an N-terminal hexahistidine-tag for crystallization, the protein construct used in this work contained a HRV-3C protease cleavage site (SLEVLFQGP) between the hexahistidine-tag and the sequence encoding for *T. brucei* FPPS protein. The hexahistidine-tag was cleaved with MBP-His_6-HRV-3C protease during the purification process and only glycine and proline of the HRV-3C recognition site remained. The use of different constructs could be an explanation for the lack of reproducibility of reported crystallization conditions. Crystallization and reproducibility of *T. brucei* FPPS crystals was challenging and an exhaustive number of 1248 crystallization conditions was screened in six screening rounds to develop a crystallization protocol that enabled to solve crystal structures down to 1.67 Å diffraction limit routinely. Apart from the classical approach of varying the buffer strength, salt and precipitant concentration, also the microseed matrix screening approach and the addition of small molecules have made a significant difference for *T. brucei* FPPS crystallization. This is consistent with experiences described in the literature reporting on the success of these two specific approaches and especially also the combination of different optimisation strategies for challenging targets.[528-529] For the

presented work, crystal structures of protein-fragment complexes with diffraction limits above 2.5 Å are especially important to enable unambiguous identification of the fragment binding mode.

T. brucei FPPS protein crystallized as needles and X-ray diffraction analysis showed slight anisotropy, which describes a direction dependent X-ray diffraction limit requiring a different cut-off of merged intensity data on each axis. For isotropic data, the isotropic-diffraction cut-off surface is spherical. However, for anisotropic data applying the same diffraction cut-off would result in loss of high-resolution data in one lattice direction. Therefore, X-ray data was processed using STARANISO, which defines a varying direction dependent anisotropic diffraction cut-off surface only including statistically significant data based on the locally averaged signal to noise (I/σ) value.[497, 500, 530] Crystals belonged to space group P 6_122 with average unit cell parameters: a = 60.5 b = 60.5 c = 341. The particularly long c axis can be troublesome as the corresponding diffraction pattern might result in overlapping reflection spots in the area detector that are challenging to deconvolute. However, the astonishing developments at synchrotrons and the implementation of improved data processing software allowed a straightforward data collection and processing of *T. brucei* FPPS crystal data sets. One important aspect was the oscillation angle per image and the application of fine slicing to reduce spot overlapping.[531] Fine slicing describes the collection of an image at a very small rotation angle (in this case at 0.25 °), which results in intensities of one reflection distributed on adjacent images and reduced overlap between reflections.[455] The XDS software performs three-dimensional profile fitting of these reflections from consecutive images to reveal interpretation over individual reflections. They are then compared with reference profiles.[532] This method led to an improved signal-to-noise ratio and lower background. Additionally, the crystal was mounted with the long cell axis along the rotation axis to further simplify data collection and all together, thanks to these recent advances and developments, a straightforward crystal structure analysis was possible.

4. An allosteric site on *T. brucei* FPPS

4.1 Introduction

The current literature status for new *T. brucei* FPPS inhibitors mainly focuses on bisphosphonate analogues binding to the active site.[116, 299, 303, 533] Two studies on human FPPS and *Pseudomonas aeruginosa* FPPS report on a previously unknown allosteric site. In a fragment screen on human FPPS, fragments bound to an allosteric site were detected and allosteric fragment binders were optimised to low-digit nanomolar inhibitors.[310] In a DSF screen and X-ray follow-up experiments two compounds bound to the same allosteric location on *Pseudomonas aeruginosa* were found.[228]

In the present work, the *T. brucei* FPPS X-ray structure was investigated and a sequence alignment was performed with human FPPS to analyse if the identified human FPPS allosteric site also exists in *T. brucei* FPPS. To complement the investigation of the *T. brucei* FPPS protein, molecular dynamic simulations were performed to search for hot spots of binding in *T. brucei* FPPS as part of a research stay at the University of Barcelona, where Moira Rachman and Maciej Majewski have contributed to this work. The previously identified allosteric site in human FPPS was also identified in *T. brucei* FPPS with a 70 % amino acid identity. This finding was the basis for testing ten reported human allosteric binders and analogues thereof on *T. brucei* FPPS. In this chapter, the identification of the allosteric site on *T. brucei* FPPS is highlighted and three *T. brucei* FPPS-ligand complex structures are presented.

4.2 Results

4.2.1 Sequence and structural similarities of human and *T. brucei* FPPS

Up to now, only X-ray complex structures with bisphosphonates in the allylic site of *T. brucei* FPPS were solved and no other binding site was described. The *T. brucei* FPPS structure was investigated in detail and a similar allosteric site on *T. brucei* FPPS was identified in similar spatial location as the previously described for human FPPS (Figure 27).

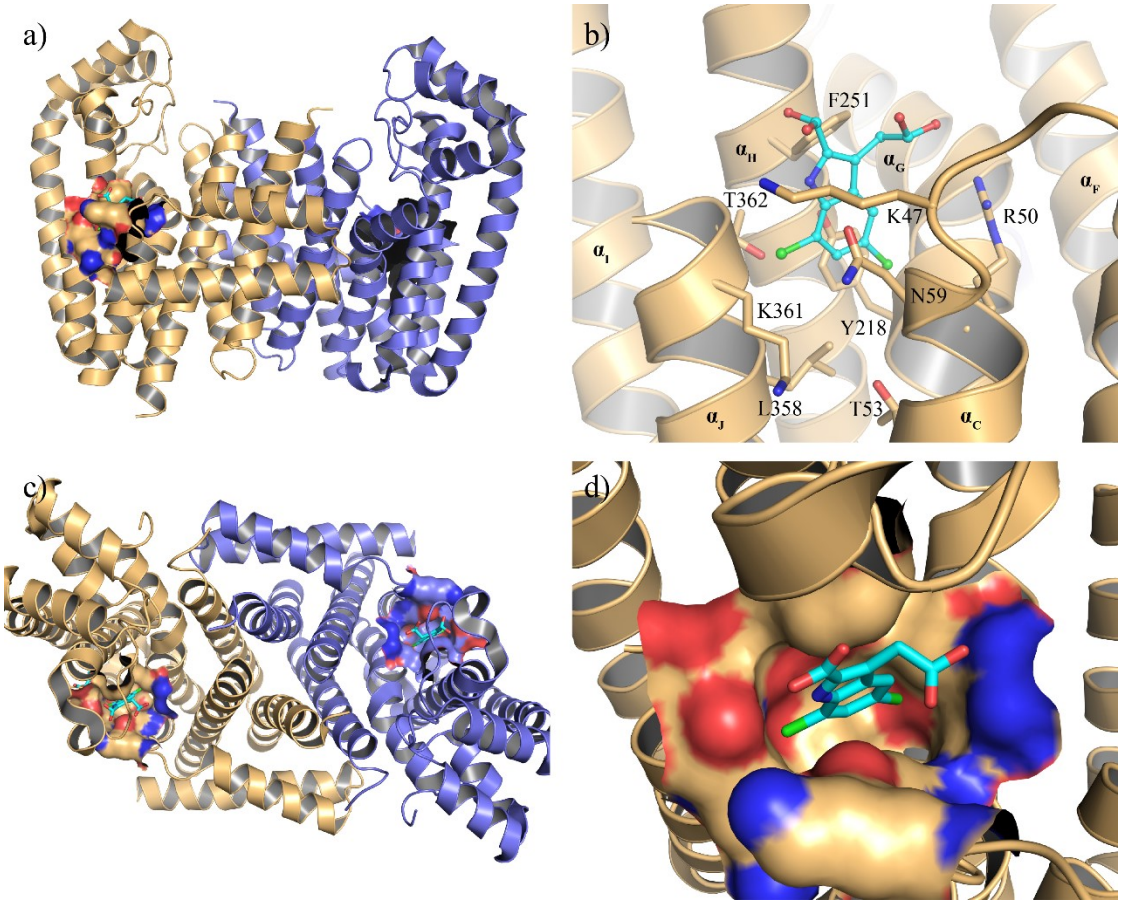


Figure 27: Allosteric site in *T. brucei* FPPS. *T. brucei* FPPS homodimer (PDB ID 6R36) with human allosteric binder (PDB ID 3N5J) as a reference for binding site location. **(a)** Overall view of the *T. brucei* FPPS homodimer in Cartoon representation with additional surface representation of residues: Asn49, Tyr218, Arg50, Leu358, Thr362, Phe251, Lys361, Thr53, Thr217 in the allosteric site. **(b)** Close-up on the allosteric site with same residues as in **(a)** stick representation. **(c)** Top view on *T. brucei* FPPS as in **(a)**. **(d)** Close-up on the allosteric site.

The allosteric site was characterised as being aligned by helices α_C, α_G, α_H. Additionally, the C-terminal helix α_J with a hydrophobic and charged, polar side and ten residues were found to form the overall site.[310] The X-ray structure of *T. brucei* FPPS was analysed and the respective overlaying residues in the *T. brucei* FPPS allosteric site were identified (Figure 27 and Table 27). In contrast to human FPPS, helix α_A in *T. brucei* FPPS is located remotely from the allosteric site and therefore, no residue on helix α_A was observed to participate in forming the allosteric site (Figure 28).

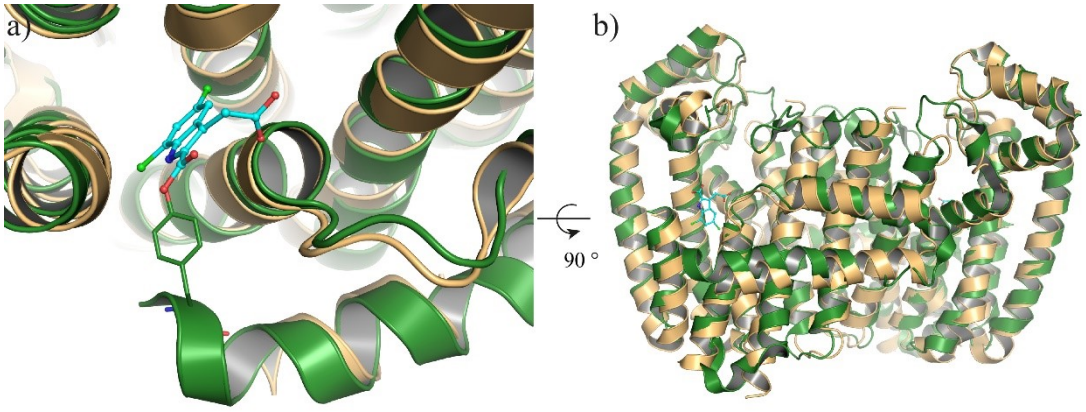


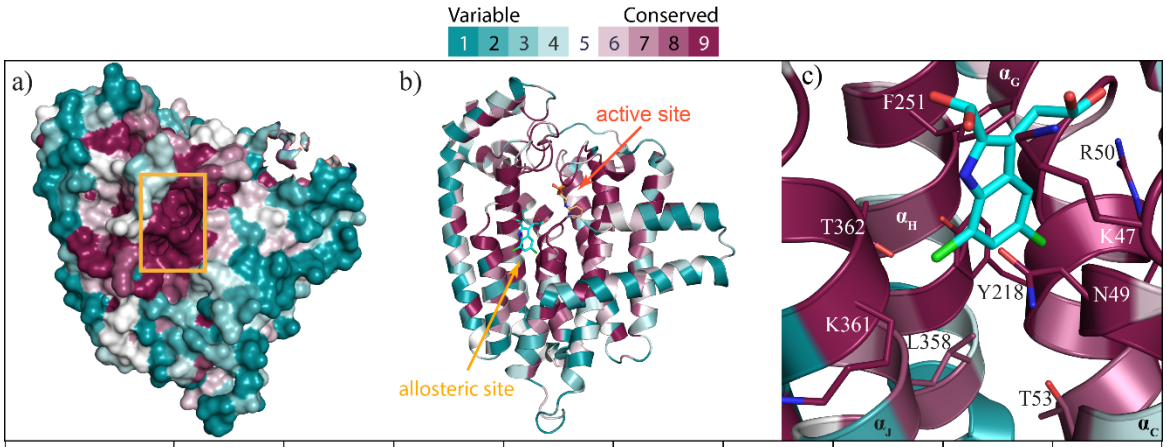
Figure 28: Residue Tyr10 on human FPPS forms the allosteric binding site. Overlay of human FPPS (PDB ID 4XQR, green) and *T. brucei* FPPS (PDB ID 6R36, yellow) with allosteric binder (PDB ID 3N5J, cyan) as a reference. **(a)** Top view on allosteric site. Stick representation of Tyr10 to show participation in forming the allosteric binding site. **(b)** Overview of human and *T. brucei* FPPS homodimer.

Comparing the ten residues that align the allosteric site on human FPPS with the residues overlaying on *T. brucei* FPPS, seven residues show an amino acid identity and two residues show an amino acid similarity (Phe206 vs. Tyr218 and Ile348 vs. Thr362). No residue is observed in *T. brucei* FPPS which matches with Tyr10 in human FPPS. This results in a 70 % amino acid identity and 90 % similarity to the allosteric site found in human FPPS (Table 27).

Table 27: Comparison of residues in allosteric site on *T. brucei* and human FPPS.

Human FPPS	Y10	K57	N59	R60	T63	F206	F239	L344	I348	K347
***T. brucei* FPPS**	-	K47	N49	R50	T53	Y218	F251	L358	T362	K361
Location	α_A	α_B-α_C	α_C	α_C	α_C	α_G	α_H	α_J	α_J	α_J

The conservation of the residues in the allosteric site across the FPPS proteins of different organisms was further analysed using the ConSurf Server[534] to investigate the degree of conservation of an amino acid in the FPPS protein throughout the different organisms. For a detailed description of the ConSurf parameters and analysed UniProt entries, see Appendix, Chapter 10.6. The *T. brucei* FPPS structure (PDB ID 2I19) was used as the input structure. It also contains the residues of the flexible loop regions whereas in the *T. brucei* FPPS structure solved as part of this thesis these residues could be assigned to the difference density. The allosteric site is well-conserved in the analysed FPPS structures and all residues forming the site were grouped in conservation groups 8 to 9 where 0 represents low conservation and 9 highest conservation. Residues Lys47, Asn49, and Lys361 are conserved across all analysed FPPS structures (Figure 29).



TbruFPPS	K47	N49	R50	T53	Y218	F251	L358	K361	T362
Colour group									
Conservation	9	9	9	8	9	9	8	9	9
Variations	none	F, N, H, Y, I, Q, T	none	T, C, S, L, A, M, I, G	F, Y, I, V, L	L, C, F	V, M, L, F, Y	none	T, V, I

Figure 29: *T. brucei* FPPS ConSurf model illustrating sequence variability. ConSurf model on *T. brucei* FPPS (PDB ID 2I19) generated by using 497 sequences manually chosen (for details, see Appendix, Chapter 10.6). **(a)** *T. brucei* FPPS homodimer as surface representation with active and allosteric site marked with an orange square. **(b)** *T. brucei* FPPS homodimer in Cartoon representation with bisphosphonate (PDB ID 2I19, yellow) in the active site and human allosteric binder (PDB ID 3N5J, cyan) overlaid in the allosteric site. The two sites are labelled. **(c)** Close-up on the allosteric site with the human allosteric binder (PDB ID 3N5J, cyan) superimposed. Residues and helices are labelled.

The conservation groups are dependent on the number and identity of chosen sequences for analysis. In this study, 497 sequences were chosen to cover a wide range of homologs with amino acid identities from 35 % to 100 % including a diverse set of organisms (see Appendix, Chapter 10.6, Figure 82). These results have to be analysed with caution as one residue can be highly conserved but still the target structure, in this case the *T. brucei* FPPS, could be one of the few structures with an amino acid variation. However, the residue in the *T. brucei* FPPS structure was consistent with the conserved residues in this position. In the few structures that show variations to the residues in *T. brucei* FPPS the variations extend over residues with very different properties. For instance, the polar Asn49 is replaced by the bulky hydrophobic residue phenylalanine, which significantly changes the properties of the site (Figure 29). Notably, this variation is also found in the closest parasitic analogue *T. cruzi* FPPS (UniProt: Q8WS26). Apart from the allosteric site, residues in the active site are also highly conserved but helices α_A, α_B and α_I on the surface of the protein show high variability.

4.2.2 Hot spot identification on *T. brucei* FPPS

The presence of high affinity interaction spots in the active and allosteric site was investigated during a two month research stay in the lab of Xavier Barril at the University of Barcelona. MDMix

molecular dynamic simulations with an organic/aqueous solvent mixture were performed on the first *T. brucei* FPPS apo structure solved as part of this thesis and on a PDB structure (PDB ID 2I19) of *T. brucei* FPPS in complex with a bisphosphonate. In a molecular dynamic simulation with three replicas using an ethanol/water solvent mixture, hydrophobic and polar binding hot spots were identified on the surface, the active and allosteric site of the *T. brucei* FPPS. (Figure 30 a). Binding hot spots for polar interactions were observed from the ethanol oxygen residing in the respective area for a long time during the simulation and similarly, binding hot spots for hydrophobic interactions were represented by the ethanol tail carbon. The binding free energy of water molecules was calculated and hydration sites were visualised by superimposing them on the *T. brucei* FPPS structure (Figure 30). In the *T. brucei* FPPS apo protein and bisphosphonate complex structure a binding hot spot for hydrophobic interactions was identified close to the chlorine atom of the overlaid human allosteric binder (Figure 30 c and f, marked with a black circle) and in the *T. brucei* FPPS apo structure another hot spot was identified close to the aromatic backbone of the indole scaffold (Figure 30 c, marked with a black circle). Furthermore, an additional hot spot for polar interactions is found close to the Asn49 side chain in the closed conformation of the PDB structure with bisphosphonate (Figure 30 g, marked with a black circle). The hot spot was not identified in the apo structure as N49 flips towards helix α_B in all three replicas (Figure 30 d).

The analysis further calculates the binding free energy of waters and predicts their displaceability. The *T. brucei* FPPS PDB structure contains two displaceable waters that overlay with the human allosteric binder in the allosteric site (Figure 30 g, marked with arrows). Additionally, polar hot spots close to the loop α_B-α_C are lacking as this region showed high conformational changes in the molecular simulations (Figure 30 d and g). However, MDMix simulations identified a water with low binding free energy close to the chlorine substitutent of the overlaid human allosteric binder that is predicted not to be displaceable in both structures (Figure 30 d and f). The *T. brucei* FPPS apo structure contains water molecule W1 close to this hot spot and this water would be hard to be displaced by a ligand as predicted by the simulation (Figure 30 d).

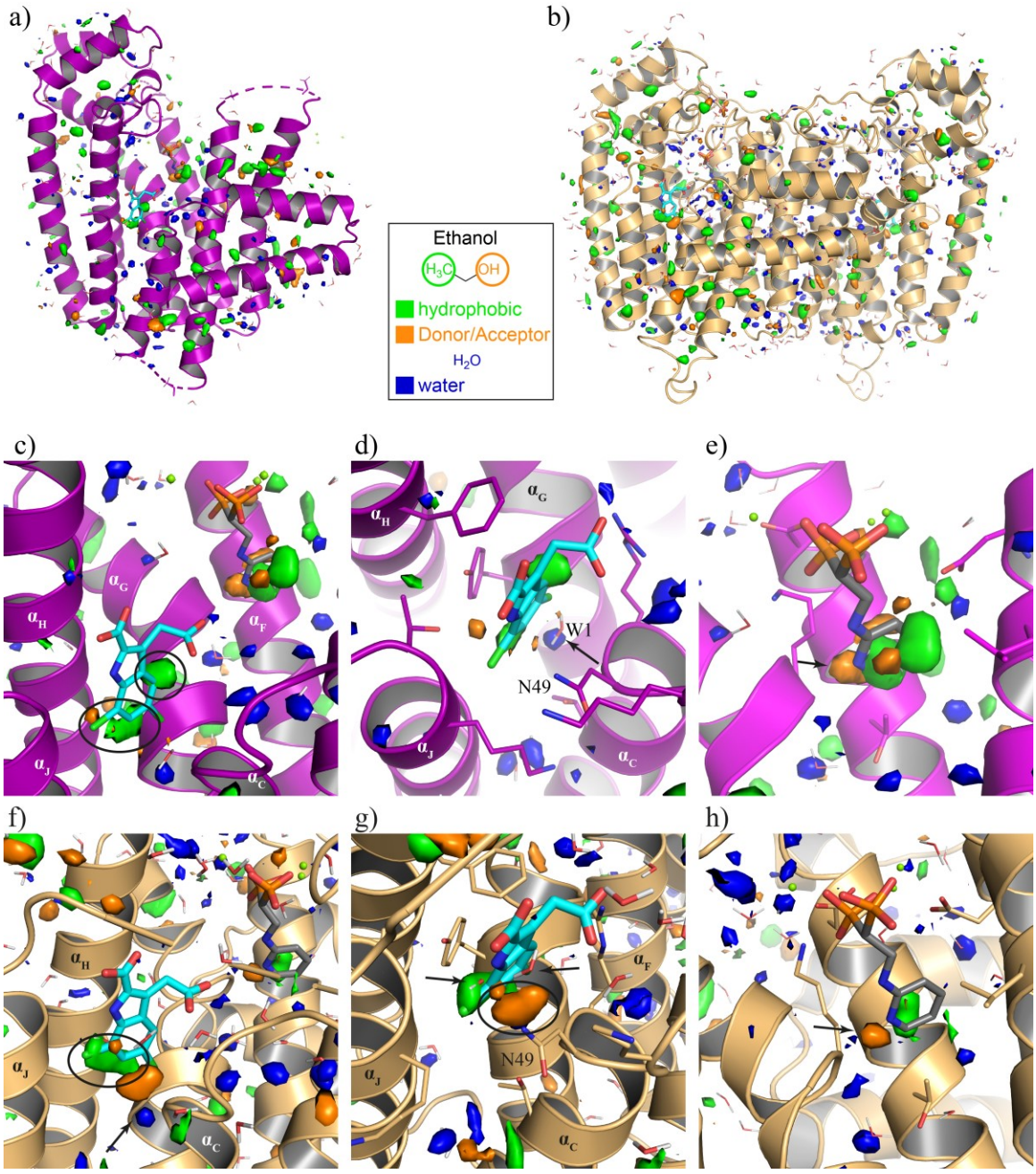


Figure 30: Hot spots of binding identified in *T. brucei* FPPS using MDMix. Result of molecular dynamic simulations with 20 % ethanol solvent mixture for 50 ns production run. Three replicas were performed with in-house solved *T. brucei* FPPS apo structure displayed in **(a)**,**(c)**, **(d)** and **(e)** and *T. brucei* FPPS in complex with bisphosphonate (PDB ID 2I19) displayed in **(b)**, **(f)**, **(g)** and **(h)**. Allosteric binder (PDB ID 3N5J, cyan) is overlaid to locate the allosteric site. Bisphosphonate (PDB ID 2I19, grey) is overlaid on in-house solved *T. brucei* FPPS apo structure (violet) and shown in *T. brucei* FPPS-bisphosphonate complex structure (PDB ID 2I19, yellow). Green spots represent hydrophobic interactions, orange spots represent donor/acceptor interactions and blue represent water sites. Overview of **(a)** in-house solved *T. brucei* FPPS apo structure and **(b)** *T. brucei* FPPS dimer structure (PDB ID 2I19) as Cartoon representation, **(c)** and **(f)** Close-up on the allosteric site and active site with allosteric binder and bisphosphonate as stick representation. Hydrophobic hot spots overlaying with allosteric binder are marked with black circles. **(d)** and **(g)** Top view on the allosteric site with residues forming the allosteric site shown as stick representation. Residue Asn49 is labelled. In **(d)** W1 is marked with an arrow and in **(g)** the two displaceable waters are marked with an arrow. In **(g)** the hot spot for polar interactions close to Asn49 is marked with a black circle. **(e)** and **(h)** Close-up on the bisphosphonate binding site with specific residues in the ligand binding site shown as stick representation. The hot spot for polar interactions close to the nitrogen of the pyridine scaffold is marked with an arrow.

In the active site, identified hot spots overlay well with the scaffold of the bisphosphonates (Figure 30 e, h). In Figure 30, bisphosphonate with a pyridine backbone (PDB ID 2I19) is overlaid. On the pyridine ring hydrophobic hot spots are identified in the open (apo protein, Figure 30 c, e) and closed conformation (PDB ID 2I19, Figure 30 f and h) and an additional hot spot for polar interactions is identified close to the nitrogen of the pyridine scaffold (Figure 30 e and h, marked with an arrow).

In summary, the identification of high affinity interaction spots in the allosteric site highlights promising features to accommodate a ligand and supports the previous sequence and structural alignment analysis concluding that the already reported allosteric site in human FPPS also exists in *T. brucei* FPPS.

4.2.3 Human allosteric binders on *T. brucei* FPPS analysed by NMR

Jahnke *et al.* described several fragment binders with affinities ranging from 80 nM to above 500 µM found in the allosteric site of human FPPS.[310] Ten reported allosteric binders and analogues thereof were selected and first tested in substance on *T. brucei* FPPS by protein-observed NMR (Figure 31). For *T. brucei* FPPS with a molecular weight of 42 kDa for the monomer predominantly clusters of peaks are observed. As no peak assignment is available, the binding mode of the studied ligands cannot be extracted from the NMR experiments. However, chemical shift perturbations upon compound addition were observed. Unexpectedly, human allosteric binders showed significantly lower affinities on *T. brucei* FPPS. Of the ten fragments tested on *T. brucei* FPPS, compound **117** (Figure 31) with a binding affinity of 7.1 µM on human FPPS resulted in *T. brucei* FPPS protein aggregation. The other nine fragments were classified as weak, medium and strong shifters according to their magnitude and number of chemical shift perturbations (for classification of shifter categories, see Chapter 2.8.5). Strikingly, compound **122** (Figure 31) with 80 nM affinity on human FPPS only showed weak shifts on *T. brucei* FPPS. Though, for compounds **118**, **119** and **121** (Figure 31) with a 6 µM, 7 µM and 200 nM binding affinity on human FPPS, strong chemical shift perturbations were detected and fragments were classified as strong shifters on *T. brucei* FPPS.

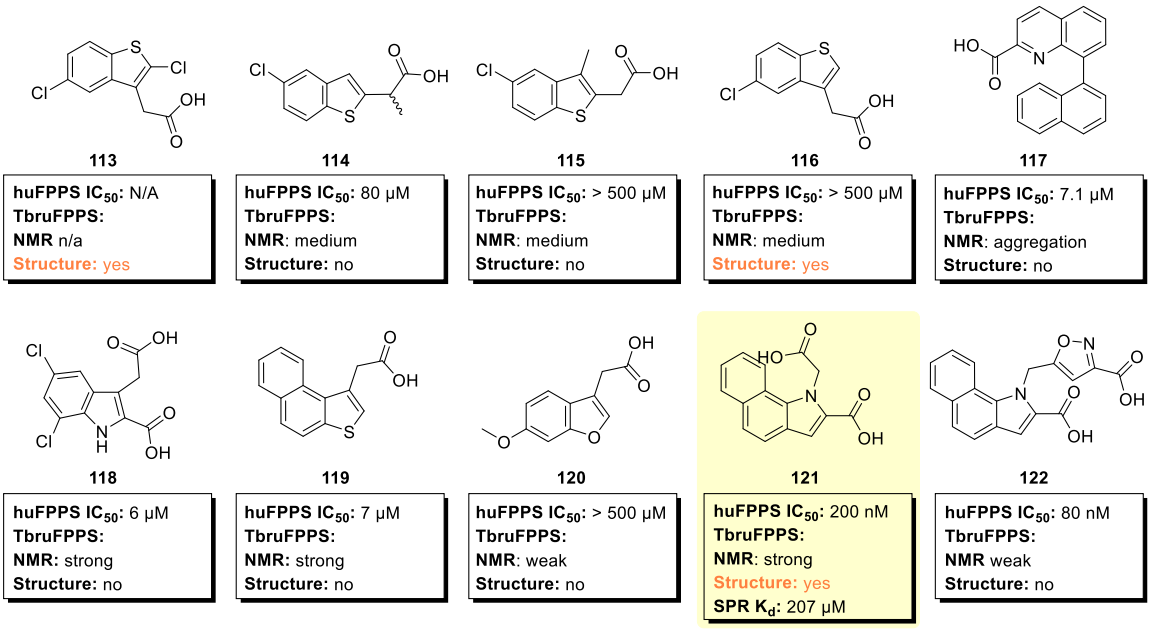


Figure 31: Reported allosteric binders on *T. brucei* FPPS. Ten human allosteric binders and analogues thereof are shown with their respective IC_{50} on human FPPS, the categorisation in aggregation, weak, medium and strong shifters as determined by [^{13}C, ^{1}H]-SOFAST-HMQC, the information if an X-ray *T. brucei* FPPS-fragment complex structure was solved (yes/no) and for compound **121** the K_d determined by SPR is noted.

In Figure 32 b, the protein-observed NMR spectrum of *T. brucei* FPPS with compound **121** is shown as representative example for a protein-observed NMR spectrum indicating a strong shifter. The spectrum of *T. brucei* FPPS with compound **117** in Figure 32 a displays the spectrum of a compound that initiated protein aggregation and precipitation. The latter is indicated by lower peak intensities in the spectrum measured in presence of the compound compared to the DMSO control.

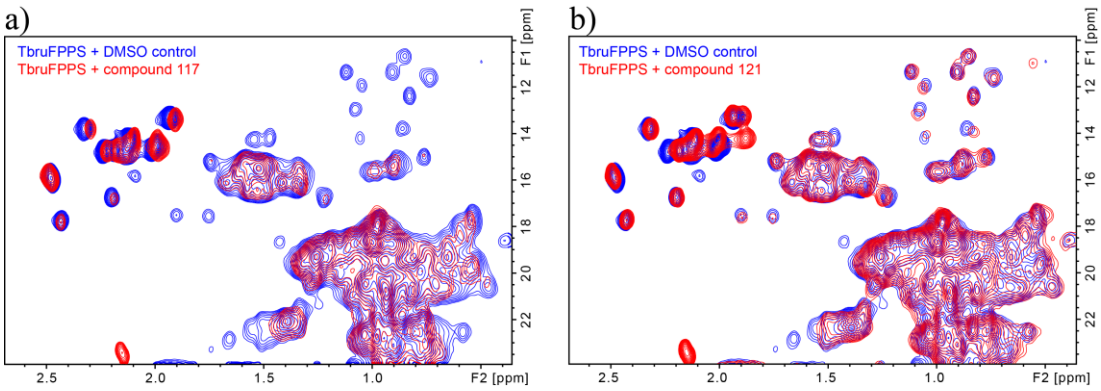


Figure 32: Protein-observed NMR spectra of *T. brucei* FPPS with human allosteric binders. The [^{13}C, ^{1}H]-SOFAST-HMQC spectra (32 scans) of 30 µM *T. brucei* FPPS with 1 mM compound measured at 310 K are shown. **(a)** Overlay of *T. brucei* FPPS with DMSO (control, blue) and with compound **117** (red). Lower peak intensities in spectrum of *T. brucei* FPPS with compound indicates protein precipitation and aggregation. **(b)** Overlay of *T. brucei* FPPS with DMSO (control, blue) and with compound **121** (red). Comparing the two spectra significant peak shifts are observed.

An SPR assay was setup to determine the binding affinity of human allosteric binders on *T. brucei* FPPS. K_ds of compounds **113** and **116** could not be determined and were above experimental test range (>500 mM). However, the K_d of compound **121** with an IC_{50} of 200 nM on human FPPS could be determined and was found to be 1000-fold weaker in binding with a K_d of 200 µM on *T. brucei* FPPS (Figure 33). This observation matches with the crystallographic results as significantly different binding modes for compound **121** in *T. brucei* FPPS and human FPPS were observed (see Discussion, Chapter 4.3).

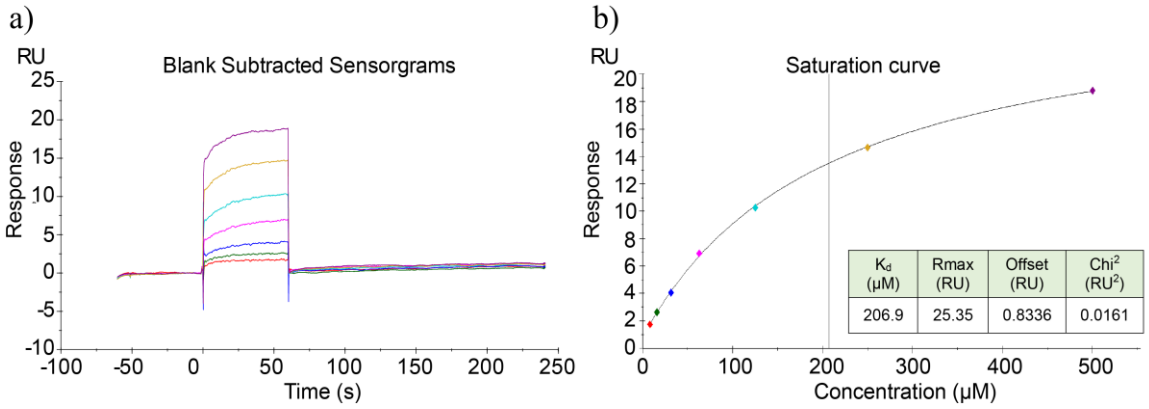


Figure 33: SPR results curve of human allosteric binder tested on *T. brucei* FPPS. Measurements were performed on a Biacore T200 at 23 °C. Biotinylated avi-tagged protein was immobilised on a SA chip and compound **121** was tested in an eight point 2-fold compound titration up to 500 µM. Samples were injected at a flow rate of 30 µL/min for 30 s followed by a dissociation of 60 s. **(a)** Blank subtracted sensorgrams of compound **121** titration: the time in seconds is shown on the x-axis and the response in RUs on the y-axis. **(b)** Saturation curve with a vertical line at the determined K_d. On the x-axis, the concentration in (µM) is shown and on the y-axis, the response in RUs.

4.2.4 Human allosteric binders bind in allosteric site on *T. brucei* FPPS

The sensitivity of X-ray crystallography allows detection of even very weak shifters and despite the results of protein-observed NMR experiments, all nine fragments excluding the compound that indicated aggregation of the protein were individually soaked into *T. brucei* FPPS crystals. Electron densities that indicated the presence of the ligands were observed for fragments **116**, **113,** and **121** (Figure 31) in the previously unknown allosteric site (Figure 34). Additional density maps for compounds **113**, **116**, and **121** can be found in the Appendix, Chapter 10.5.1, Figure 78.

Crystal structures were refined to a diffraction limit of 2.33 Å (compound **113**, occ. 0.76), 2.10 Å (compound **116**, occ. 0.80), and 2.55 Å (compound **121**, occ. 0.90). These fragments were also medium to strong shifters in the [^{13}C, ^{1}H]-SOFAST-HMQC NMR experiments. Compounds **113** and **116** exhibit the same binding mode and form two H-bonds to Arg50 and Asn49 through its carboxyl-group. Compound **121** has two carboxyl groups and also forms with one of them H-bonds to Arg50 and Asn49 and the second carboxyl group points towards the C-terminus. The sequence of the last four residues in the C-terminus is KRQR (residues 364 – 367) and

therefore, contains one lysine and two arginines which could form further H-bonds with the second carboxyl group to the protein backbone. However, these residues are not modelled in the presented structures as no clear electron density is observed.

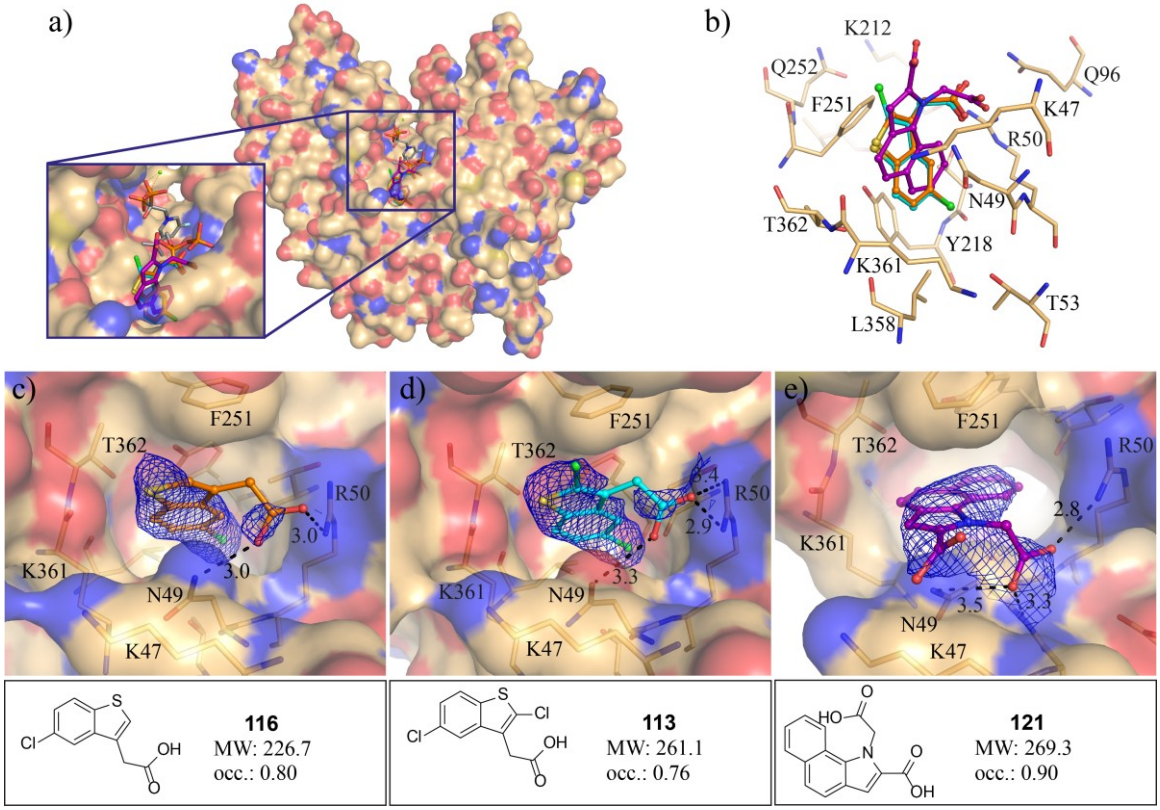


Figure 34: Human allosteric binders and analogues thereof on *T. brucei* FPPS. (a) Side view on FPPS monomer (PDB ID 6R37) with bisphosphonate and IPP (from PDB ID 3DYF) as a reference for location of active site. Close-up on allosteric site with overlay of compound **113**, **116** and **121** (PDB ID 6R38, 6R37, 6R39). **(b)** Overlay of compound **113**, **116**, **121** with ligand site residues in stick representation (PDB ID 6R37) **(c), (d), (e)** Close-up on *T. brucei* FPPS allosteric site. Protein backbone shown as surface representation and important residues in the ligand binding site are labelled and shown as sticks. Hydrogen-bonds are indicated by black dashed lines. 2Fo-Fc electron density map is contoured at 1 σ. **(c)** Compound **116** (PDB: 6R37). **(d)** Compound **113** (PDB ID 6R38). **(e)** Compound **121** (PDB ID 6R39).

4.3 Discussion

T. brucei FPPS has an overall sequence identity of 44 % and similarity of 64 % to human FPPS. In contrast to human FPPS, *T. brucei* FPPS has two additional loop insertions of residues 66 – 72 and 184 – 194 which do not align the allosteric and active site and whether these loops have any functional meaning is yet unknown.[300] By comparing the amino acid sequences and protein structures of human and *T. brucei* FPPS, the previously identified allosteric site in human FPPS is also found in *T. brucei* FPPS with a 70 % amino acid identity and 90 % similarity. Binding hot spots for hydrophobic and polar interactions were identified in this site using MDMix simulations. One strength of the MDMix approach is that water molecules are included and the displaceability of water molecules can be studied. In the MDMix simulation of

T. brucei FPPS in the open and closed conformation one water site was identified with low ΔG_{bind} indicating that this water is likely difficult to be replaced by a ligand. Furthermore, hot spots for hydrophobic interactions were identified for the open and closed conformation and another polar interaction spot was only found in the closed conformation of *T. brucei* FPPS. Hot spots were previously shown to overlay with key interactions of ligands, such as on Hsp90 and HIVp[495] and therefore, the identification of hot spots in the allosteric site of *T. brucei* FPPS predicts binding of putative ligands in this previously unknown site and overlays with human allosteric binders on *T. brucei* FPPS.

ConSurf analysis showed that the allosteric site is highly conserved across a diverse set of organisms. This result could be explained by the recent discovery of a feedback mechanism on human FPPS with allosteric inhibition of its own product, FPP.[320] Possibly, allosteric inhibition of FPPS by FPP not only plays an important role in the regulation of FPPS in the sterol biosynthesis pathway in humans but also in a plethora of other organisms. Therefore, the allosteric site is conserved to facilitate binding of FPP.

Interestingly, significant differences in the binding modes and binding affinities were observed when testing ten reported human allosteric binders on *T. brucei* FPPS. In fact, the biaryl-allosteric inhibitor **117** with an IC_{50} of 7.1 µM on human FPPS initiated protein aggregation and compound **122** with an 80 nM binding affinity on human FPPS was only categorised as a weak shifter in protein-observed NMR. No X-ray structures of *T. brucei* FPPS with these two compounds could be determined. Jahnke *et al.* did also not report on the structure of human FPPS with compound **122** to investigate its binding mode. In general, the two compounds are not of fragment size as compounds **117** and **122** have a molecular weight of 299.33 and 336.30, respectively. The amino acid differences of a tyrosine instead of a phenylalanine and a threonine instead of a isoleucine on *T. brucei* FPPS compared to human FPPS may be sufficient to make binding of these compounds in the allosteric site unfavourable or even impossible.

A *T. brucei* FPPS structure in complex with compound **116** was solved and the carboxyl groups of the binding mode in human FPPS and *T. brucei* FPPS are almost superimposable. In both structures the carboxyl groups form an H-bond to the side-chain nitrogens of Arg50 and Asn49. However, the chloro-benzothiophene moiety is by nearly 180° rotated in *T. brucei* FPPS compared to human FPPS (Figure 35).

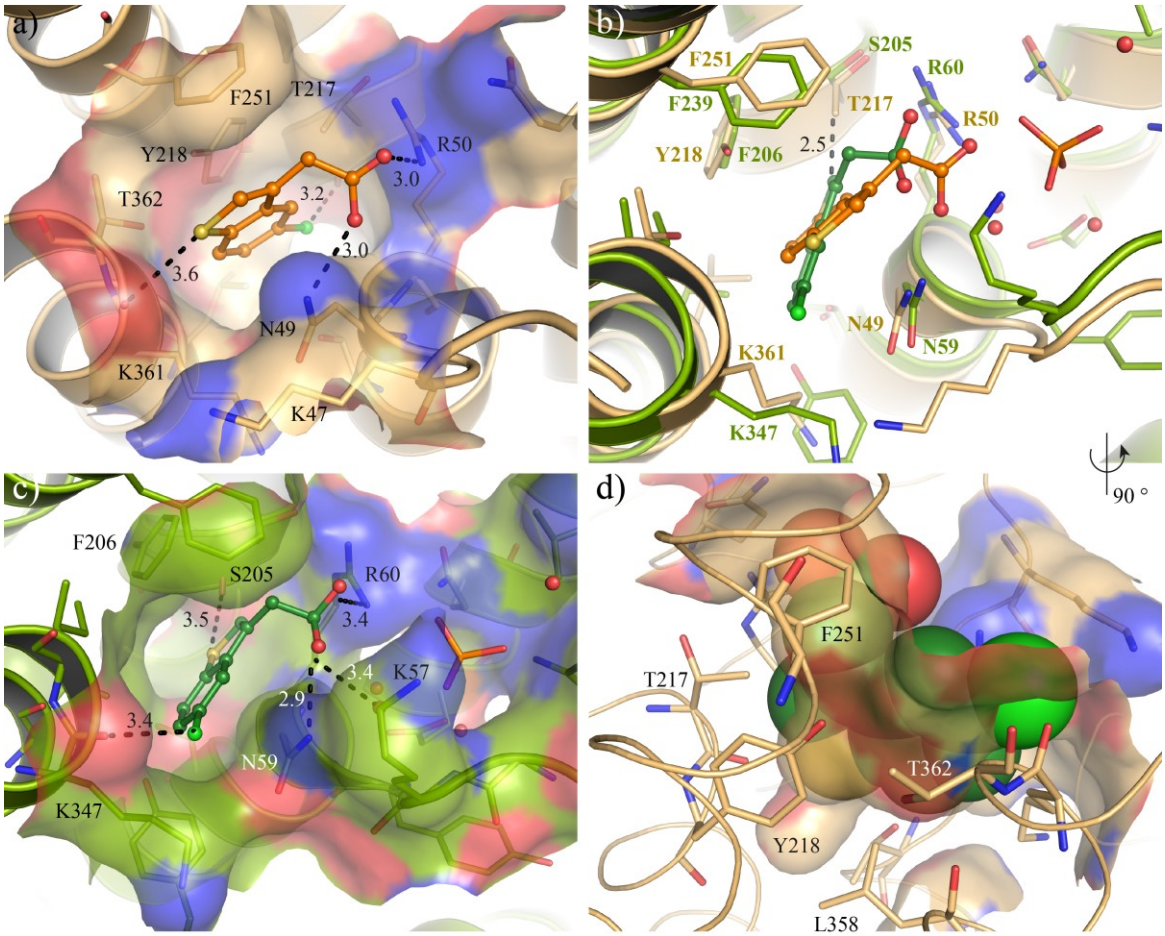


Figure 35: Benzothiophene-allosteric binder in human and *T. brucei* FPPS. *T. brucei* (PDB ID 6R37, light orange) and human FPPS (PDB ID 3N1W, green) with benzothiophene **116** in allosteric site. In all four images important residues in the ligand binding site are depicted as sticks and labelled. Compound **116** in *T. brucei* FPPS is shown in orange and compound **116** in human FPPS is shown as green ball and stick representation. Observed or putative interactions are indicated by black dashed lines. **(a)** Surface representation of the *T. brucei* FPPS allosteric site with compound **116**. **(b)** Overlay of *T. brucei* FPPS and human FPPS with compound **116**. Too close interaction of 2.5 Å of compound **116** human binding mode with T217 of *T. brucei* FPPS are depicted. **(c)** Surface representation of the human FPPS allosteric site with compound **116 (d)** Surface representation of the *T. brucei* FPPS ligand site with compound **116** in CPK representation with the binding mode adopted in human FPPS to visualise that the pose of the human enzyme would not fit in the parasitic one.

In the *T. brucei* FPPS complex, the sulphur of the benzothiophene shows noncovalent interactions with the backbone carbonyl oxygen of Lys361 and the chlorine atom is located on top of the backbone carbonyl oxygen of Thr217. In contrast, in human FPPS, the sulphur atom is located in 3.5 Å proximity to the backbone carbonyl oxygen of Ser205 and the chlorine atom points towards the C-terminus. It is oriented towards the backbone carbonyl oxygen of Lys347. At the position of Ser205 in human FPPS, Thr217 is found in *T. brucei* FPPS and therefore, the allosteric site is larger in human FPPS than in *T. brucei* FPPS. If compound **116** would adopt the same binding mode as in human FPPS, it would clash with the protein backbone as the benzothiophene moiety would then only have a 2.5 Å distance to Thr217 (Figure 35 b and d).

Additionally, the comparison of the binding modes and affinities of compound **121** in human FPPS and *T. brucei* FPPS revealed a 1000-fold lower binding affinity in *T. brucei* FPPS compared

to human FPPS. In human FPPS, upon compound **121** binding, the C-terminal helix α_A and the N-terminal helix α_J are shifted to form the binding site (Figure 36).

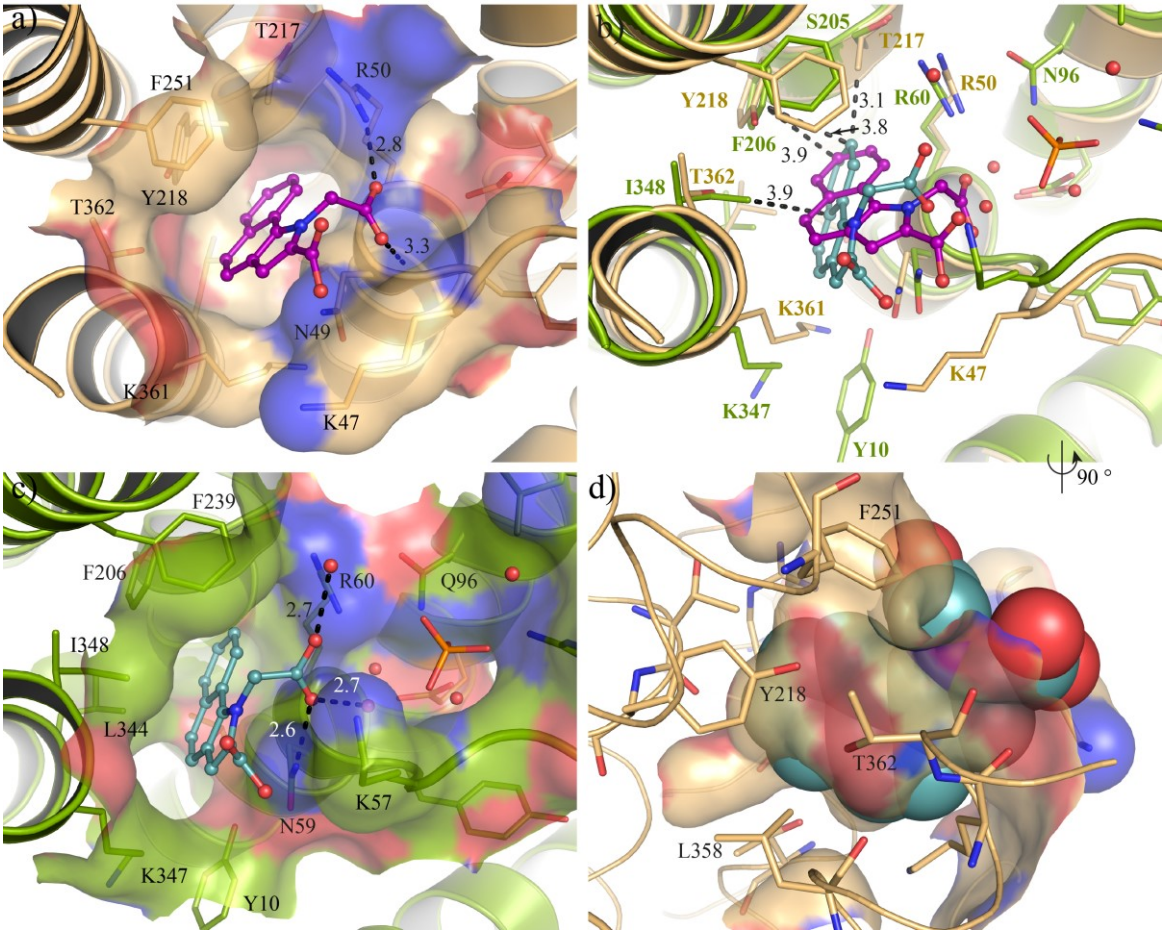


Figure 36: Benzoindole-allosteric binder in human and *T. brucei* FPPS. *T. brucei* (PDB ID 6R39, light orange) and human FPPS (PDB ID 3N6K, green) with benzoindole **121** in the allosteric site. In all four images important residues in the ligand binding site are depicted as sticks and labelled. Compound **121** in *T. brucei* FPPS is shown in magenta and compound **121** on human FPPS is shown in light blue as ball and stick representation. Observed or putative interactions are indicated by black dashed lines. **(a)** Surface representation of the *T. brucei* FPPS allosteric site with compound **121**. Ligand interactions with N49 and R50 are shown. **(b)** Overlay of *T. brucei* FPPS and human FPPS with compound **121**. Interactions of compound **121** in human FPPS with I348 and F206 are shown. Additionally, too close interactions of *T. brucei* FPPS T217 with compound **121** in the human allosteric binding mode are depicted. **(c)** Surface representation of human FPPS allosteric site with compound **121 (d)** Surface representation of the *T. brucei* FPPS ligand site with compound **121** in CPK representation with the binding mode adopted in human FPPS binding mode in sphere representation to visualise the pose of the human enzyme would not fit into the parasitic one.

The ligand establishes via its carboxylate two 2.7 Å H-bonds to two water molecules and one 2.6 Å H-bond to the nitrogen of the Asn59 side chain. Furthermore, the benzoindole moiety shows π-stacking with a 3.8 Å – 3.9 Å distal hydrophobic interaction with Phe206. Another 3.9 Å hydrophobic interaction is found with Ile348. The second carboxyl-group does not show any polar interactions in the refined structure (PDB ID 3N6K) but the C-terminal residues (RRK) are not visible in the electron density, nevertheless, additional electrostatic interactions could be possible with these basic residues.

In *T. brucei* FPPS, compound **121** is bound with 45° rotated orientation with its carboxyl group located closer to the Asn49 backbone. Interestingly, in human FPPS, Lys347 is moved out of space whereas Lys361 remains in *T. brucei* FPPS at the same position. *T. brucei* FPPS exhibits Tyr218 in place of Ph206 with close proximity to the benzoindole moiety and the Tyr218 hydroxyl group would clash with the benzoindole scaffold of **121**. The ligand is not ideally placed to form π-stacking and hydrophobic interactions with Phe206, which could explain the deviating binding affinities of compound **121** towards human and *T. brucei* FPPS. The binding mode of the ligand in human FPPS does not fit into the *T. brucei* FPPS binding site as C_{10} of the benzoindole moiety would undergo too close interactions of 3.1 Å with Thr217. Therefore, a significantly different binding affinity may result as a combination of the amino acid exchange of Phe206 vs. Tyr218 and Ser205 vs. Thr217 between the human FPPS and *T. brucei* FPPS resulting in a reduced space to properly accommodate the ligand with similarly efficient protein-ligand interactions.

5. NMR fragment screening identified different binding sites on *T. brucei* FPPS

5.1 Introduction

The FBDD field covers a wide range of fragment screening technologies with inherent advantages and disadvantages. It requires a careful choice of the appropriate screening technique for each individual target.[329, 348] This project focused on the identification of binding fragments on *T. brucei* FPPS using NMR and X-ray fragment screening. In a first step, the Novartis core and fluorine fragment libraries with 1818 compounds in total were screened by ligand-observed NMR. Hits that were validated by protein-observed NMR were subsequently evaluated by X-ray crystallography. Fragments were identified in various binding sites including the active and allosteric site using PanDDA analysis for the identification of even less populated binders in X-ray protein-ligand complex structures. Optimisation by SAR by archive/catalogue demonstrates fragment hopping of an active site fragment analogue into the allosteric site which clearly presents the challenges of fragment optimisation. This aspect of fragment-based drug discovery and the hit rate of NMR fragment screening including follow-up by crystallography will be addressed in this chapter.

5.2 Results

5.2.1 Diverse fragment hits identified for *T. brucei* FPPS by NMR screening

Fragment screening by ligand-observed NMR is a commonly applied approach to explore new binding sites and to identify fragments with even very weak affinities in an economic and efficient workflow.[535] The Novartis core fragment library containing 1336 compounds was screened by wLOGSY and T1ρ in mixtures of eight compounds on *T. brucei* FPPS. Screening in mixtures allowed the exploration of a large chemical space in short time with only 14 mg of protein for the full Novartis core fragment library. In this fragment screen, 85 compounds were categorised as fragment hits with strong effects of ≥20 % in T1ρ, and a positive wLOGSY signal (Table 28).

Table 28: Identified hits in NMR fragment screen of Novartis core fragment library.

Number of compounds	wLOGSY effect	T1ρ effect	Hit criteria fulfilled
85	positive	≥ 20 %	yes
173	positive	< 20 %	partly
4	negative	> 20 %	partly
1074	negative	none, <10 %	no

Figure 37 shows an example of a wLOGSY and T1ρ spectrum of an NMR mixture. The spectra of a T1ρ 10 ms and 200 ms delay experiment measured on a sample containing *T. brucei* FPPS and compound mixture were normalised to the compound reference spectra. After normalising, compound **131** (for chemical structure, see Figure 40) was identified as a hit with a 70 % T1ρ effect and a positive wLOGSY. In contrast, no effect was observed for compound **c** (chemical structure not shown) in the T1ρ experiment and a negative wLOGSY signal was detected, which classifies compound **c** as a non-binder.

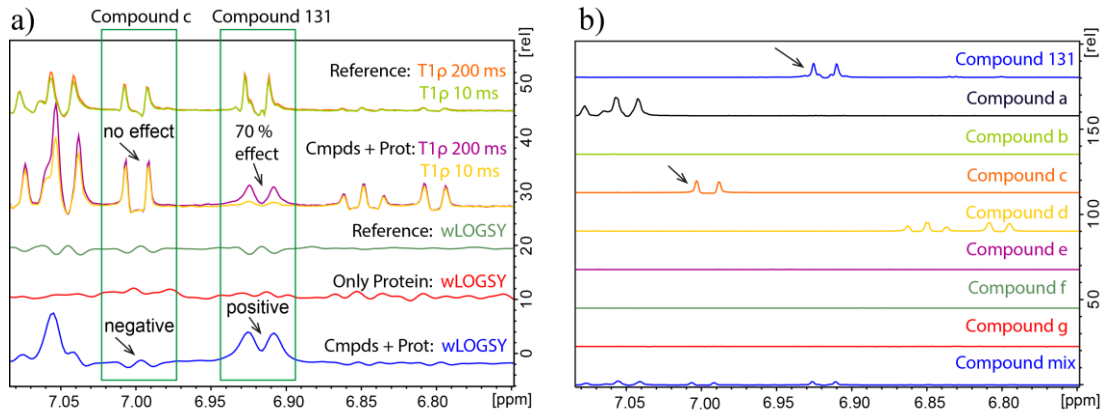


Figure 37: Ligand-observed NMR spectra of fragment mixture. Compounds reference spectra were recorded on samples containing 200 µM compound mixture on a Bruker Avance III 600 MHz quadruple at 296 K. Protein samples contained 10 µM *T. brucei* FPPS protein and 200 µM compound mixture and were recorded on a Bruker Avance III 600 MHz at 296 K. **(a)** NMR spectra shown from top to bottom: Overlay of T1ρ 200 ms (orange) and 10 ms (green) delay experiments of compounds reference, overlay of T1ρ 200 ms (magenta) and 10 ms (yellow) delay experiments of protein sample, wLOGSY of compounds reference (green), only *T. brucei* FPPS (red), wLOGSY of protein sample. **(b)** Individual 1D ^{1}H-zgesgp reference spectra of eight compounds in the compound mixture from top to bottom and 1D ^{1}H-zgesgp spectrum of compound mixture is shown. These reference spectra identify which peaks correspond to which compound. Chemical structures of compounds **a** – **g** are not shown.

After this first screening phase, a hit rate of 4.8 % was determined. As this work was in close synergy with a fragment-screening project on *T. cruzi* FPPS performed by Joy Petrick, NIBR, Basel, Switzerland, fragment hits identified for both targets were combined and compared. At this stage of the project, the overall goal was the identification of either a compound that binds to both trypanosomal targets or only to one target without any effect on the other target. Hence, identified fragment hits were categorised as “intersection” hits binding to both proteins or as “distinct” hits,

binding to only one protein. For this categorisation, *T. brucei* FPPS fragment hits that also partly fulfilled hit criteria for *T. cruzi* FPPS and are therefore not clearly "distinct" or "intersection" hits were excluded from follow-up in single compound NMR experiments. A compound that partly fulfilled hit criteria for *T. cruzi* FPPS had a positive wLOGSY and T1ρ effect of <20 % or a negative wLOGSY and a T1ρ effect >20 % on *T. cruzi* FPPS. When comparing *T. brucei* FPPS hits with *T. cruzi* FPPS hits, 29 *T. brucei* FPPS hits were excluded from the total of 85 *T. brucei* FPPS hits, based on the single compound follow-up criteria described above (Figure 38). Nine "distinct" hits for *T. brucei* FPPS, 59 "distinct" hits for *T. cruzi* FPPS and 47 "intersection" hits showing binding to *T. brucei* and *T. cruzi* FPPS were identified. Identified hits were classified in "distinct" hits and "intersection" hits to compare the number of identified hits on the two trypanosomal targets.

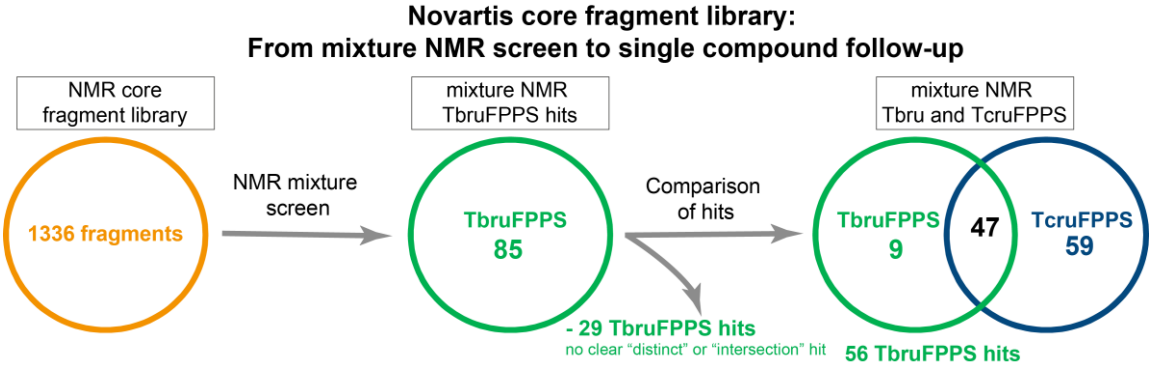


Figure 38: NMR fragment screen of Novartis core fragment library on *T. brucei* FPPS . The number of fragments screened is shown in an orange circle, *T. brucei* FPPS hits in a green circle and *T. cruzi* FPPS hits in a blue circle. From left to right: Number of fragments of the NMR core fragment library that were screened by ligand-observed NMR experiments in mixtures. Number of hits identified for *T. brucei* FPPS. By comparison of fragment hits between *T. brucei* and *T. cruzi* FPPS, *T. brucei* FPPS hits were excluded as these compounds partly fulfilled hit criteria on *T. cruzi* FPPS (positive wLOGSY and T1ρ effect of <20 % or a negative wLOGSY and a T1ρ effect >20 %) and could not be categorised as clear "intersection" (fulfilling hit criteria on *T. brucei* and *T. cruzi* FPPS) or "distinct" hits (only clear hits on either of the targets). The number of "distinct" fragment hits on *T. brucei* and *T. cruzi* FPPS and "intersection" hits of the two FPPS proteins compounds are shown on the right. The flow diagram ends with the combined 115 fragment hits identified by NMR fragment screening in mixtures of eight fragments.

The combined 115 fragment hits were followed-up by single compound ligand-observed NMR experiments on *T. brucei* to exclude mixture effects that could occur from a synergetic binding event of two fragments. When combining "distinct" *T. brucei* FPPS hits and "intersection" hits of *T. cruzi* and *T. brucei* FPPS, 56 *T. brucei* FPPS fragment hits were identified in the NMR mixture experiments from which 30 fragment hits were validated in single compound NMR experiments (Figure 39, second circle). This corresponds to a 54 % validation rate. Additionally, three hits that were "distinct" *T. cruzi* FPPS hits in mixture NMR experiments bound to *T. brucei* FPPS in single compound NMR experiments. This can occur due to competition of compounds for a binding site on the protein in the mixture and therefore, not resulting in a strong enough effect for the individual compound in the mixture.

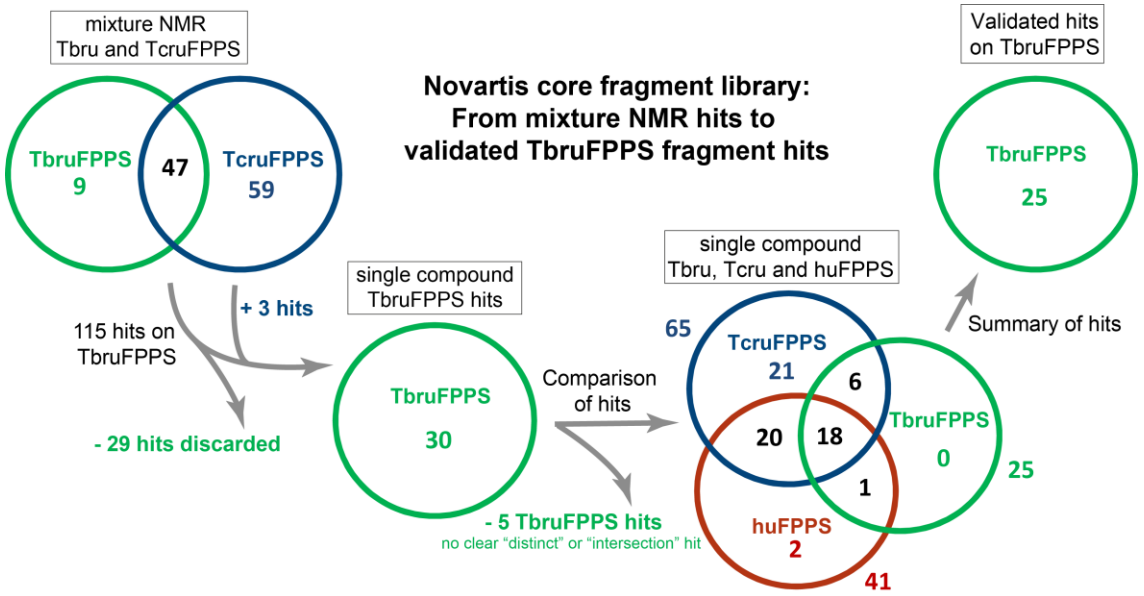


Figure 39: From mixture NMR hits to validated *T. brucei* FPPS fragment hits. The number of *T. brucei* FPPS hits is shown in a green circle, *T. cruzi* FPPS hits in a blue circle and human FPPS in a red circle. From left to right: The number of "distinct" fragment hits on *T. brucei* and *T. cruzi* FPPS and "intersection" hits of *T. brucei* and *T. cruzi* FPPS identified by mixture NMR screening experiments are depicted. Of the 56 *T. brucei* FPPS hits (nine "distinct" *T. brucei* FPPS hits and 47 "intersection" hits), 29 hits were discarded. Three additional hits bound to *T. brucei* FPPS in single compound ligand-observed NMR experiments but were "distinct" *T. cruzi* FPPS hits in mixture NMR experiments. Thus, thirty fragment hits were validated hits for *T. brucei* FPPS in single compound ligand-observed NMR experiments. By comparison of fragment hits between *T. brucei*, *T. cruzi* FPPS and human FPPS, five *T. brucei* FPPS hits were excluded and not considered further as thus compounds partly fulfilled hit criteria on *T. cruzi* FPPS (positive wLOGSY or T1ρ effect of <20 % and ≥10 %) and could not be categorised as clear "intersection" or "distinct" hits. In combination of testing all compounds on *T. brucei*, *T.cruzi* and human FPPS, 65 fragment hits bound to *T. cruzi* FPPS, 41 fragment hits bound to human FPPS and 25 fragment hits bound to *T. brucei* FPPS (third circle diagram). On the right, the total number of validated *T. brucei* FPPS hits is shown.

All 115 fragment hits were also tested on *T. cruzi* and human FPPS. Testing identified fragment binders in human FPPS was included at this stage to investigate selectivity of fragments on the three targets. The combination of validated fragment hits on *T. brucei* FPPS and *T. cruzi* as well as including hits binding to human FPPS resulted in new "distinct" and "intersection" hits (Figure 39, third circle diagram). No *T. brucei* FPPS "distinct" hit could be reported after the validation in single compound ligand-observed NMR experiments because either a previously "distinct" hit also bound to, or at least partially met hit calling criteria for *T. cruzi* FPPS. All hits that partially met hit calling criteria were excluded from the final hit list which was five fragment hits out of the 30 validated *T. brucei* FPPS hits. In the combination of fragment hits on the three tested targets, 18 fragments bound to the three FPPS, six fragments bound to *T. brucei* and *T. cruzi* FPPS and one fragment bound to *T. brucei* and human FPPS. Two fragments were "distinct" human FPPS binders although the tested compounds were initially chosen to bind to at least one trypanosomal target in the mixture NMR fragment experiments (Figure 39, third circle diagram). However, this can be the result of no effect on both trypanosomal targets in the single compound NMR experiments.

As no "distinct" *T. brucei* FPPS hit was part of the final hit list, all 25 validated fragment hits (Figure 39, on top right) were further followed-up by protein-observed NMR in order to get all the structural information that can be extracted from the fragments. Selectivity towards human FPPS can be still achieved at a later stage. Validated 25 fragment hits showed structural diversity containing fragments with chloride-, hydroxyl-, amine-, methyl sulfonyl-moieties and chemical scaffolds, such as imidazole, pyrimidine, pyridine and bi- and tricyclic systems. Sixteen out of the 25 fragment hits are publicly known and shown in Figure 40.

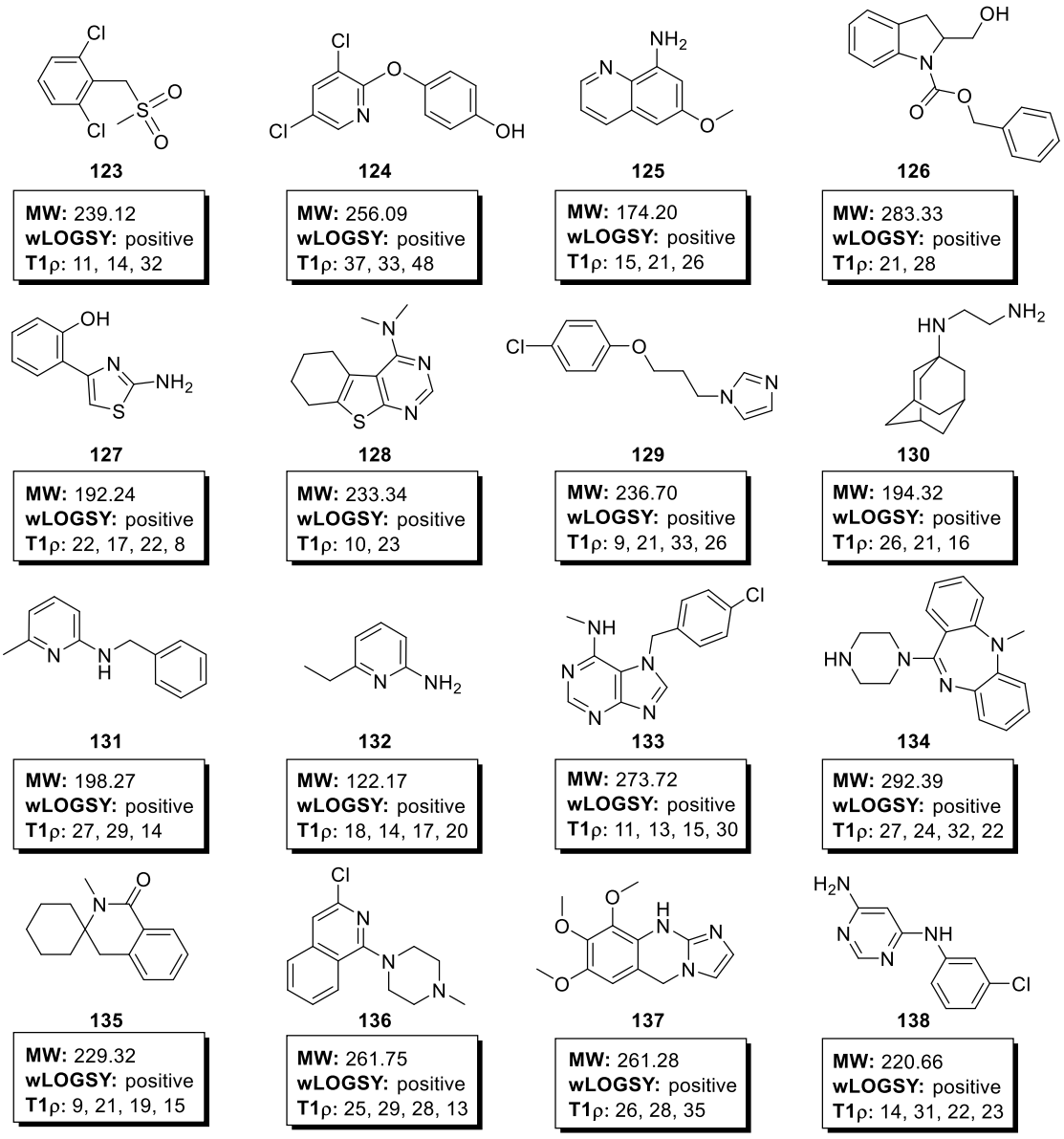


Figure 40: Fragment hit structures of the Novartis core library screen. Sixteen structures of fragment hits and their molecular weight are depicted. Results of single compound NMR experiments with 10 µM *T. brucei* FPPS protein and 200 µM compound are shown. For each compound, the wLOGSY and T1ρ results from the single compound experiment are shown. The wLOGSY result is either positive or negative and qualitatively judges the compound peaks in a sample with *T. brucei* FPPS protein and compound, as compared to a compound reference. The T1ρ results are the reductions of the intensities of compound peaks in %, when comparing T1ρ experiments with 200 ms and 10 ms delays, normalised to the same experiment without protein present. The reduction for at least two peaks is noted.

In addition to the Novartis core fragment library, the Novartis fluorine fragment library with 489 compounds was screened in mixtures of up to 32 compounds by ^{19}F-CPMG experiments, and 34 compounds were identified as binding with a ^{19}F-CPMG effect ≥20 %, which corresponds to a hit rate of 7.0 %.

Table 29: Identified hits in NMR fragment screen of Novartis fluorine fragment library.

Number of compounds	^{19}F-CPMG effect	Hit criteria fulfilled
34	≥20	yes
50	≥ 10 and ≤ 20	no
444	none, < 10	no

The hit list was combined with *T. cruzi* FPPS hits and 72 compounds were binding on either *T. brucei*, *T. cruzi* or both targets (Figure 41). In contrast to the NMR fragment screen of the Novartis core fragment library, no hits were excluded at this stage and all hits were classified in either "distinct" hits or "intersection" hits even if hits also partly fulfilled hit criteria on the respective other FPPS target. Interestingly, high hit rates were observed on *T. cruzi* FPPS and therefore, the threshold for a hit was chosen at a ^{19}F-CPMG effect of >40 % to enable a fast and efficient follow-up.

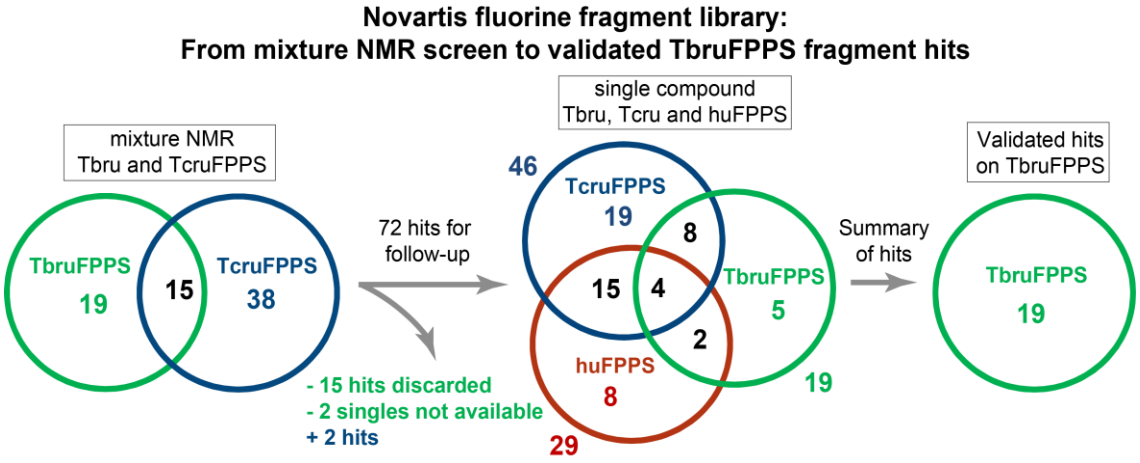


Figure 41: From NMR fragment screening to validated *T. brucei* FPPS fragment hits. *T. brucei* FPPS hits are shown in green circles, *T. cruzi* FPPS hits in blue circle, human FPPS hits in red circle and intersecting hits in black. From left to right: Number of "distinct" fragment hits on *T. brucei* and *T. cruzi* FPPS and "intersection" hits of the two FPPS proteins identified by NMR fragment screening in mixtures of eight compounds are shown. Single compound ligand-observed NMR follow-up result on *T. brucei*, *T. cruzi* and human FPPS is illustrated in the centre and summary of validated hits on *T. brucei* FPPS is shown on the right.

Seventy of the 72 identified hits (the two missing ones were *T. brucei* hits) were available for follow-up in single compound NMR experiments for hit validation. Out of the combined 32 *T. brucei* FPPS "distinct" and "intersection" hits tested, 17 hits were validated in single compound experiments, which corresponds to 51 % validation rate. Additionally, two compounds

that were only *T. cruzi* hits in NMR mixture experiments were identified as binders in the single compound NMR experiment, which results in 19 hits on *T. brucei* FPPS after single compound NMR validation experiments. The combination of hit lists on *T. brucei* and *T. cruzi* FPPS and the addition of testing the compounds on human FPPS identified new groups of "distinct" and "intersection" binders. The hit criterion for human FPPS was chosen to be a ^{19}F-CPMG effect of >30 %, different from *T. brucei* FPPS and *T. cruzi* FPPS. In Figure 41, the summary of identified hits is shown based on the different hit criteria for the three different targets. Eight compounds were identified as "distinct" hits on human FPPS although compounds for single compound NMR experiments were selected due to binding events on *T. brucei* or *T. cruzi* FPPS in NMR mixture experiments. In total, five compounds were "distinct" hits for *T. brucei* FPPS and 14 compounds were "intersection" hits with either *T. cruzi* or human FPPS or showed binding on all three targets (Figure 41). Identified and publicly known compounds **139** – **149** are shown in Figure 42.

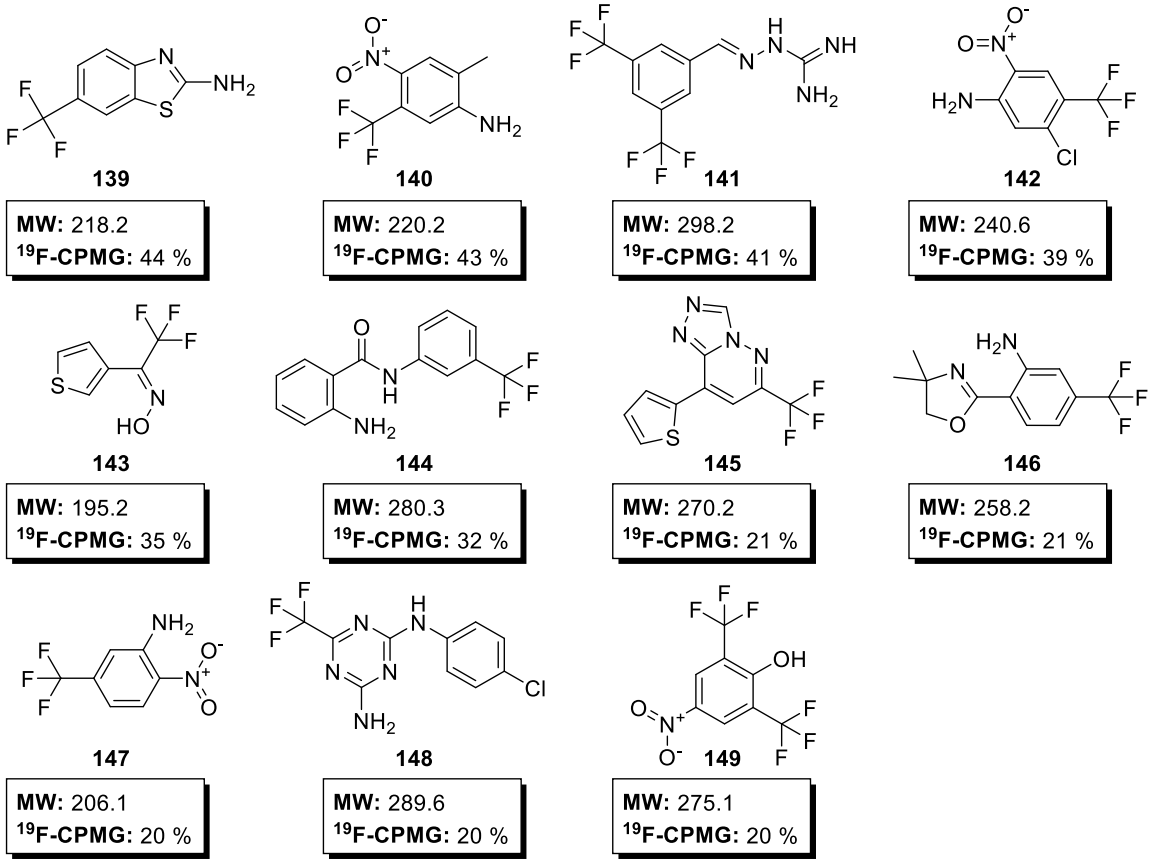


Figure 42: Fragment hit structures of the Novartis fluorine library screen. Eleven structures of fragment hits and their molecular weight are depicted and for each compound, the ^{19}F-CPMG result is shown. For this, the peak intensities of the *T. brucei* FPPS with compound ^{19}F- CPMG spectra are normalised to the only compound reference. Then, the minimising percentage of the fluorine peak comparing the ^{19}F- CPMG 80 ms and 400 ms spectrum of the *T. brucei* FPPS protein with compound sample is noted.

5.2.2 NMR screening hit validation by protein-observed NMR

The 25 fragment hits of the Novartis core library and the 19 fragment hits of the Novartis fluorine library, which were detected in mixture NMR experiments and followed-up by single compound NMR, were further validated by protein-observed NMR as described in Chapter 2.8.5. A control experiment with the addition of DMSO was performed to exclude false positive fragment hits and one peak shifted upon DMSO addition (Figure 43). For this reason, the spectrum of *T. brucei* FPPS apo protein with the same amount of DMSO as in the compound NMR experiment was used as an apo protein reference spectrum. DMSO-sensitive peaks were not considered for compound binding determination. All 25 fragment hits of the Novartis core library and 19 fragment hits of the Novartis fluorine library showed peak shifts upon compound addition and were therefore considered validated in [^{13}C, ^{1}H]-SOFAST-HMQC experiments. Compounds were categorised as weak, medium and strong shifter and spectra overlays of *T. brucei* FPPS DMSO control with the addition of a weak, medium and strong binding compound are visualised in Figure 43. Six compounds of the Novartis core fragment library and one compound of the Novartis fluorine library showed peak shifts but at the same time reduced peak intensities indicating protein precipitation (Table 30).

Table 30: Fragment hits categorised as weak, medium and strong shifters according to protein-observed NMR.

	Nr. of hits	Weak	Medium	Strong
Novartis core library	25	7	15 (6)[1]	3
Novartis fluorine library	19	4 (1)[1]	10	5

[1] Compounds also showed protein precipitation.

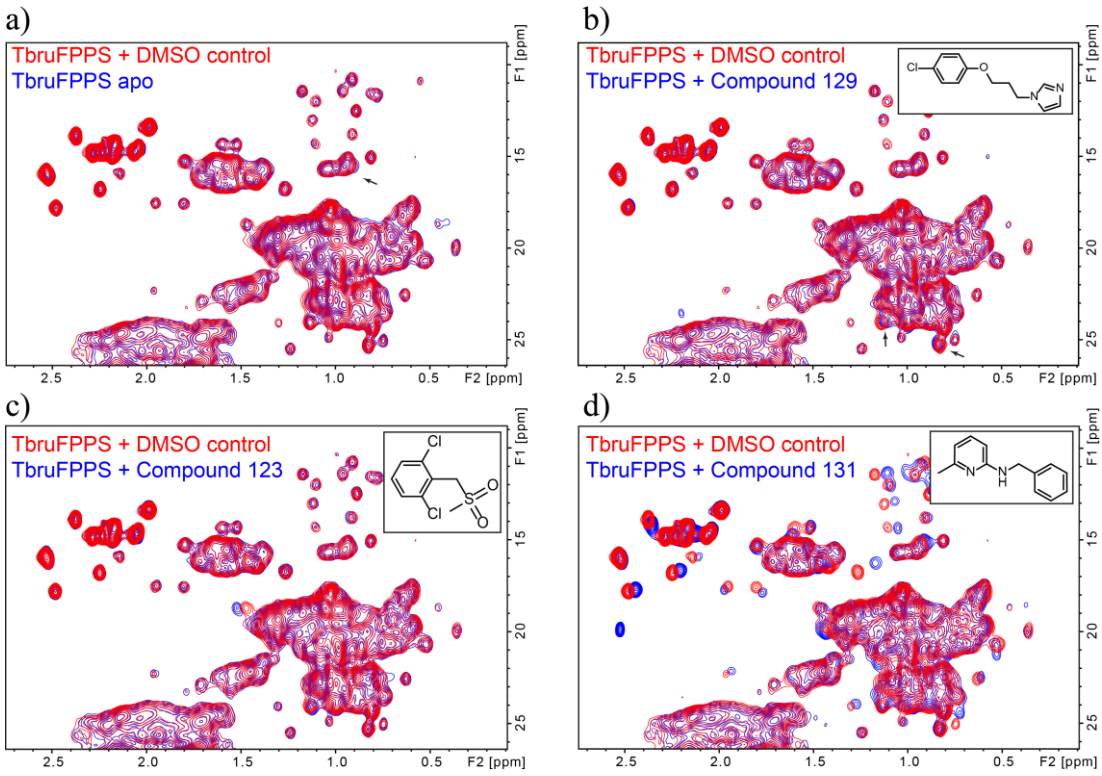


Figure 43: Protein observed NMR spectra of apo protein, DMSO reference and compounds 123, 129, 131. [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) with 30 µm *T. brucei* FPPS protein measured at 310 K are shown. **(a)** Overlay of *T. brucei* FPPS protein without (blue) and with 1 % v/v d_6-DMSO (red). Overlay of *T. brucei* FPPS protein DMSO control (red) with **(b)** compound **129**, **(c)** compound **123** and **(d)** compound **131** at 1 mM compound concentration are shown. In **(a)** and **(b)** weak shifts are indicated by arrows.

5.2.3 K_d estimation of fragment hits by NMR

In order to estimate the K_d of validated fragments, protein-observed NMR titration experiments were performed. Five validated hits identified from the Novartis core fragment library categorised as strong shifters were chosen for a 2-fold dilution row and chemical shift perturbations were analysed. As expected for low affinity fragments, tested fragments were in the fast exchange regime with k_{off} rates greater than the chemical shift differences. Peak shifts from the free state to the bound state were followed. Spectra were overlaid and curve fitting was performed in fitKD (Figure 44). Saturation was observed towards 2 mM or 4 mM compound concentration.

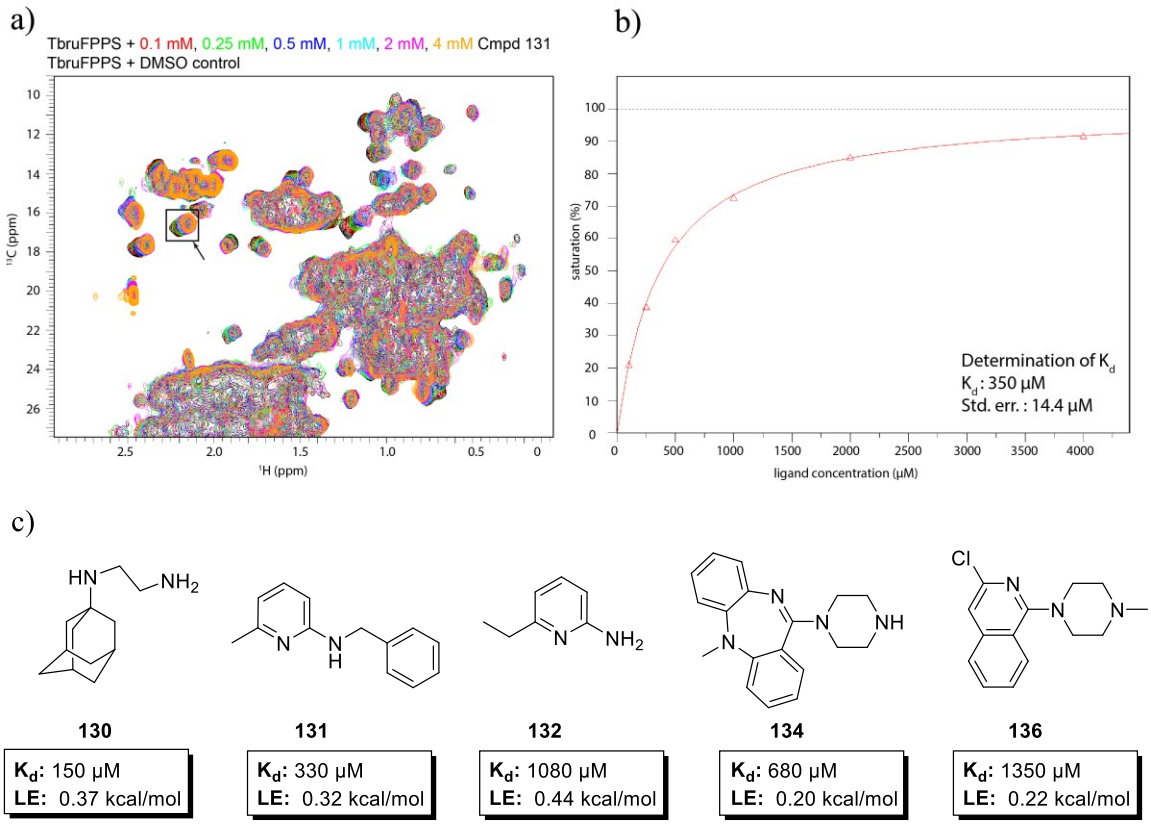


Figure 44: K_d plot and estimated K_d by NMR of five compounds. (a) Overlay of [^{13}C, ^{1}H]-SOFAST-HMQC spectra measured of samples with 30 µm *T. brucei* FPPS and the following concentrations: 0.1 mM, 0.25 mM, 0.5 mM, 1mM, 2 mM, and 4 mM of compound **131** at 310 K. Peak shift for which in **(b)** the K_d curve is indicated with an arrow. **(b)** Ligand concentration (µM) is plotted against the saturation (%) and K_d curve as determined by titration experiment is shown in red. **(c)** Five fragment hits with their respective K_d estimated by [^{13}C, ^{1}H]-SOFAST-HMQC titration experiments measured of sample with 30 µM *T. brucei* FPPS and compound concentrations ranging from 0.1 mM – 2 (4) mM at 310 K are shown. Ligand efficiency (LE) based on kcal/mol of the compounds was calculated and is noted.

Between two and eight peak shifts were selected from NMR titration experiments and the final K_d was estimated as an average of K_ds from selected peaks. The K_d plot of one peak shift upon addition of compound **131** to *T. brucei* FPPS protein is shown in Figure 44, but the final K_d of **131** was estimated based on four different peaks, which explains the variation between the K_d in the plot and the final K_d given. The estimated K_d of tested compounds ranged between 146 µM and 1.62 mM and ligand efficiency (LE) was calculated as followed based on kcal/mol:

Compound **131:**

$$LE = RT\,x\ln\frac{K_d}{HAC} = -0.59\,x\,\ln\frac{327.5}{15} = 0.32\,kcal/mol$$

For compound **131**, a ligand efficiency of 0.32 kcal/mol was calculated. The ligand efficiency of compounds **130** – **132**, **134,** and **136** ranged between 0.20 and 0.44 kcal/mol (Figure 44). Interestingly, compound **132**, a smaller analogue of **131**, was identified with a K_d of around 1 mM

and a ligand efficiency of 0.44. This high ligand efficiency expresses a strong binding free energy likely resulting from the interactions of two nitrogen functionalities and suggests a favoured interaction between the protein and the compound. However, the binding location and the relevance of the binding mode could not be determined by protein-observed NMR and therefore, X-ray crystallography experiments were performed.

5.2.4 NMR fragment hits probe various binding sites - Identification of non-bisphosphonate allosteric site and active-site binders

For all 25 fragment hits of the Novartis core fragment library and the 15 medium and strong shifters of the Novartis fluorine fragment library, crystal soaking experiments were setup as described in Methods, Chapter 2.9.6. Different soaking protocols with 5 mM to 100 mM compound concentration for 30 min up to 2 days were performed over the course of this thesis. For each fragment, at least two crystal soaks were harvested and in some cases also co-crystals. In total, 413 crystal soaks were harvested with compounds of the Novartis core and fluorine fragment library and 163 X-ray data sets were collected at a resolution <3.3 Å (Figure 45).

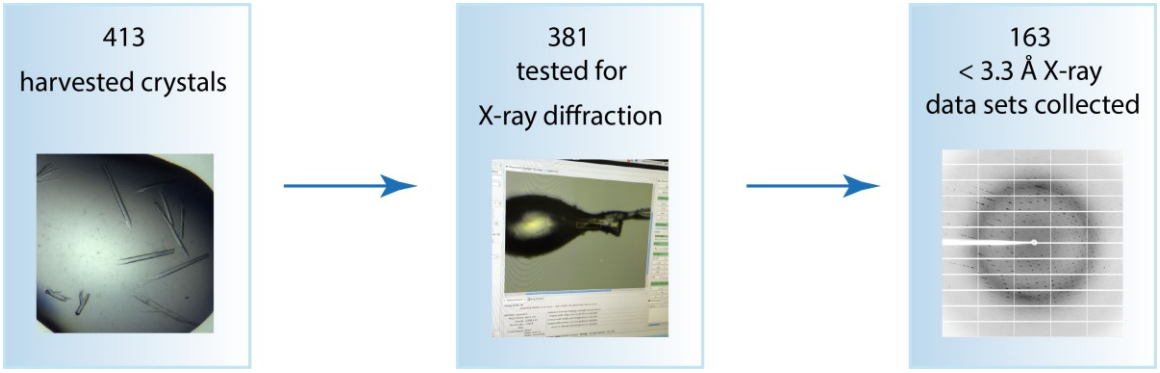


Figure 45: Crystallization and data collection outcome of Novartis fragment hits. From left to right the number of harvested crystals, tested crystals for X-ray diffraction and X-ray data sets collected with a diffraction limit below 3.3 Å of crystal soaks with NMR fragment hits is visualised.

Data sets were processed and analysed using Global Phasing Pipedream and PanDDA software. All X-ray data sets collected at PXII (SLS Villigen, Switzerland) were combined into one single PanDDA analysis. This includes data sets of apo crystals as well as crystals soaked with human allosteric binders, fragment analogues and fragments synthesised over the course of this thesis. For best results, PanDDA analysis requires at least 30 apo protein X-ray data sets and the analysis further improves if the PanDDA run is setup using more data sets.[463-464] The 357 data sets refined and analysed in PanDDA showed a diffraction limit ranging from 1.65 to 3.90 Å (only one data set with a diffraction limit above 3.3 Å) and a mean diffraction limit of 2.38 Å. R_{free} values ranged from 0.245 to 0.488 and six data sets with R_{free} values above 0.40 after refinement were excluded by the PanDDA software for further analysis (Figure 46). Data sets showed higher

R-values than expected for their high diffraction limit due to anisotropy of the needle-shaped crystals. The c axis of crystals was on average 341.9 Å, in the same range as for the apo protein crystals. The data collection and analysis with such a long cell axis is discussed in Chapter 3.3. For additional information on cell volume, axis length and angles, see Appendix, Chapter 10.4.1.

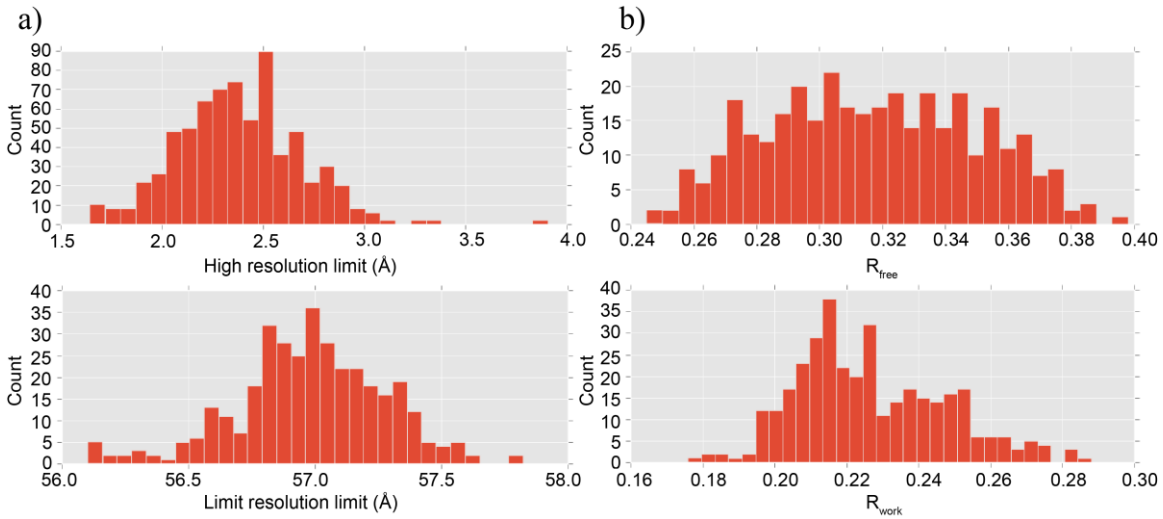


Figure 46: Diffraction limit and R_{work}/R_{free} value distribution of analysed data sets. On the left: The bar diagram shows on the x-axis the high resolution limit in Å (top) and the low resolution limit in Å (bottom) and on the y-axis the count of respective X-ray data sets. On the right: The bar diagram shows on the x-axis the R_{free} value (top) and the R_{work} (bottom) and on the y-axis the count of respective X-ray data sets.

The remaining 357 data sets were processed by *PanDDA.analyse* and 137 data sets were classified as interesting. A data set is qualified as interesting when at least one PanDDA event (difference density) was found while compared to the reference data set. In total, 194 PanDDA events were identified in 13 different sites. A PanDDA site is determined as a cluster of PanDDA events that are grouped together. In Figure 47, the different sites are visualised on the cartoon representation of a *T. brucei* FPPS structure that was solved as part of this thesis.

PanDDA events were manually inspected and in 31 PanDDA events, ligands were modelled and resulted as distributed over five sites. The remaining PanDDA events were artefacts of low quality electron density maps or additional blobs in the event maps which were too small to allow for the assignment of the soaked compound. Thus, they were discarded. Fitted ligands were grouped in low confidence, medium and high confidence hits according to their fit into the respective density and data sets were exported using *PanDDA.export* (Figure 47).

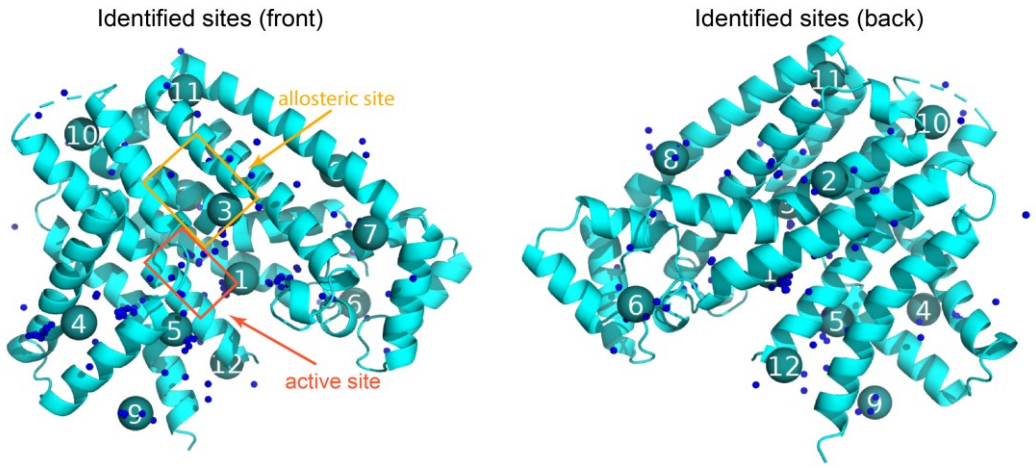


Figure 47: PanDDA summary of X-ray data for the FBS using Novartis fragments. From top to bottom: The figure illustrates the number of identified PanDDA events and the active (red) and allosteric (yellow) site is labelled. PanDDA sites with events are overlaid on *T. brucei* FPPS protein structure in Cartoon representation (front and back side). Each blue dot distributed over the protein represents one PanDDA event of one data set. Each labelled sphere corresponds to one PanDDA site which is a cluster of PanDDA events and results in 12 different PanDDA sites.

Fitted ligands in 31 PanDDA events were a combination of compounds from the Novartis core and fluorine fragment library, human allosteric binders, fragment analogues and synthesised compounds (Figure 48). Structures of *T. brucei* FPPS in complex with human allosteric binders had already been deposited in the PDB as part of this thesis and are not further discussed in this chapter (see Chapter 4.2.4). Out of the four PanDDA events fitted with fragment follow-up compounds, only one data set with a PanDDA event was successfully refined and the other data sets with one or more of these criteria were discarded: Ligands that were modelled with low confidence and could not be modelled unambiguously, and ligands with negative electron density after refinement. The refined *T. brucei* FPPS structure in complex with the one follow-up compound will be presented and discussed as part of the fragment analogue search later in the thesis (Chapter 5.2.8).

On exported X-ray data sets with modelled ligands, refinement was performed with PHENIX on an ensemble model of the ground state without the ligand and bound state with the ligand using GIANT SCRIPT, which is part of PanDDA. After the first PHENIX refinement cycle, the ground and bound state model was splitted. The ground state was modelled in the ground state map, and the bound state containing the ligand was modelled in the bound state map. If a residue had a different conformation in the ground state than in the ligand-bound state, then the occupancy of the two conformations was matched with the occupancy of the ligand. The ensemble model of the ground and bound state contains an overlay of both conformations with the respective occupancies. In the refinement process, data sets with one or more of the criteria already described for discarding data sets with follow-up compounds were excluded. Out of eight unique ligands from the Novartis core and fluorine fragment library, seven X-ray *T. brucei* FPPS-compound complex structures were solved with compounds in four different binding sites (Figure 48). The crystal structure of one additional unique ligand in the allosteric site was discarded as during refinement negative difference

electron density was observed around the ligand and an unambiguous modelling of the ligand was not possible.

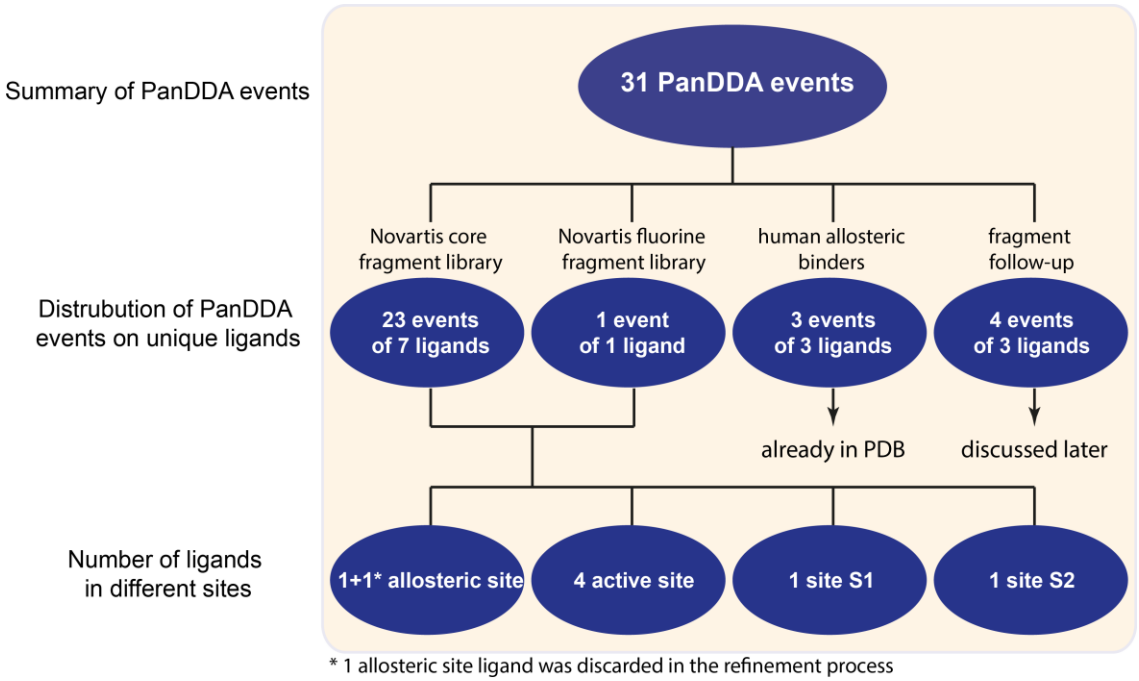


Figure 48: Flow diagram of compounds identified in different *T. brucei* FPPS binding sites. The flow chart starts with the number of PanDDA events and shows of which crystal soaks these events were identified. It further summarizes the number of unique ligands from the Novartis core and fluorine fragment library and visualises as a final the distribution of the ligands at the different binding sites. The crystal structure of one unique ligand was discarded during refinement.

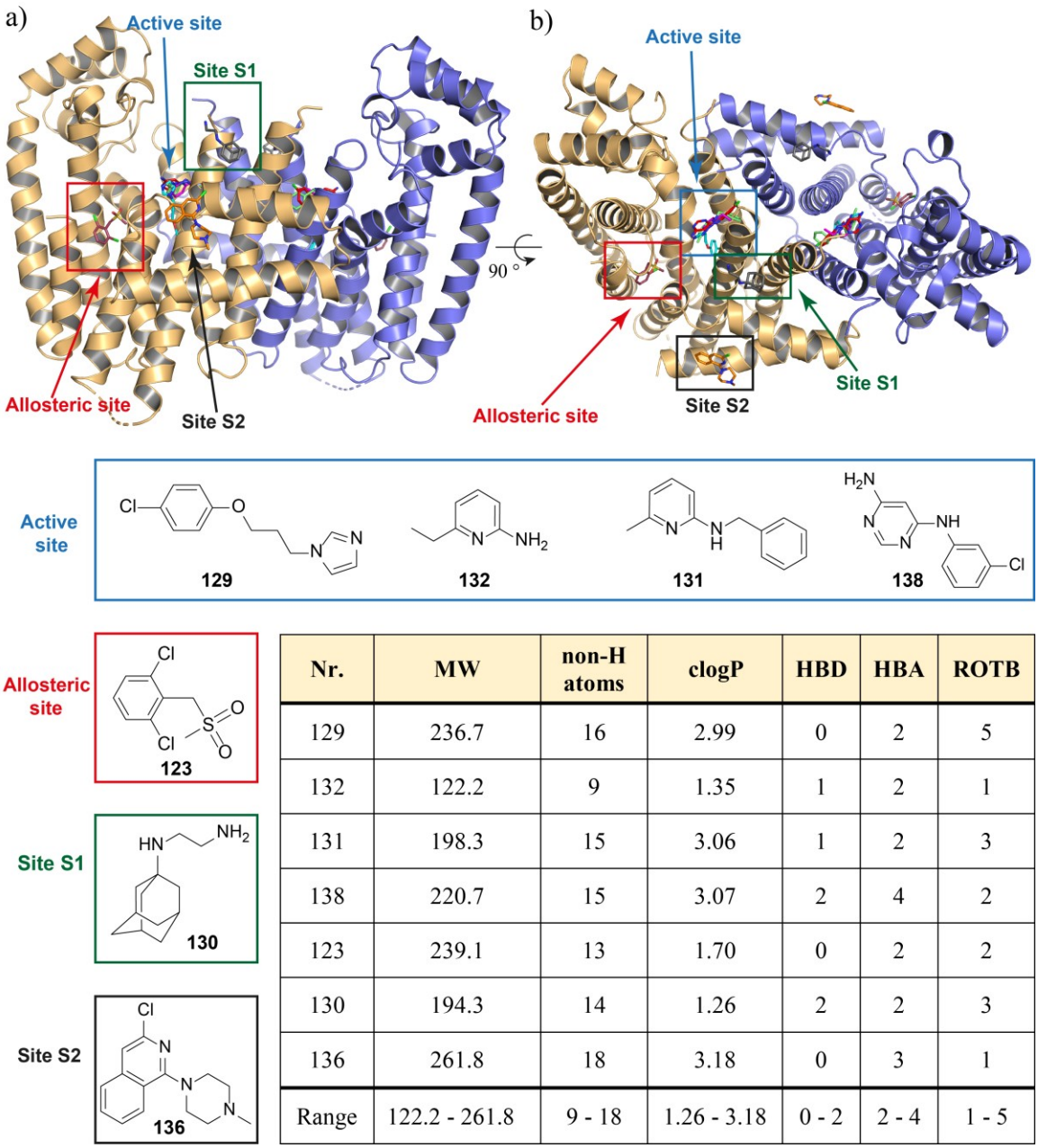
Soaking protocols of X-ray crystallography experiments varied largely (Table 31) and for each compound different soaking protocols were tested depending on the crystal stability in presence of the respective compound. Overall, protein-ligand crystal structures were solved from different crystallization conditions and cryo protectant strategies with compound concentrations ranging from 5 mM to 50 mM for 3 hours to 1 day. For compound **131**, soaking and co-crystal experiments were performed. Co-crystals grew within three days and one co-crystal structure exhibited the highest resolution with the strongest electron density to cover the bound ligand **131**. The diffraction limit of refined X-ray structures ranged from 1.97 Å to 2.32 Å and ligands were modelled with occupancies ranging from 0.41 to 1.0 (Table 31). Calculation of PanDDA maps enables to identify even less populated binders with occupancies as low as 20 %.[536] PanDDA analysis has proven to reveal low occupancy ligands that would not have been identified by conventional X-ray analysis and thus significantly improved the hit rate (for discussion, see Chapter 6.3).

Table 31: Crystallization strategy and X-ray results for identified compounds.

Compound	Binding site	Soaking/ Co-crystal	Crystal condition	Soaking concentration	Soaking duration	Resolution [Å]	occupancy
129	active	Soaking	A	50 mM	1 day	2.32	0.41
132	active	Soaking	B	5 mM	1 day	2.14	0.70
131	active	Co-crystal	C	50 mM	3 hours	2.19	0.73
138	active	Soaking	A	50 mM	3 hours	2.29	0.76
123	allosteric	Soaking	A	50 mM	1 day	2.11	0.68
130	S1	Soaking	A	50 mM	3 hours	1.97	1.00
136	S2	Soaking	D	25 mM	o/n	2.20	0.74

A	0.12 M caesium chloride, 12 % w/v PEG 3350, 12 % v/v DMSO Drop setup: 1:1:0.5 = cmpd. diluted in reservoir: cryo solution (15 % PEG 3350, 45 % glycerol): seed buffer
B	0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % v/v ethylene glycol Drop setup: 1:1:0.5 = cmpd. diluted in reservoir: protein buffer: seed buffer
C	0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % v/v DMSO Drop setup: 1:1:0.5 = cmpd. diluted in reservoir: protein buffer: seed buffer crystal transferred to 1:1 mix of 50 mM compound in reservoir and 30 % ethylene glycol + 12% PEG 3350
D	0.12 M caesium chloride, 12 % w/v PEG 3350, 6 % v/v xylitol Drop setup: 1:1:0.5 = cmpd. diluted in reservoir: protein buffer: seed buffer

T. brucei FPPS complex crystal structures were solved with fragments ranging from a molecular weight of 122 to 262 and clogP values from 1.26 to 3.18 and no significant molecular weight and clogP trend could be observed (Figure 49). Notably, four fragments (**129**, **131**, **132**, and **138**) were identified in the active site. All four fragments were enriched in amines and contain a scaffold with a nitrogen-containing aromatic ring and at least two hydrogen-bond acceptors. In contrast, the allosteric binder **123** was lacking amino groups and instead contained a rotatable methyl-sulfonyl moiety and two chlorine atoms at the phenyl ring. The largest fragment, fragment **136** with the highest lipophilicity, was found in a surface-exposed new site called S2 and a three-dimensional fragment **130** was found in an additional site S1 (Figure 49).



Nr.	MW	non-H atoms	clogP	HBD	HBA	ROTB
129	236.7	16	2.99	0	2	5
132	122.2	9	1.35	1	2	1
131	198.3	15	3.06	1	2	3
138	220.7	15	3.07	2	4	2
123	239.1	13	1.70	0	2	2
130	194.3	14	1.26	2	2	3
136	261.8	18	3.18	0	3	1
Range	122.2 - 261.8	9 - 18	1.26 - 3.18	0 - 2	2 - 4	1 - 5

Figure 49: Overview of *T. brucei* FPPS with ligands of NMR fragment screen in different binding sites. *T. brucei* FPPS protein backbone of crystal structure in complex with compound **132** is shown with fragments of the NMR fragment screen in **(a)** front view and **(b)** top view. The allosteric site (red), active site (blue), site S1 (green) and site S2 (black) are marked and labelled. Chemical structures of compounds grouped by binding site are shown below. The table contains each compound with the molecular weight (MW), number of non-hydrogen atoms (non-H atoms), clogP value, number of hydrogen-bond donors (HBD), number of hydrogen-bond acceptors (HBA) and number of rotatable bonds (ROTB). The last row shows the respective range of values.

5.2.5 Identifying non-bisphosphonate active-site binders – a new binding mode

Four of the solved X-ray complex structures have a fragment bound to the active site. Compounds **132**, **131** and **138** bind in the active site horizontal to the antiparallel helices α_D and α_F and the compounds point towards the dimer interface. Compound **132** binds via a 3.0 Å H-bond

and a 2.9 Å H-bond to Asp103. The pyridine ring is positioned on top of Tyr99 for further π-stacking interactions (Figure 50).

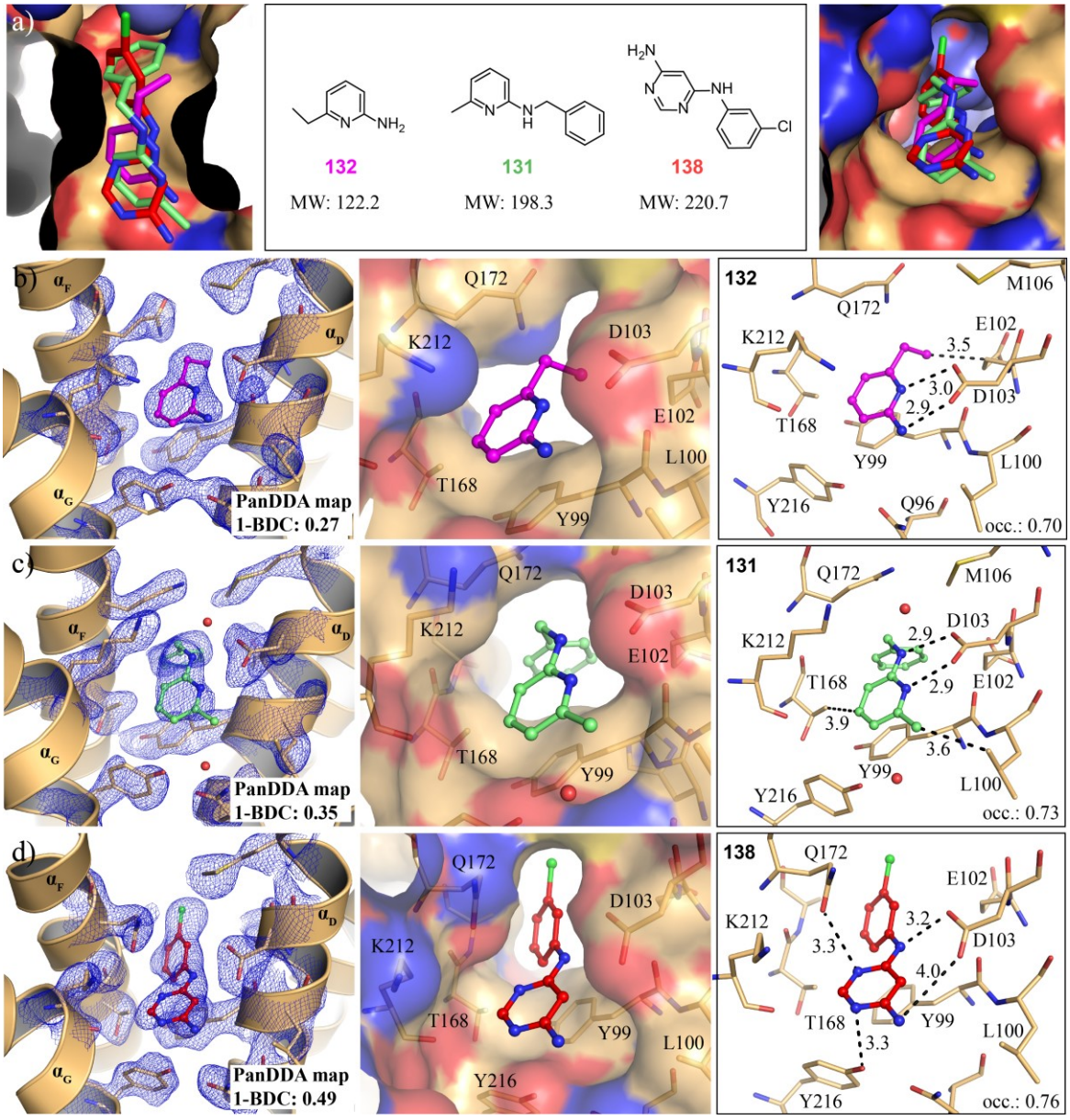


Figure 50: Three fragments in the active site identified in NMR fragment screen. (a) Surface representation of the *T. brucei* FPPS in complex with compound **138** (red) as stick representation. Compounds **132** (pink) and **131** (green) are superimposed and shown as stick representation. On the left: the side view and on the right the top view on the compound binding site is shown. In the centre: Chemical structures and molecular weight of compounds are noted. The *T. brucei* FPPS structure in complex with **(b)** compound **132**, **(c) 131** and **(d) 138** is shown. **(b)**, **(c)** and **(d)** On the left: the *T. brucei* FPPS is shown as Cartoon representation with PanDDA event map (blue, 2 σ) on compounds and residues in the binding site. The 1-BDC value of the PanDDA map is noted. In the centre: the surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of the *T. brucei* FPPS protein is presented. On the right, residues as sticks in the binding site are shown and the occupancy of modelled ligands is noted. Compounds are illustrated as ball and stick representation. Atomic distances are indicated by black dashed lines. Visual helices, residues and atomic distances are labelled.

Compound **131** is an extended analogue of **132** with an additional benzyl-ring attached to the amine and forms two H-bonds to Asp103 and π-stacking interactions with Tyr99. However, both

fragments bind with reversed topology. In contrast to **131** and **132**, compound **138** has a pyrimidine ring and establishes via the amino linking the chloro-substituted phenyl ring a 3.2 Å H-bond to Asp103. The primary amino group is 4.0 Å away and too far for a proper H-bond interactions. Apart from that, **138** forms two H-bond interactions to Tyr216 and Gln172 and shows π-stacking interactions with Tyr99. Its phenyl ring is positioned with a 45 ° rotation with respect to the pyrimidine ring and the meta-substituted chlorine atom points towards helix α_F of the homodimer. The chlorine atom is positioned in 3.5 Å distance to Met106 and interacts via 3.9 Å distance with an angle of 102.3 ° to the nitrogen at Asn131 on the dimer interface.

All three fragments show partial difference densities in the PanDDA maps at 2 σ that fully encompass the compounds and the side chains of the residues forming the binding site (Figure 50). Upon binding of compounds **131**, **132**, and **138**, the side chain of Lys212 is shifted and adopts a similar conformation in all three structures. However, in contrast to the protein-ligand structures with compounds **131** and **132**, in the binding mode of **138**, the side chain of Gln172 is rotated towards the dimer interface to make space for compound binding.

In contrast to **131**, **132**, and **138**, the fourth fragment **129** in the active site adopted a different binding mode oriented along helix α_D with an occupancy of 0.41. Figure 51 shows a comparison of the event map and the finally refined 2Fo-Fc and Fo-Fc map of the ligand. The PanDDA map showed clear partial-difference density for **129** with the strongest density around the chlorine atom, and an unambiguous identification of the binding mode was possible. The final 2Fo-Fc map is the result of both states. In the ground state, residues in the ligand binding site would clash with the ligand (Figure 52).

For additional mFo-DFc difference electron density maps, 2mFo-DFc electron density maps, refined 2Fo-Fc and Fo-Fc electron density maps, and the PanDDA event maps for compounds **132**, **131**, **138**, and **129**, see Appendix, Chapter 10.5.2, Figure 79.

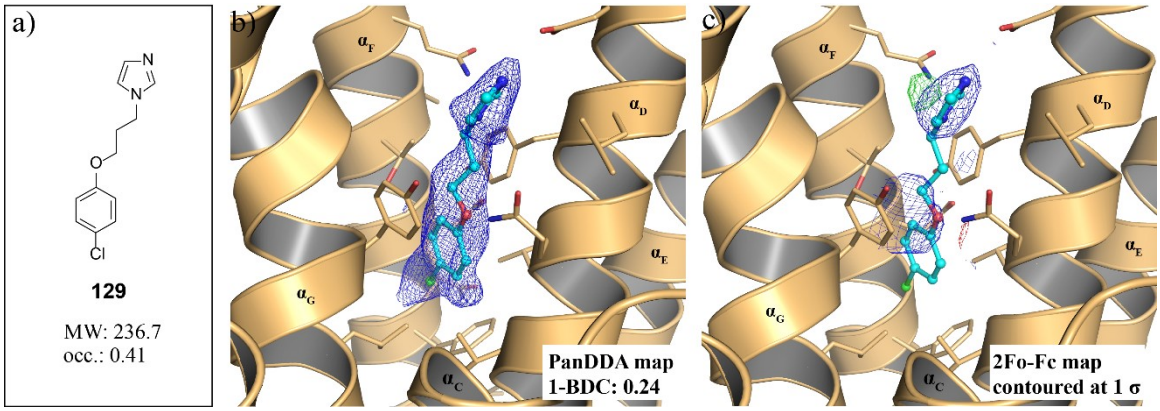


Figure 51: Comparison of PanDDA map and 2Fo-Fc map of *T. brucei* FPPS in complex with compound 129. **(a)** Chemical structure of compound **129** is depicted. **(b)** and **(c)** *T. brucei* FPPS-compound **129** complex structure is shown as Cartoon representation with residues around the ligand as stick representation. Compound **129** is shown as ball and stick representation in cyan. **(b)** PanDDA event map (blue, 2σ) of ligand (carve: 1.9) is shown. **(c)** Final 2Fo-Fc map contoured at 1 σ and Fo-Fc map contoured at 3 σ on ligand (carve: 2.5) is shown.

Compound **129** has a flexible four atom long linker between a para-substituted phenyl ring and an imidazole ring. The compound binds along helix α_D and upon binding, the side chains of Tyr216, Thr217 on helix α_D and Arg50 on α_C are shifted to open up a long elongated binding site (Figure 52a). The para-substituted chlorine atom is positioned surrounded by mostly hydrophobic residues Leu220, Phe161 and Val164. In a reported study on protein-ligand interactions in PDB entries, the mean distance for polar interactions of chlorine atom with oxygen was identified to be 2.9 Å and generally, interaction distances for halogen-bonds are described to be below the van der Waals radii of the two interacting partners which is 3.27 Å for a Cl···O interaction.[537-538] In the present complex structure, the closest Lewis base for a halogen-bond is the backbone oxygen of Phe161 in a distance of 3.4 Å and therefore too far for a halogen-bond. However, halogen-bonds are weaker than hydrogen-bonds and the position of **129** allows a 3.2 Å H-bond to Gln96 and the imidazole-nitrogen establishes a 2.8 Å H-bond to Asp103 (Figure 52).

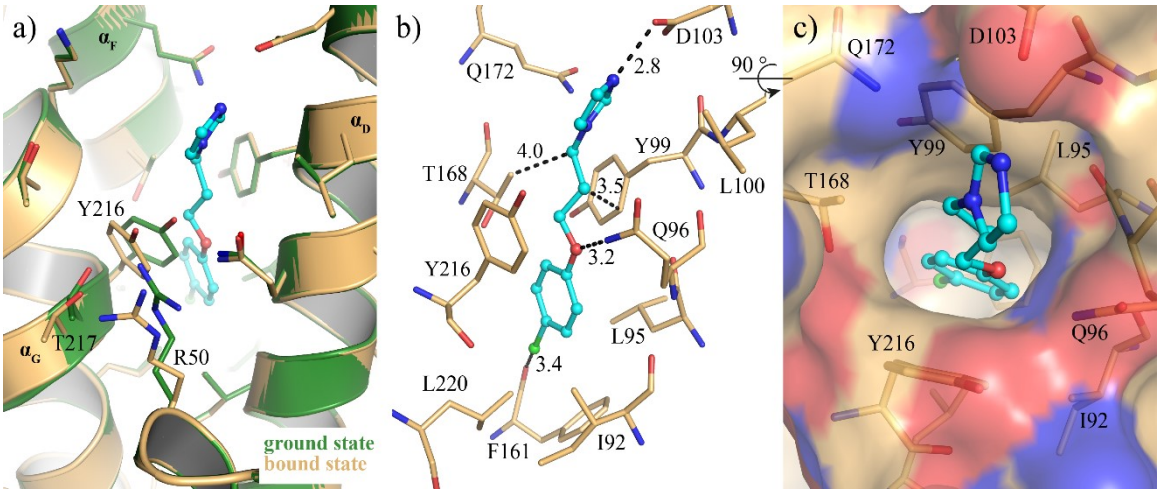


Figure 52: Binding mode of elongated active-site binder. The figure shows the structure of *T. brucei* FPPS in complex with compound **129**. Compound **129** is depicted as ball and stick representation (cyan). **(a)** Overlay of *T. brucei* FPPS ground state (green) and bound state (light orange) with residues in the binding site as stick representation. Residues that show significant conformational changes in ground and bound state are labelled. **(b)** Stick representation of residues in the binding site of the bound state. Distances between the protein and ligand are indicated by black dashed lines. Residues and atomic distances are labelled. **(c)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of ligand binding site in the bound state with visible residues labelled.

All four active-site binders (compounds **129**, **131-132**, and **138**, see Figure 49) superpose by their nitrogen-containing aromatic ring and these in turn are in the same location as the heterocycle ring of the bisphosphonates (Figure 53). Up to now, no structure has been reported with a ligand expanding along helix α_D. Instead, a structure of *T. brucei* FPPS in complex with bisphosphonate **25** (Figure 53) having an eight-membered aliphatic side chain was determined and the aliphatic chain points towards the dimer interface, similarly to compounds **132** and **138**.[301] For the reported structure in complex with **25**, the side chain orientation of Tyr99 on α_D was made responsible to prevent a binding mode along helix α_D. In the *T. brucei* FPPS structure in complex

with **129**, the Tyr216 on α_G which is opposite to Tyr99 is shifted and in that way opens up the narrow binding site along α_D.

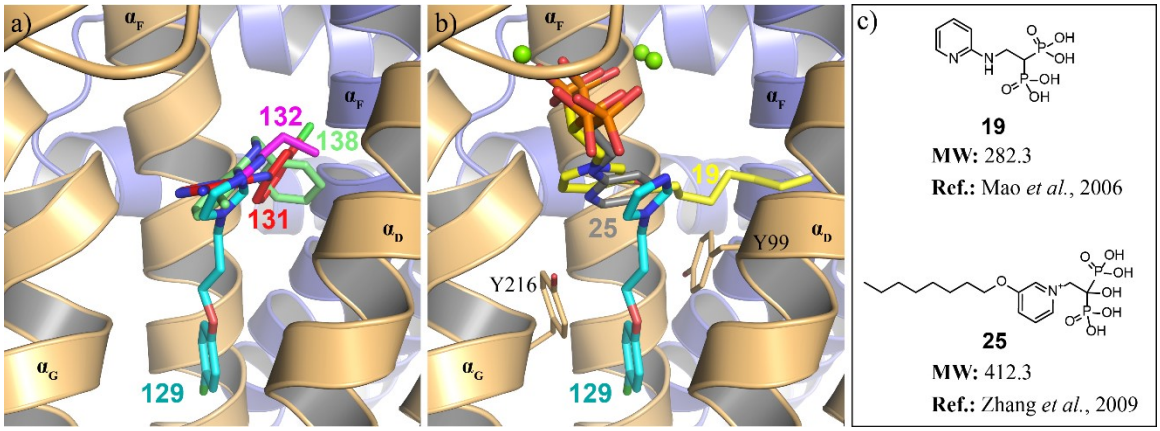


Figure 53: Overlay of *T. brucei* FPPS active-site binders and bisphosphonates. (a) and **(b)** Cartoon representation of the *T. brucei* FPPS in complex with **129** (cyan). **(a)** Compounds **131**-**132** and **138** are overlaid (protein backbone of respective structure is not shown). **(b)** Bisphosphonates **19** (PDB ID 2I19) and **25** (PDB ID 3EFQ) are overlaid (protein backbone of respective structure is not shown). Stick representation of Tyr216 and Tyr99 in the orientation found with bound **129**. **(c)** The chemical structure of bisphosphonates **19** and **25**. The molecular weight (MW) and reference is noted.

5.2.6 Methyl sulfonyl-fragment overlays with human allosteric binders in allosteric site

Out of the seven crystal structures in complex with fragments of the Novartis core fragment library, crystal structure analysis of compound **123** revealed binding to the allosteric site and the ligand was modelled with an occupancy of 0.68 (Figure 54). The PanDDA event map showed clear difference density for the dichloro-substituted benzyl moiety but the difference density for the methyl-sulfonyl moiety and the residues around the ligand appear blurry. For residue Leu358 on helix α_J no difference density is observed. This can be due to the high flexibility of the side chains around the ligand and to the high diffraction limit of 2.11 Å of the crystal structure.

Compound **123** forms two H-bonds to the terminal nitrogens of Arg350 and Gln49. Additionally, for **123** a short distance of 3.1 Å is indicated that might speak for a halogen-bond to the backbone oxygen of Thr217. However, due to the limited accuracy of the structural data and the reduced occupancy of **123**, interpretation of the density is difficult. The second chlorine atom is too distant to the protein backbone for halogen-bond interaction. An overlay with human allosteric binders and analogues **113**, **116**, and **121** shows that **123** binds in the same location. One chlorine atom attached to the benzyl moiety exactly overlays with the chlorine atoms of **113** and **116**. However, the benzyl moiety is rotated by 45 ° compared to the indole moiety of **113** and **116** and reaches deeper into the binding site between helix α_C and α_J (Figure 54).

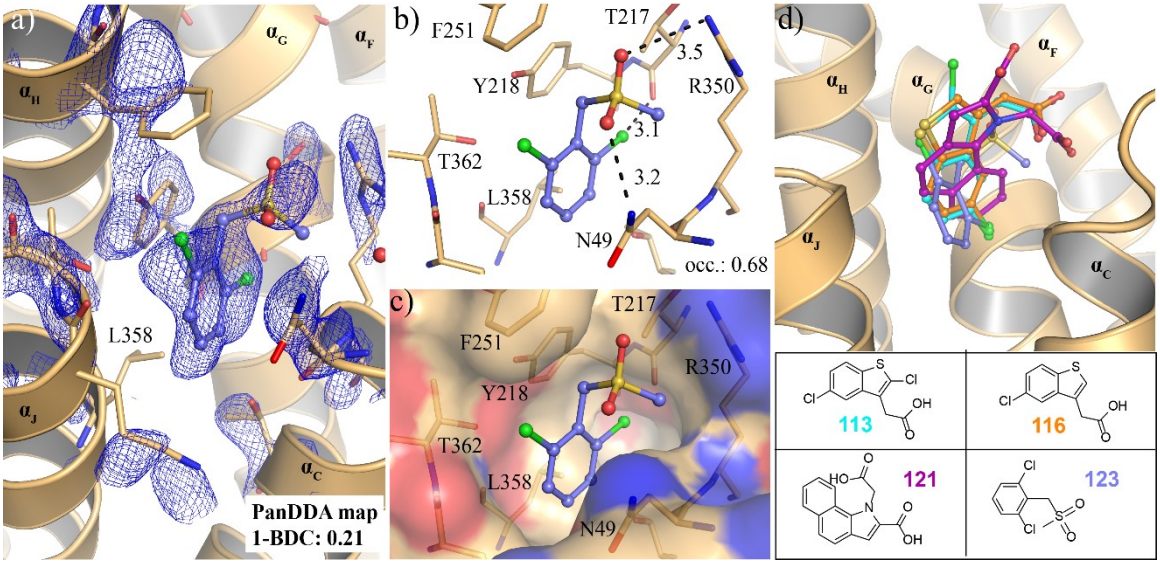


Figure 54: Allosteric binding of methyl sulfonyl-fragment. Complex structure of *T. brucei* FPPS with compound **123** (ball and stick representation, light blue) is shown. **(a)** Cartoon representation with PanDDA map around the ligand and residues in the binding site are visualized at 2 σ. The 1-BDC value of the PanDDA map is noted. **(b)** Residues as stick representation around the ligand are shown and the occupancy of compound **123** is noted. **(c)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of the ligand binding site with visible residues are labelled. For **(b)** and **(c)** distances between the protein and ligand are indicated by black dashed lines. Residues and atomic distances are labelled. **(d)** Cartoon representation of compound **123** overlaid with human allosteric binders **113** (cyan), **116** (orange) and **121** (violet).

5.2.7 Surface-exposed binding of NMR fragment hits

Two fragments of the seven NMR-detected fragment screening hits were identified in two different surface-exposed binding sites called S1 and S2 in this thesis (see Chapter 5.2.4, Figure 49). Compound **130** has an adamantyl substituent attached to a flexible four atom chain with a terminal amino group. The compound binds on top of helices α_B and adjacent to helices α_D and α_E in the surface-exposed site S1 (Figure 55 a and b). For **130**, a 136 µM K_d was estimated by NMR and a ligand efficiency of 0.37 kcal/mol was calculated. The loop between α_D and α_E is disordered and residues were not modelled into the density. However, residues of the α_D-α_E loop could further show interactions with **130** and contribute to the high binding affinity of **130** (Figure 55 a and b). For the additional density maps of the compound, see Appendix, Chapter 10.5, Figure 79.

Compound **136** consists of a chloro-isoquinoline moiety with an attached methyl piperazine ring and binds on the surface on top of helices α_A and α_B in S2 (Figure 55 c and d). Binding of **136** was also identified in protein-observed NMR and a K_d of around 1.4 mM was estimated. Although **130** has modest binding affinity, the binding sites of **130** and **136** are not of interest for a fragment follow-up campaign. The focus of this work was on fragments binding to the allosteric and active site as described in the following chapters.

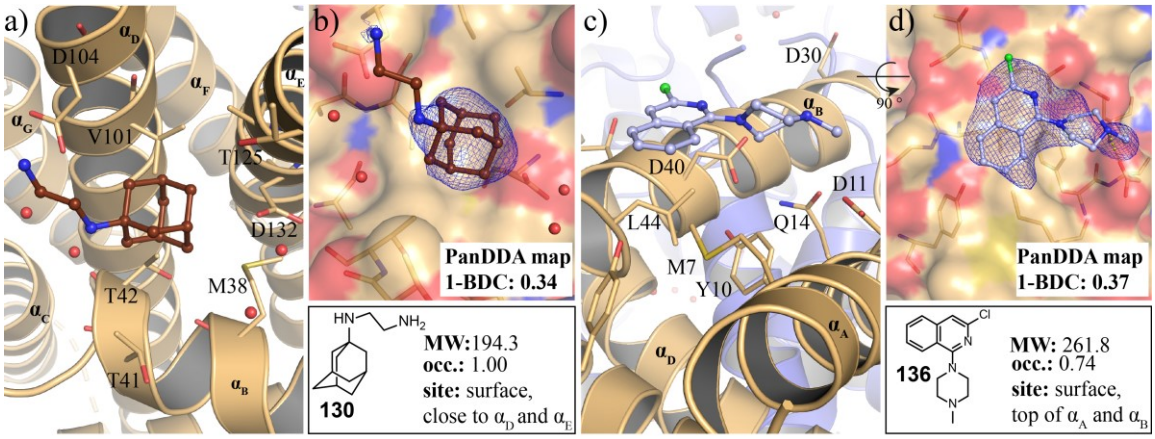


Figure 55: Surface-exposed binding of compounds 130 and 136. (a) and **(c)** Cartoon representation of *T. brucei* FPPS. **(b)** and **(d)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of *T. brucei* FPPS. Protein structure in complex with **(a)** and **(b)** compound **130** (ball and stick representation, brown) and **(c)** and **(d)** compound **136** (ball and stick representation, light blue) is depicted. In **(b)** and **(d)** the PanDDA difference map at 2 σ of the compounds is shown. Residues and helices are labelled.

5.2.8 Site hopping of fragment analogue from active to allosteric site

Compound **129** was identified in an NMR fragment screen and X-ray follow-up provided evidence for the binding pose in the active site. A search for fragment analogues of **129** retrieved four fragments from the Novartis archive. Fragments were tested in protein-observed NMR and were categorised as very weak, weak, medium and strong shifters.

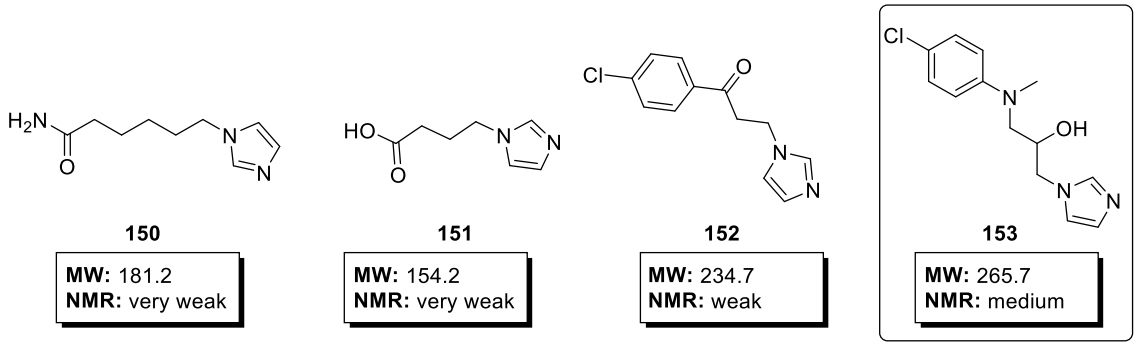


Figure 56: Fragment analogues of flexible active-site binder 129. Chemical structures, molecular weight and classification of compounds in very weak, weak, medium and strong shifter according to protein-observed NMR are shown. NMR results are obtained from [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiments with 30 µm *T. brucei* FPPS and 1 mM compound measured at 310 K.

Of the four tested analogues, only compound **153** was classified as a medium shifter. Compounds **150**, **151** and **152** only showed weak chemical shift perturbations upon compound addition and were classified as very weak or weak shifters by NMR. In order to determine the binding location of the compounds, crystal soaking experiments were performed. *T. brucei* FPPS crystals dissolved immediately upon the addition of 50 mM compound **152**. However, diffraction data sets were collected of crystals soaked with imidazole hexylamide **150** and the corresponding

butanoic acid **151** at 50 mM compound concentration for 21 hours or 3 hours and 30 minutes. Unfortunately, no additional electron density for ligand fitting could be identified in the electron density map. In contrast, an crystal structure of *T. brucei* FPPS in complex with **153** was solved from crystal soaking experiments at 50 mM compound concentration for one day. Ligand binding was detected using PanDDA analysis as part of the PanDDA run described in Chapter 5.2.4. Surprisingly, compound **153** was found to occupy the allosteric site (Figure 57). For the additional mFo-DFc difference electron density map, the 2mFo-DFc electron density map, the finally refined 2Fo-Fc and Fo-Fc electron density maps, and the PanDDA event maps, see Appendix, Chapter 10.5, Figure 80.

In contrast to the compound in the active site, the analogue had a less flexible linker of only four atoms with a hydroxyl group linking the aniline moiety and the imidazole ring. The compound forms a 3.0 Å H-bond to the hydroxyl group of Tyr218 although this interaction is weak due to an angle of 122 ° between the O-H···O far from the ideal angle of linearity (Figure 57). The nitrogen atoms of the imidazole ring and the chlorine atoms are positioned too far from the protein backbone to form key interactions. However, **153** forms hydrophobic interactions to Phe251, Thr217, Thr53 and Tyr218. The limited interactions of the ligand to the protein are also represented in the low ligand occupancy of 0.54.

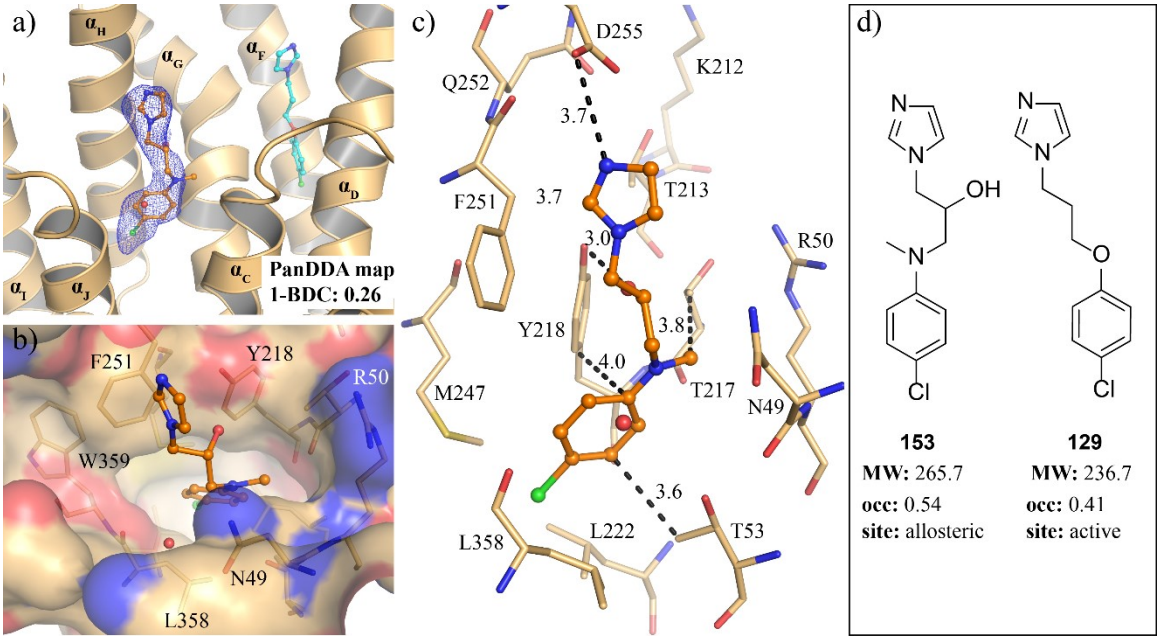


Figure 57: Fragment analogue of an active-site binder in the allosteric site. (a) Cartoon representation **(b)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) and **(c)** Stick representation of *T. brucei* FPPS (light orange) in complex with compound **153** (ball and stick representation, orange) is shown. **(a)** PanDDA differential density map at 2 σ indicates the bound compound. Compound **129** from the *T. brucei* FPPS structure in complex with **129** (stick and ball representation, cyan) is depicted in the active site. Helices are labelled. **(b)** Top view on allosteric site. **(c)** Side view on allosteric site. Atomic distances are indicated by black dashed lines and atomic distances, residues and helices are labelled.

In a second round of fragment analogue search, SAR by catalogue and archive considering compounds **129** and **153** was performed, and ten additional analogues were tested in protein-observed NMR (Figure 58). For all analogues, the imidazole-moiety was retained, but variations of the linker and different attachments on the benzyl moiety were taken into account. For most of the compounds more than one chemical property was varied compared to the parent compounds, which complicates the investigation which variation resulted in additional chemical shift perturbations in protein-observed NMR. SAR studies showed that additional methyl substitutions on the benzyl moiety as in **154** leads to protein aggregation. The addition of a methoxy group and a methyl group on the benzyl moiety as in compound **158** diminishes protein binding. Compounds **160** and **162** have an amide group in a respective three or four atom linker, which almost completely diminished protein binding. Similarly, compound **161** with an additional methyl substitution on the benzyl and imidazole ring only showed very weak shifts. Compound **163** was the only tested compound with a bicyclic system but again only weak chemical shift perturbations were detected in protein-observed NMR. Of the ten tested compounds, four compounds were classified as medium to strong shifters. Compounds **155**, **157** and **156** showed different substitutions on the phenyl ring, and compound **159** had a methyl and amide group in the linker (Figure 58).

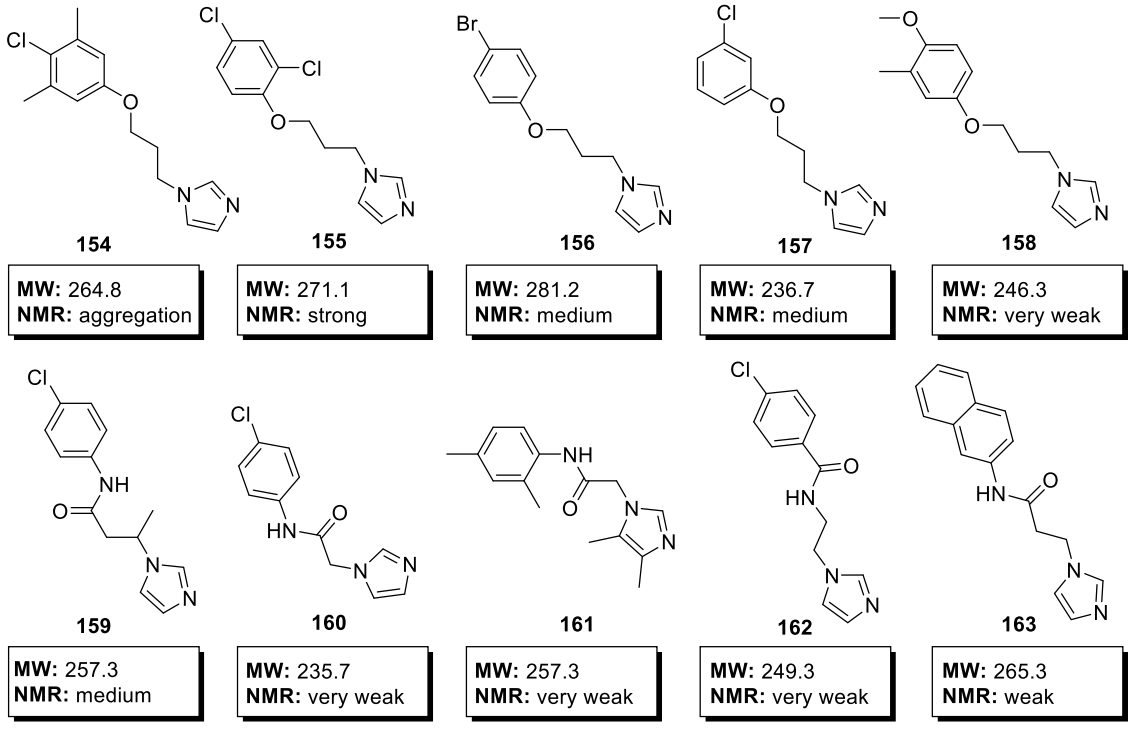


Figure 58: Fragment analogues of active-site binder 129 and allosteric site binder 153. Chemical structures, molecular weight and classification of compounds in very weak, weak, medium and strong shifter according to protein-observed NMR are shown. NMR results are obtained from [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiments with 30 µm *T. brucei* FPPS protein and 1 mM compound measured at 310 K.

Compound **155** with a dichloro substitution was the only compound classified as a strong shifter by protein-observed NMR and was subjected to X-ray experiments. Unfortunately, diffraction data collection of crystal soaks with 10 mM compound for six hours or one day were unsuccessful. At this stage, no additional follow-up was conducted. Instead a broad X-ray fragment screen by crystallography was performed to identify additional diverse fragment binders (Chapter 6).

5.3 Discussion

Generally for NMR fragment screening, a combination of different ligand-observed NMR experiments, such as saturation transfer difference (STD), wLOGSY and T2 filtered experiments, are recorded for cross-validation.[348, 381] STD experiments are probably the most commonly used ligand-observed NMR experiment.[348] However, experiences at Novartis have shown that the STD experiment is prone to the identification of false positive hits if small amounts of protein aggregates are present. These false positive fragment hits needs further follow-up experiments for confirmation. To avoid artificially increasing the hit rate and to avoid time consuming NMR experiments for validity check, the combination of wLOGSY and T1ρ experiments were the method of choice for this work. Both experiments can be recorded on the same sample and are of relatively short duration. Fragment hits identified by these two methods were combined and in total, 85 fragment hits were identified in mixture NMR experiments. The additional aspect of exploring differences and similarities of fragment hits between the *T. brucei* and *T. cruzi* FPPS was added at this stage by comparing identified hits with *T. cruzi* FPPS fragment hits of a parallel project performed by Joy Petrick, NIBR, Basel, Switzerland. At this early fragment screening stage, the combination of fragment hits on *T. cruzi* and *T. brucei* FPPS was also chosen as a pre-selection to reduce the number of hits since only fragment hits were retained that gave clear results for both *T. brucei* and *T. cruzi* FPPS.

The validation rate of *T. brucei* FPPS hits identified in the NMR mixture experiments to validated hits in the single compound NMR experiments was 53.6 % of fragments of the Novartis core fragment library and 55.9 % of fragments of the Novartis fluorine fragment library. Screening in mixtures accelerates the fragment screening process but at the same time it can result in mixture effects due to combined binding of two fragments, which further need to be excluded in follow-up experiments. For single compound follow-up NMR experiments, fragment testing of *T. cruzi* and *T. brucei* FPPS hits on human FPPS was included. Two compounds of the Novartis core fragment library were selective human FPPS binders in consequence these two hits were discarded for the trypanosomal FPPS targets. For a putative drug candidate, selectivity towards the trypanosomal FPPS target will be crucial in the treatment of parasitic infections in humans. To obtain first insights on these binding property, fragment hits were compared between the three FPPS targets. At this

early stage, selective fragments of the Novartis core fragment library towards *T. brucei* FPPS were not observed but selectivity can be introduced throughout the fragment optimisation process at later stage.

Unexpectedly, the hit rate for *T. cruzi* FPPS is significantly higher than for *T. brucei* FPPS and all fragments that bind to *T. brucei* FPPS also bind to either *T. cruzi* FPPS, human FPPS or both targets. The *T. cruzi* FPPS allosteric site is blocked by a phenylalanine as described in the ConSurf analysis in Chapter 4.2.1, which limits fragment binding in this site on first sight. As a limitation, ligand-observed NMR does not indicate the binding location and potentially *T. brucei* and *T. cruzi* FPPS have various different binding sites.

Protein-observed NMR was the first orthogonal method for hit validation and all 25 fragments of the Novartis core and additional 19 fluorine fragments of the Novartis fluorine fragment library were validated and classified as weak, medium and strong shifters. Interestingly, compounds with the highest T1ρ effect were not the compounds classified as strong shifters. However, the average T1ρ effect ranged between 20 – 30 % and was not significantly different. No resonance assignment of the amide backbone was available. Therefore, the actual binding site of the compounds could not be determined by protein-observed NMR. Although it was recently argued in the literature that structural models by NMR are a powerful alternative to X-ray crystallography, this method is not applicable to all proteins and requires time-consuming experiments.[466] In fact, only 10 % of fragment optimisation campaigns did not include a crystal structure at any stage.[340] In the current work, validated NMR *T. brucei* FPPS fragment binders were subjected to crystallographic experiments and the overall goal was to determine protein-ligand crystal structures of as many NMR validated fragments as possible to characterize their binding location. In X-ray crystallography experiments seven fragments at four different sites using PanDDA analysis were identified, which correspond to a 17 % validation rate of the combined 41 fragments that were tested in X-ray crystallography. Interestingly, only strong shifters of the Novartis fluorine fragment library were subjected to crystallographic experiments but no crystal structure of *T. brucei* FPPS with a fluorine fragment could be determined. Additional soaking experiments with higher compound concentrations and crystallization experiments applying other crystallization conditions may resulted in further protein-ligand structures but such conditions were not tested as part of this thesis.

In the described screening cascade, X-ray crystallography was employed as an orthogonal method after primary NMR fragment screening as a prefilter. However, a comparative study on endothiapepsin using six fragment screening methods in addition to X-ray crystallography found that remarkably 44 % of fragments solved in an X-ray structure would not have been identified in the other tested biophysical methods. This case study and also the setup of X-ray fragment screening platforms, such as XChem (Diamond, Grenoble) or the HTX lab (EMBL, Grenoble) support the notion of X-ray crystallography as a primary fragment screening technique. Therefore, a large scale X-ray fragment screen was performed on *T. brucei* FPPS and will be described in the next chapter.

As part of the X-ray fragment screen also PanDDA analysis used herein for the detection of the seven fragment binders in X-ray crystallography will be discussed. Notably, of the seven fragment binders identified in the X-ray follow-up of NMR fragment screening, four fragments occupied the active site.

In this work, the first fragment screen on *T. brucei* FPPS is described and for the first time non-bisphosphonate active-site binders on *T. brucei* FPPS were identified. Three fragments overlay with the backbone of bisphosphonates horizontal to helices α_D and α_F in the active site and the fourth fragment **129** occupies a long extended binding site along helix α_D (Figure 53). Over the course of the thesis, two *T. brucei* FPPS structures in complex with **129** were determined and compound binding was identified using PanDDA analysis. Compound **129** had a low occupancy of 0.41 and was a weak shifter in protein-observed NMR. The binding pose is unique compared to reported bisphosphonates. Even the reported bisphosphonate **25** with a molecular weight of 412.3 comprising an eight-membered aliphatic side chain bound horizontal to helices α_D and α_F and did not introduce conformational changes to accommodate compound **25** at the site along helix α_D (see Chapter 5.2.5, Figure 50).[301] SAR studies of **129** based on were performed and in a first fragment analogue screen of the described active-site binder, analogue **153** was identified, however, remarkably now at the allosteric site. This result distinctly shows the challenges of fragment-based drug discovery. On the one hand, fragments with a molecular weight of < 300 Da can probe otherwise unexplored binding sites but on the other hand, minor changes on the molecule can lead to pocket hopping and fragment reorientation. Thus, fragments do not necessarily maintain their binding mode during a fragment optimisation program.[539-540]

Apart from active-site binders, a methyl-sulfonyl fragment **123** was identified in the allosteric site and its binding pose overlaid with those of human allosteric binders on *T. brucei* FPPS as described in Chapter 4.2.4. Interestingly, more fragments were identified in the active site than the allosteric site although in a reported fragment screen on human FPPS no fragments in the active site were found.[310] Fragment-based screening is a powerful technique to identify new binding sites. In this work, two fragments, **130** and **136**, were identified in new surface-exposed binding sites S1 and S2 but these sites were not of interest for this work.

.

6. X-ray fragment screening
Understanding the crystal system

6.1 Introduction

X-ray crystallography has advanced remarkably in the last two decades thanks to astonishing developments in crystal setup, handling, beamline automation, X-ray data collection and data analysis and has reached the throughput necessary for the use as a primary fragment screening method.[372, 423] The method has the power to readily elucidate binding modes that can guide a structure-based fragment optimisation strategy.

In this work, fragment screening by X-ray crystallography was performed at the XChem facility (Diamond, Oxford) and at the HTX lab (EMBL/ESRF, Grenoble). In total, 1149 compounds were screened as single crystal soaks at a compound concentration up to 45 mM. A diverse set of fragment hits in various binding locations were identified and will be described in this chapter. Unfortunately, the X-ray fragment screen at Diamond was unsuccessful because the diffraction limit of X-ray data sets was not sufficient to identify binding fragments with confidence. In the X-ray screen at EMBL/ESRF Grenoble, seven binding fragments were identified and the respective X-ray structures were submitted to the PDB. The potential reasons for this difference in the outcome of the X-ray screens will be addressed as well as the structural diversity and properties of identified fragment hits.

6.2 Results

6.2.1 XChem fragment screen – crystals with low diffraction limit

An X-ray fragment screen was performed at the XChem facility at Diamond, Oxford. At Diamond, crystals were set up on Swissci 2-drop and Swissci 3-drop crystal plates. Long thin crystal needles were observed and crystal morphology was comparable to crystals at Novartis (Chapter 3.2.4). Crystals appeared within 24 hours and grew to their full size within 48 hours at a crystallization rate of 90 %. For an X-ray fragment screen, 709 crystals were harvested in total, of which 278 crystals were soaked with Edelris library fragments, 407 crystals were soaked with DSPL fragments and 24 were apo-protein for X-ray diffraction tests. Immediately before harvesting, glycerol was added as a cryoprotectant and data sets were collected at the I04-1 Diamond light source. Of the 709 crystals, 604 data sets could be collected. Processed data sets had a mean diffraction limit of 3.15 Å and 376 data sets had a diffraction limit >3.0 Å. After further processing in DIMPLE, PanDDA analysis was performed on 396 data sets excluding data sets with

R_{free} values above 0.4 after refinement. Eighty-three PanDDA events were identified, and the majority of these events was found in data sets with diffraction limits below 2.5 Å where PanDDA analysis showed events as a result of blurry electron densities around residues. In 14 diffraction data sets with 17 PanDDA events, fragment (compounds **164 – 177**, see Appendix, Chapter 10.2.5) fitting was attempted. However, ligands could not be fitted with high confidence as PanDDA event maps did not show sufficiently strong difference density to cover the full ligand and the overall diffraction limit was inadequate for ligand fitting with high confidence. Therefore, as compounds **164 – 167** were available in the Novartis archive, crystal soaking experiments of the putative XChem hits were reproduced at Novartis but these crystal soaking experiments did not reveal any additional structures with bound fragments.

In parallel, an X-ray fragment screen on *T. cruzi* FPPS was performed by Joy Petrick, NIBR, Basel, Switzerland on the XChem screening sample at Diamond. Protein-ligand complex crystal structures with a diffraction limit <2 Å were solved and structures showed bound fragments in various sites including the active and allosteric site. These compounds were also screened as part of the X-ray screen on *T. brucei* FPPS at Diamond but the diffraction limit of data sets was insufficient to identify binding events. For seven of the identified *T. cruzi* FPPS active site hits (compounds **178 – 184**), crystal soaking experiments were reproduced on *T. brucei* FPPS at Novartis without success.

In summary, the X-ray fragment screen at Diamond failed and also all attempts to reproduce crystal soaking experiments at Novartis did not result in any novel structure with a bound fragment. Nevertheless, the seven *T. cruzi* FPPS active-site binders **178 – 184,** which were identified in the X-ray fragment screen on *T. cruzi* FPPS were also tested on *T. brucei* FPPS by protein-observed NMR (Figure 59). Except for compound **164**, all compounds were indicated as very weak or weak shifters in protein-observed NMR (for a definition of the shifter classification, see Chapter 2.8.5).

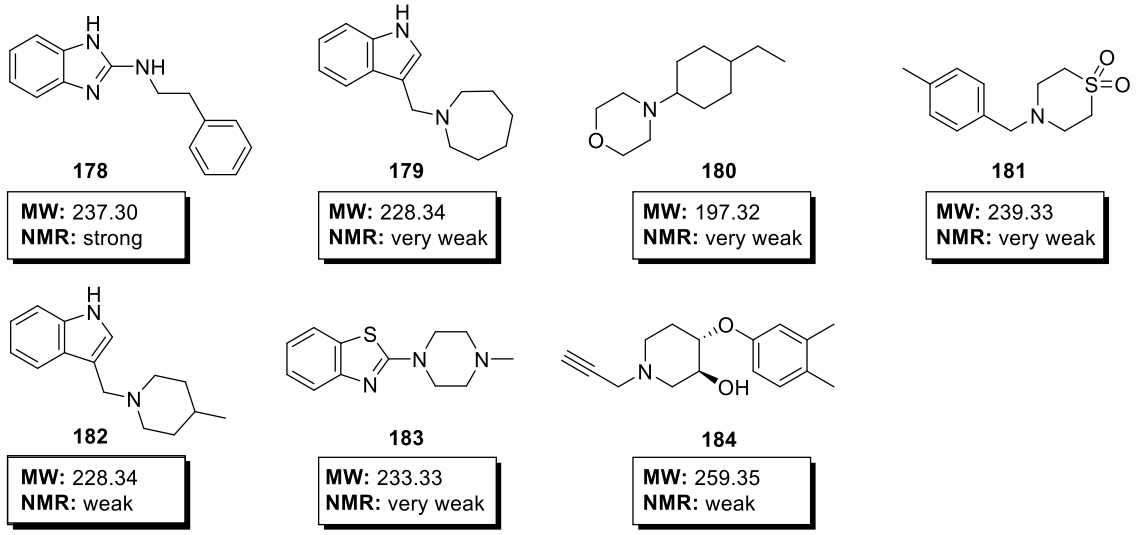


Figure 59: *T. cruzi* FPPS active-site binders on *T. brucei* FPPS in protein-observed NMR. Chemical structures, molecular weight and classification of compounds in very weak, weak, medium and strong binder according to protein-observed NMR are shown. NMR results are obtained from [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiments with 30 µm *T. brucei* FPPS protein and 1 mM compound measured at 310 K.

6.2.2 X-ray fragment screen at EMBL/ESRF using CrystalHarvester

The HTX lab offered an X-ray fragment screening campaign under the iNEXT framework (submitted proposal ID2848) using the CrystalDirect harvester.[459, 519] The CrystalDirect harvester replaces manual crystal loop harvesting, which potentially harms crystals and is a major improvement particularly for fragile crystal systems such as the needles studied in this thesis. Interestingly, during initial plate testing at EMBL/ESRF, a 90 % crystallization rate was obtained but first collected diffraction data sets had a resolution >3 Å, which was insufficient for X-ray fragment screening. In an onsite visit the setup at Novartis and Grenoble was compared and the harvesting step was different in two aspects: First, at Novartis crystals were harvested manually and at Grenoble the CrystalDirect harvester was used and second, no cryoprotectant was added to the crystal drops at EMBL/ESRF Grenoble. Crystal soaking drops at Novartis contained glycerol but using the automated CrystalDirect harvester the crystallization solution is removed before harvesting and no cryoprotectant is needed. However, in an onsite visit at EMBL/ESRF, different crystal setups were compared and it was found that crystals diffracted up to 1.63 Å when glycerol was added whereas diffraction limit was around 3 Å without glycerol (for Discussion, see Chapter 6.3). Hence, glycerol was added to the crystallization drop together with the compound solution for crystals harvested as part of the X-ray fragment screen.

In total, 506 diffraction data sets with 464 unique ligands were collected at beamline MASSIF-1, MASSIF-3 and ID30B-1 at ESRF Grenoble. Diffraction limit of data sets ranged from 1.63 Å to 3.37 Å with a mean diffraction limit of 2.12 Å. R_{free} values ranged from 0.25 to 0.60 and

the six data sets with R_{free} values above 0.4 after refinement were excluded (Figure 60). For cell volume, axis length and angles, see Appendix, Chapter 10.4.2.

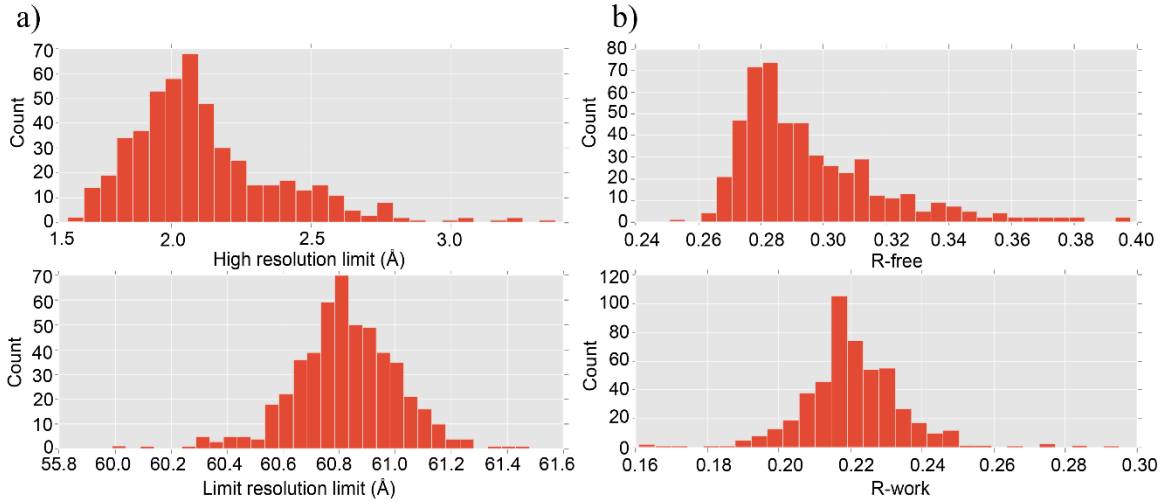


Figure 60: Resolution limit and R_{work}/R_{free} value distribution of analysed data sets. On the left: The bar diagram shows on the x-axis the high resolution limit in Å (top) and the low resolution limit in Å (bottom) and on the y-axis the count of respective diffraction data sets. On the right: The bar diagram shows on the x-axis the R_{free} value (top) and the R_{work} (bottom) and on the y-axis the count of respective diffraction data sets.

PanDDA analysis identified 69 PanDDA events distributed over 10 sites as shown in Figure 61. According to PanDDA inspections, 14 ligands were modelled at six different PanDDA sites and these data sets were exported for additional refinement in PHENIX. Refinement was performed using an ensemble model as described for the X-ray data analysis of Novartis structures in Chapter 5.2.4.

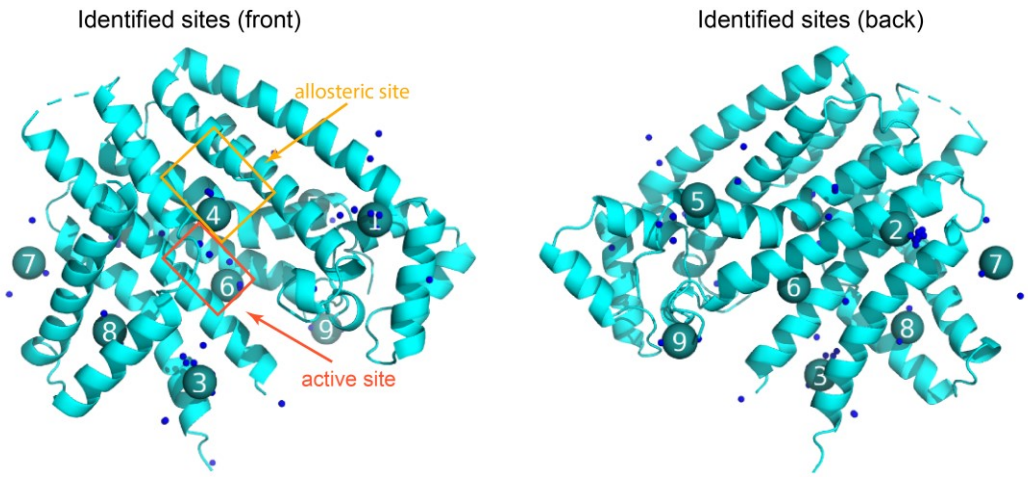
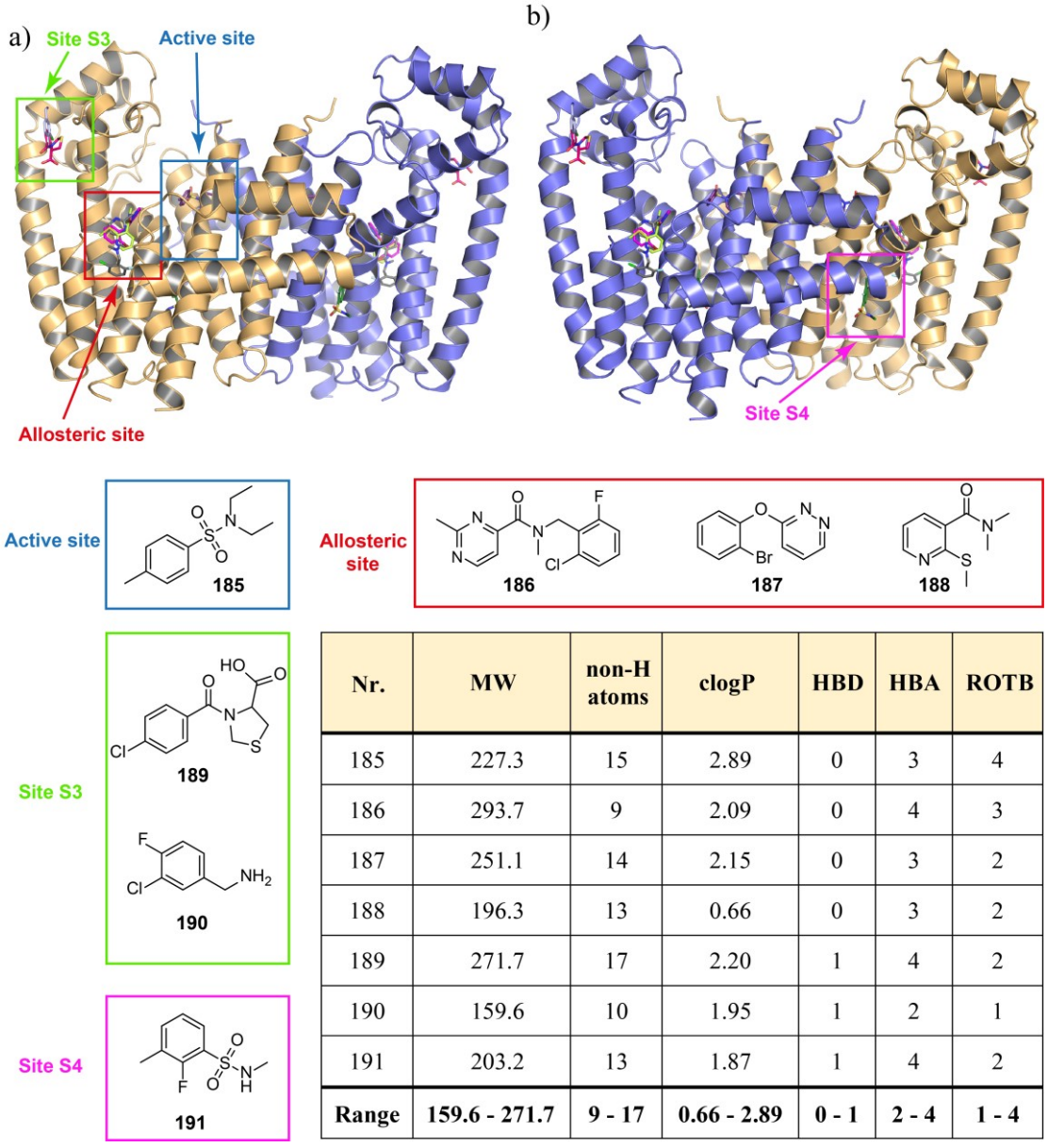


Figure 61: Identified PanDDA events of the X-ray fragment screen at EMBL/ESRF Grenoble. PanDDA sites with events are overlaid on *T. brucei* FPPS protein structure in Cartoon representation (front and back side). The allosteric site (yellow) and the active site (red) are marked and labelled. Each blue dot represents a PanDDA event of one X-ray data set. More than one PanDDA event can result from one diffraction data set if additional electron density is identified at different sites on *T. brucei* FPPS. Each labelled sphere corresponds to one PanDDA site which is a cluster of PanDDA events and results in 10 different PanDDA sites.

Of the initial 14 modelled ligands in the PanDDA inspection, seven ligands were excluded during further refinement if at least one of the following criteria applied to them: Ligands that were modelled with low confidence and could not be modelled unambiguously, and ligands with negative electron density after refinement. In summary, seven crystal structures were refined with three ligands in the allosteric site, one ligand in the active site, two ligands in a surface-exposed site between helices α_H and α_J, here called site S3, and one ligand between α_F and α_G on the opposite site of the allosteric site, here called site S4 (Figure 62).



Nr.	MW	non-H atoms	clogP	HBD	HBA	ROTB
185	227.3	15	2.89	0	3	4
186	293.7	9	2.09	0	4	3
187	251.1	14	2.15	0	3	2
188	196.3	13	0.66	0	3	2
189	271.7	17	2.20	1	4	2
190	159.6	10	1.95	1	2	1
191	203.2	13	1.87	1	4	2
Range	**159.6 - 271.7**	**9 - 17**	**0.66 - 2.89**	**0 - 1**	**2 - 4**	**1 - 4**

Figure 62: Overview of *T. brucei* FPPS with ligands discovered by X-ray screen in different binding sites. *T. brucei* FPPS protein backbone of crystal structure in complex with compound **185** is shown with fragments of the X-ray fragment screen in **(a)** front view and **(b)** back view. The active site (blue), allosteric site (red), site S3 (light green) and site S4 (pink) are marked and labelled. Chemical structures of compounds grouped by binding sites are shown below. The table contains each compound with the molecular weight (MW), number of non-hydrogen atoms (non-H atoms), clogP value, number of hydrogen-bond donors (HBD), number of hydrogen-bond acceptors (HBA) and number of rotatable bonds (ROTB). The last row shows the respective range of values.

Identified ligands ranged in molecular weight from 159.6 to 271.7 with a clogP value between 0.66 and 2.89. The most hydrophilic fragment is identified in the allosteric site. Notably, fragments in the active and allosteric site contain no hydrogen-bond donor but in contrast, a minimum of three hydrogen-bond acceptors. All three allosteric fragments have a nitrogen-containing aromatic scaffold. Of all the seven identified fragments, five contain at least one halogen atom. Fluorine, bromine and chlorine atoms are represented in the five different chemical structures. In the following, the binding of the different fragments will be analysed in detail.

Compound **185** occupies the active site with its diethyl group positioned between helices α_D and α_F pointing towards helix α_F at the dimer interface (Figure 63). The PanDDA map shows difference density covering the complete compound and strong difference density covering the residues in the binding site. For additional density maps, see Appendix, Chapter10.5.4, Figure 81.

The side chain of Gln172 on helix α_F is rotated by 3.6 Å and the side chain of Asp103 on helix α_D by 1.9 Å to open up a binding pocket for **185** (Figure 63 d). In the ground state *T. brucei* FPPS has a glycerol from the crystallization condition bound, which is replaced by the phenyl moiety of **185**. The methyl substituted phenyl group is positioned close to helix α_D with its methyl group pointing towards Gln96. The compound is positioned at a distance of 4.1 Å and 4.0 Å to the backbone nitrogen and side chain oxygen of Gln172.

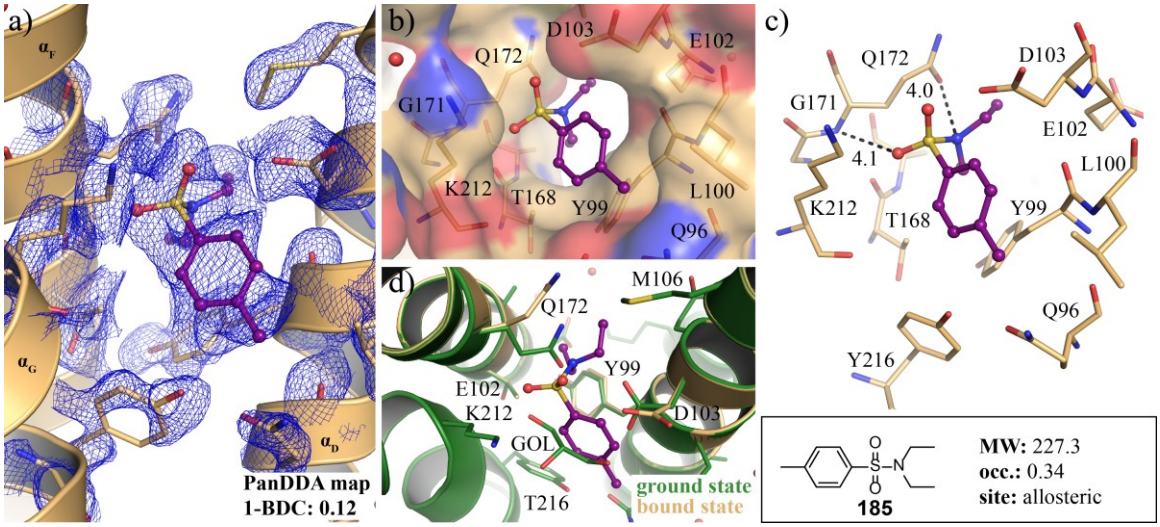


Figure 63: Active-site binder identified in the X-ray fragment screen. *T. brucei* FPPS protein structure in complex with compound **185** (ball and stick representation, violet) is depicted. **(a)** and **(d)** Cartoon and stick representation, **(b)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) and **(c)** only stick representation of compound **185** binding site in *T. brucei* FPPS. In **(a)** the PanDDA difference map of the compound and residues in the binding site are overlaid. Helices are labelled. In **(c)** residues are labelled and atomic distances are shown with a black dashed line. **(d)** Overlay of ground (green) and bound state (light orange) is shown and residues are labelled.

Apart from active-site binding, three crystal complex structures with compounds in the allosteric site were solved (Figure 64). Compound **186** has an amide linker between a methyl pyrimidine and ortho-substituted phenyl ring. It is buried in the allosteric site between α_C and α_J and the chloro-fluorine substituted phenyl ring is rotated towards helix α_G. The pyrimidine ring is

positioned in 3.8 Å distance to the hydroxyl group of Tyr218, which does not allow the formation of an H-bond. However, compound **186** shows hydrophobic interactions of the phenyl ring with Met247 and Thr53. The fluorine is not positioned to form any polar interactions and in summary, the compound's binding mode does not feature many strong protein-ligand interactions, which is also mirrored by its low occupancy of 0.37, which was only possible to be identified in the PanDDA difference density map using PanDDA analysis. Additional mFo-DFc difference electron density maps, 2mFo-DFc electron density maps, refined 2Fo-Fc and Fo-Fc electron density maps, for compound **186** and all following compounds identified in the fragment screen by X-ray crystallography at EMBL Grenoble can be found in the Appendix, Chapter 10.5.4, Figure 81.

Compound **187** also occupies the allosteric site. The pyridazine ring appears to be scattered over multiple arrangements and it only weakly contributes to difference density in the PanDDA map. Nevertheless, strongest difference density is found around the bromine atom to give confidence for binding of the compound and the location of the phenyl ring. In this binding mode, the pyridazine ring forms a 3.3 Å H-bond to Asn49 and the second nitrogen is positioned in a 3.7 Å distance to the same residue. The compound features main hydrophobic interactions with the side chains of Phe251 and Lys361 and the bromine atom is positioned at an angle of 147.1 ° and in 3.5 Å distance to the backbone oxygen of Thr271 and functions as a halogen-bond donor.

Compound **188** has a pyridine scaffold with a methyl-thioether substituent and an attached carboxamide group. It is bound to the allosteric site close to helices α_C and α_J. The carboxyl oxygen forms a 2.8 Å H-bond to one water molecule. This water molecule is positioned too far for any H-bonds to the protein backbone. The sulphate is placed in 3.7 Å distance to the oxygen of Asn49. Binding is dominated by hydrophobic interactions of the aromatic ring with Thr217, Leu358 and Asn49. Additionally, one methyl-group of the tertiary amine is at a distance of 3.3 Å to the aromatic ring of Phe251 for further hydrophobic interactions.

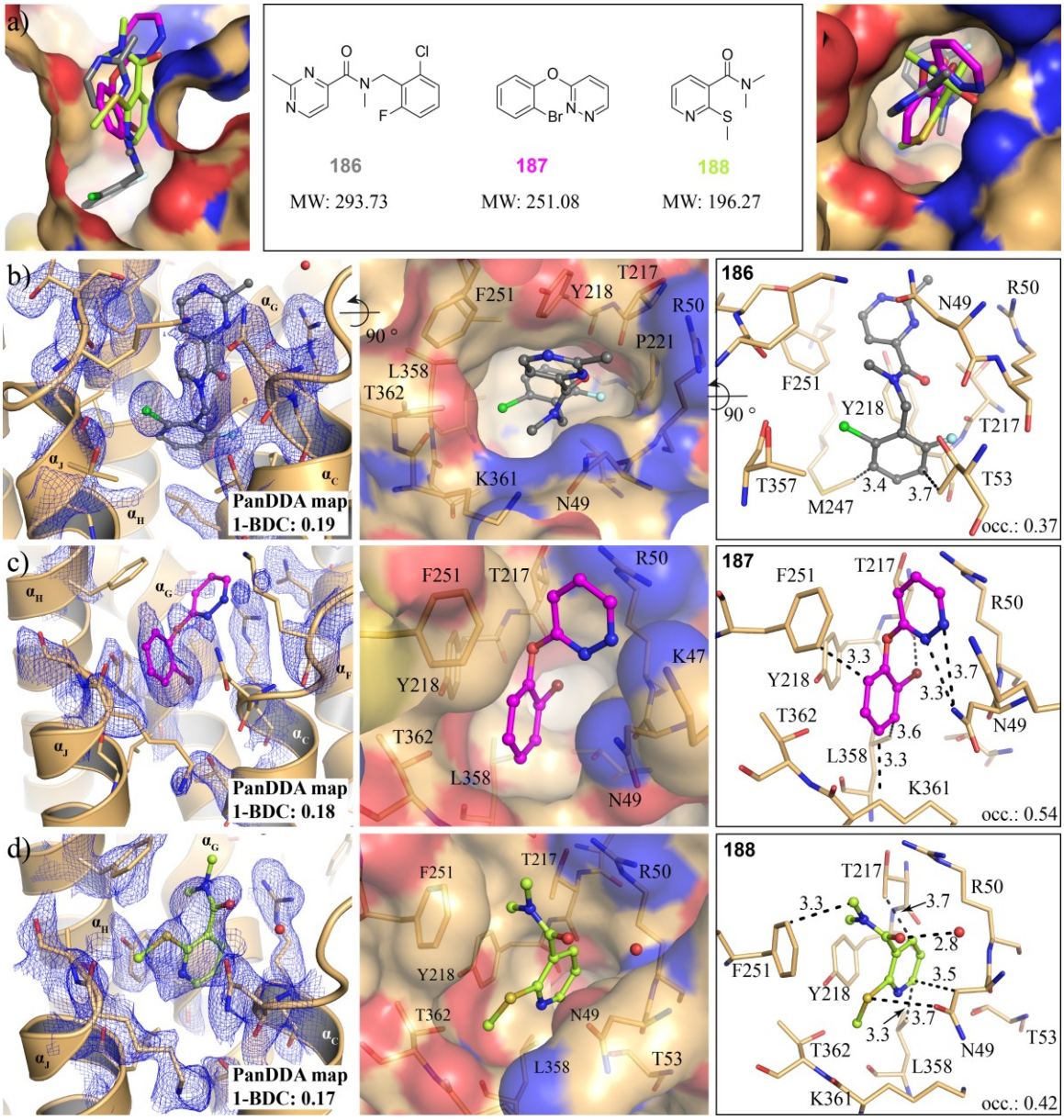


Figure 64: Three fragments in the allosteric site identified in X-ray fragment screen. (a) Surface representation of *T. brucei* FPPS in complex with compound **186** (grey) in stick representation. Compounds **187** (pink) and **188** (light green) are overlaid and shown as sticks. On the left: the side view and on the right the top view on the compound binding site is shown. In the centre chemical structures and molecular weight of compounds is depicted. *T. brucei* FPPS structure in complex with **(b)** compound **186**, **(c) 187** and **(d) 188** are shown. **(b)**, **(c)** and **(d)** On the left: *T. brucei* FPPS is shown in Cartoon representation with PanDDA map around the compounds and residues in the binding site. The 1-BDC value of the PanDDA map is noted. In the centre, *T. brucei* FPPS protein as surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) is shown. On the right, residues as sticks around the compounds are shown and occupancy of modelled ligand is noted. Compounds are illustrated as stick and ball representation. Atomic distances are indicated by black dashed lines. Helices, residues and atomic distances are labelled.

Two compounds, compound **189** and **190**, were identified at a new binding site, called S3, between helices α_H (residues Gln254 - Phe261) and α_I (residues Glu318 – Tyr326) remote from the active and allosteric site (Figure 65). There is no binding site on the unliganded *T. brucei* FPPS and a cryptic pocket appears upon ligand binding. The side chains of Gln323 and Gln318 rotate upon binding of **189** and the chloro-benzoyl moiety of compound **189** reaches between helix α_I and α_H.

It forms π-stacking interactions with Phe261 on helix α_H and hydrophobic interactions with Gln323 on helix α_I. In the bound state, the residue Glu318 on helix α_I is positioned with its carboxyl group on top of the chlorine atom of compound **189** and the thiazolidine moiety is solvent-exposed pointing towards the surface of *T. brucei* FPPS (Figure 65).

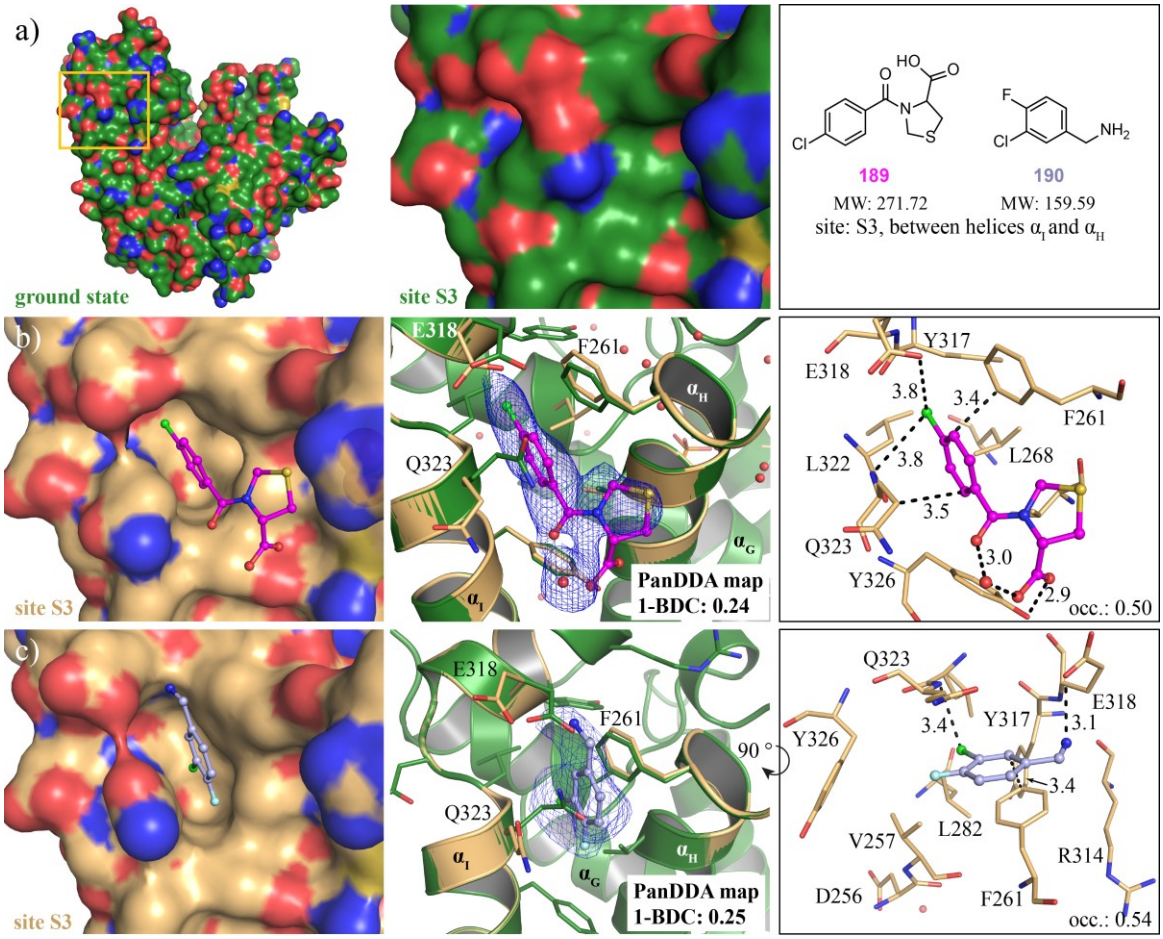


Figure 65: Fragment binding in Site S3. (a) On the left: surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of *T. brucei* FPPS in ground state (green), in the centre: close-up on Site S3 and on the right: chemical structure and molecular weight of compounds **189** and **190** and description of Site S3 location. **(b)** and **(c)** *T. brucei* FPPS in complex with **(b)** compound **189** (pink) and **(c)** compound **190** (light blue). Compounds are illustrated in stick and ball representation. On the left: surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of *T. brucei* FPPS binding site is shown. In the centre: overlay of ground (green) and bound (yellow) state of *T. brucei* FPPS in Cartoon representation and residues of the binding site are shown in sticks. PanDDA map at 2 σ of compounds is shown and the 1-BDC value of the map is noted. On the right: residues of the binding site are shown in sticks and occupancy of modelled ligand is noted. Atomic distances are indicated by black dashed lines. Helices, residues and atomic distances are labelled.

However, with its carboxyl group, **189** forms an H-bond to Tyr326 and an H-bond to a water molecule. Additionally, the carbonyl oxygen forms another 3.0 Å H-bond to the same water molecule. Compound **190** binds in the same binding location and the side chains of Gln323 and Glu318 are shifted by 3.5 Å and 2.8 Å, respectively, to open up the binding site. The compound is positioned along the α_I and α_H helices and forms π-stacking with Phe261 and a 3.1 Å H-bond to Glu318. The chlorine substitution is positioned at a distance of 3.4 Å to the backbone nitrogen of

Gln323 at an angle of 117 ° (Figure 65), which is not an ideal location head-on of the σ-hole on the chlorine atom to form a halogen-bond. Overall, both compounds have a low occupancy and extensive structure-based drug design would be needed to optimise their binding interactions. Additionally, the biological relevance of the site needs to be investigated, which is out of the scope of this thesis.

Compound **191** is buried between helices α_G and α_F on the opposite side of the allosteric site in a new binding location, here called S4 (see Figure 62). The sulphonamide group of **191** forms a 3.0 Å H-bond to Arg160 and further shows hydrophobic interactions with its methyl group to Val164, and its phenyl group to Leu219 and Leu220. The relevance of this newly identified binding site for additional fragment-to-lead optimisation was not further analysed as part of this thesis. In the following, only compounds of the active and previously identified allosteric site were further investigated by NMR.

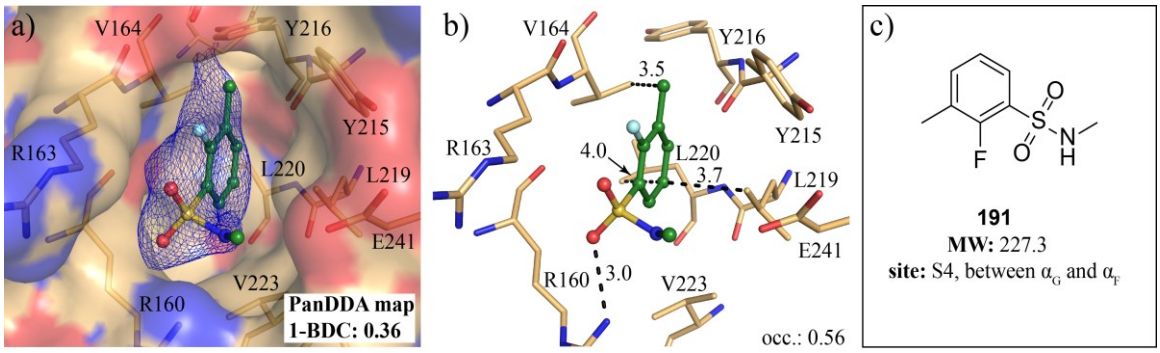


Figure 66: Sulphonamide fragment in new binding site S4. (a) Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of *T. brucei* FPPS in complex with compound **191** (ball and stick representation, dark green). The PanDDA difference map at 2 σ of the compound is shown. **(b)** Stick representation of residues in the binding site and occupancy of modelled ligand is noted. **(c)** Chemical structure of compound **191** is shown. Molecular weight (MW) and binding site is noted.

6.2.3 NMR-follow up of X-ray fragment screen

Allosteric and active-site binders identified in the X-ray fragment screen at EMBL/ESRF Grenoble were validated by protein-observed NMR and categorised as weak, medium and strong shifters. Protein-observed NMR experiments were performed as described in Chapter 2.8.5. The follow-up by protein-observed NMR was used to validate identified fragment binders by an orthogonal method and although crystal soaking experiments were carried out at compound concentrations up to 45 mM, chemical shift perturbation were observed for three fragments tested at 1 mM compound concentration (Figure 67).

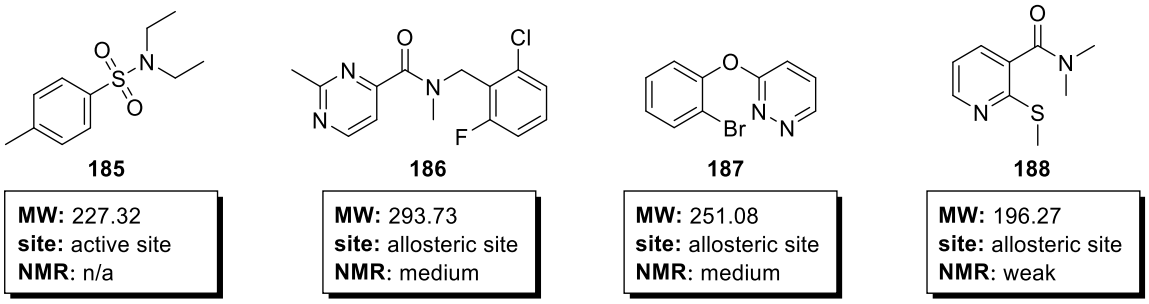


Figure 67: Protein-observed NMR of allosteric and active-site binders from X-ray fragment screen. Chemical structures of compounds **185** – **188** are shown and the molecular weight (MW), binding site and classification in weak, medium or strong shifter as determined by [^{13}C, ^{1}H]-SOFAST-HMQC is noted. Results of [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) with 30 µm *T. brucei* FPPS protein measured at 310 K are shown.

6.3 Discussion

The XChem facility (Diamond, UK) offers a fast and efficient workflow for high throughput crystal soaking experiments utilizing TeXRank[461] for crystal ranking, a Crystal Shifter for crystal harvesting, XChemExplorer[462] for data processing and PanDDA[463-464] analysis to detect weak fragment binders. In contrast, the fragment screening workflow at the HTX lab (EMBL/ESRF, Grenoble, FR) is built around the automatic CrystalDirect harvester to replace manual crystal harvesting in crystal loops.[459, 519]

An X-ray fragment screen on *T. brucei* FPPS was performed at both facilities to identify fragment binders. The XChem fragment screen at Diamond was carried out first, but strikingly, the screen failed most likely due to insufficient diffraction quality of data sets mostly of minor resolution. Similarly, initial data sets collected for the X-ray fragment screen at EMBL/ESRF showed diffraction limits >3 Å. Generally, glycerol is added to crystal drops as a cryoprotectant if no component such as glycerol, ethylene glycol, low PEG is part of the crystallization buffer but crystals were harvested with the automatic CrystalHarvester at EMBL/ESRF, the addition of a cryoprotectant was considered as obsolete. However, after extensive investigation during an onsite visit at EMBL/ESRF to understand the discrepancy between the X-ray diffraction limit observed at Novartis and at EMBL/ESRF, it was concluded that for the presented crystal system, glycerol has a second function as an additive and needs to be added to the crystal buffer. Therefore, new crystal plates were setup and initial data sets as part of the X-ray fragment screen at EMBL/ESRF were re-collected from crystals with glycerol added to the crystal drop. After this finding, 506 crystals with a mean diffraction limit of 2.12 Å were collected as part of the X-ray fragment screen at EMBL/ESRF. Therefore, it is also expected that the X-ray diffraction limit of data sets collected as part of the XChem fragment screen would have been improved if glycerol was added together with the compound or for longer periods before harvesting. However, the XChem fragment screen was not repeated as part of this thesis.

The X-ray fragment screen at EMBL/ESRF comprised data sets with 464 unique ligands and resulted in seven protein-ligand structures with fragments at four different sites, which corresponds to a 1.5 % hit rate. Three fragments were detected in the allosteric site and only one fragment in the active site whereas in the NMR fragment screen, four fragments were found in the active site and only one fragment in the allosteric site (see Chapter 5.2.4). All fragments except for one fragment in the active site exhibited a nitrogen-containing aromatic ring. Interestingly, for protein-ligand structures solved in an X-ray follow-up of the NMR fragment screen and as part of the X-ray fragment screen at EMBL/ESRF, the mean molecular weight of fragments in the active site was 201 in contrast to the mean molecular weight of 245 for fragments in the allosteric site. Hence, the discovered fragments in the allosteric site already exhibited a more complex structure and these fragments also contained carbonyl groups and more frequently halogen substitutions. Nevertheless, only a small number of fragments were identified which results in limitations of statistical analysis.

Notably, the mean occupancy of the seven protein-ligand structures solved with fragments identified in the NMR-fragment screen is 0.71, which is in contrast to the mean occupancy of 0.46 for ligands identified as part of the X-ray fragment screen at EMBL/ESRF. The crystal soaking protocol of the X-ray fragment screen is significantly different to the crystal soaking protocol performed at Novartis for X-ray follow-up of identified fragment hits by NMR. Crystals at EMBL/ESRF were soaked at lower concentrations for shorter periods, which could explain the lower occupancy of compounds identified in the X-ray fragment screen. However, due to the highly deviating soaking protocols and the screening of two different libraries, an interpretation on the different hit rates in the binding sites would not be well-founded.

In both screens, PanDDA analysis of crystal structures had a major impact on the hit rate and enabled the identification of even weakly bound fragments with occupancies as low as 0.34. This approach uses the high number of data sets collected without a bound ligand to generate a statistical model, so called ground state model that describes the different protein conformations detected in crystals as accurately as possible. Thus, it allows to detect differences even only occurring in 20 % of the unit cells. These deviations would be considered as crystallographic artefacts in conventional analyses and such low occupancy ligands would not have been detected without PanDDA analysis.[463, 536] However, these ligands contain valuable information for fragment follow-up and can help to guide the optimisation process in order to increase binding affinities of initial fragment hits, as it is described in the next chapter (Chapter 7).

7. Structure based fragment merging and growing by medicinal chemistry

7.1 Introduction

Biophysical fragment screening is an indispensable tool to identify fragments with affinities up to the double digit millimolar range. Resulting initial fragment hits need then to be optimised in a tailored fragment optimisation program. Ideally, in a first approach fragment binding is further investigated following SAR by catalogue or archive and structural data are obtained to analyse the binding mode and identify protein-ligand interactions.[348] Following this, structure-based fragment optimisation by medicinal chemistry is performed using one of the three major fragment evolution strategies: fragment linking[348, 483-485], merging[363, 486-487] and growing[363]. In this work, active site fragments identified by an NMR-fragment screen and followed-up by X-ray were the starting point for structure-based fragment evolution using both, the fragment merging and growing approach. In addition synergies from the *T. cruzi* and *T. brucei* FPPS fragment screens were utilized by superimposition and *T. brucei* and *T.cruzi* FPPS active-site binders overlaid with potential for fragment merging. In total, ten fragments were synthesised and tested by protein-observed NMR. Synthesised fragments were subjected to X-ray crystallography and one *T. brucei* FPPS structure in complex with a merged fragment will be presented.

7.2 Results

7.2.1 Fragment merging of two active-site binders

Compound **129** and **138** were identified in an NMR fragment screen and X-ray crystallography revealed an orthogonal binding pose in the active site. Compound **129** occupies the active site close to Asp103 horizontal to helices α_D, whereas **138** binds along helix α_D and points towards Phe161 (Figure 68). Upon binding of the two compounds, different conformational changes in *T. brucei* FPPS are observed. For **129**, the aromatic side chains of Tyr216 and Tyr99 shift to open up a binding site along helix α_D. For **138**, the side chains of Gln172 and Asp103 rotate out of space and give room to accommodate the chloro-substituted phenyl ring. Compound **129** could not bind to *T. brucei* FPPS in the conformation that the protein adopts upon binding of compound **138** and vice versa. To make binding of the merged compound feasible, the side chains as observed in both protein-fragment complex structures have to shift to open up a binding site that can fit the merged ligand resulting from the two orthogonal fragment binders. The imidazole ring

of **129** overlays with the pyrimidine ring of **138** indicating a potential strategy for fragment merging. Accordingly, a merged ligand of these two fragments was synthesised. The overlaying scaffold was inspired by **129** and the imidazole ring was retained.

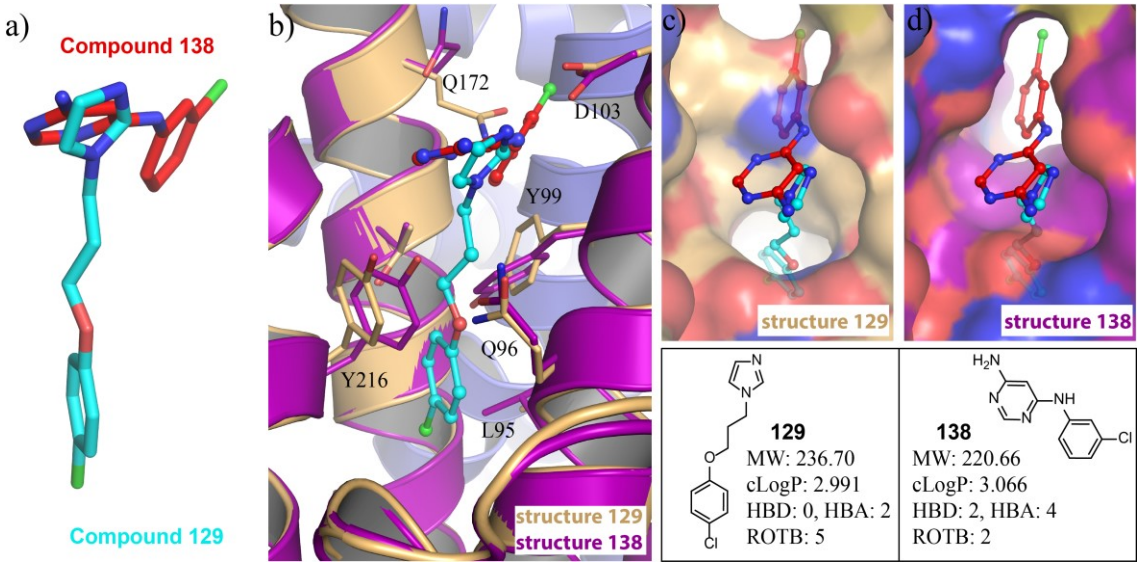
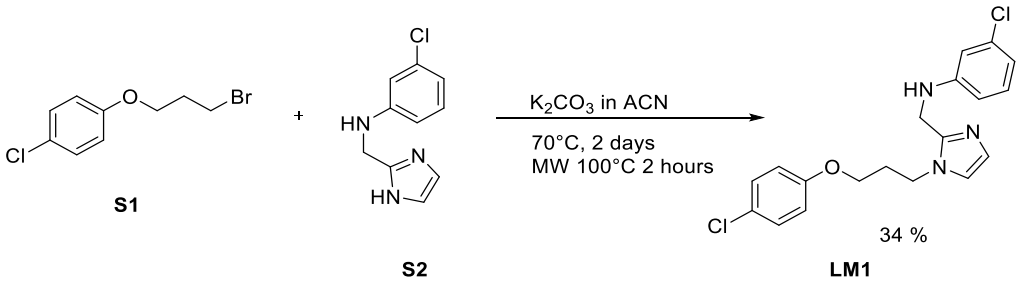


Figure 68: Overlay of *T. brucei* FPPS active-site binders for fragment merging approach. (a) Overlay of compound **129** (cyan) and **138** (red) in conformation as bound to the *T. brucei* FPPS active site (protein backbone of *T. brucei* FPPS in complex with the respective compound is not shown) **(b)** Cartoon representation of *T. brucei* FPPS in complex with **129** (protein backbone light orange, compound cyan) overlaid with *T. brucei* FPPS in complex with **138** (protein backbone violet, compound red). **(c)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of *T. brucei* FPPS in complex with **129** overlaid with compound **138 (d)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, violet: carbon) of *T. brucei* FPPS in complex with **138** overlaid with **129.**

The following labelling concept is applied for reaction schemes: starting materials, intermediates and final products that were not obtained in sufficient yield for testing are labelled with an **S**, and products that were further tested are labelled with **LM** and a running number. The synthesis scheme for the merged ligand **LM1** was inspired by Zhang *et al.*[302]. **LM1** was synthesised via one-step starting from 1-(3-bromopropoxy)-4-chlorobenzene (**S1**) and N-((1*H*-imidazol-2-yl)methyl)-3-chloroaniline (**S2**). In a nucleophilic substitution, 1-(3-bromopropoxy)-4-chlorobenzene (**S1**) was selectively attached to the imidazole ring in acetonitrile and finally **LM1** was obtained in a 34 % yield.



Scheme 1: Synthesis of merged compound via imidazole. 1-(3-bromopropoxy)-4-chlorobenzene (S2) and S2: N-((1*H*-imidazol-2-yl)methyl)-3-chloroaniline are synthesised to **LM1**: 3-chloro-N-((1-(3-(4-chlorophenoxy)propyl)-1*H*-imidazol-2-yl)methyl)aniline.

LM1 had a calculated clogP of 4.95, showed poor solubility of 16 µM determined by NMR but was a medium shifter in protein-observed NMR despite its low solubility. This suggests a potentially high binding affinity. Crystallization trials to determine the binding mode and to investigate SAR were unsuccessful. Possibly the solubility of the compound was insufficient to obtain a crystal structure with the bound fragment. In a next step, the flexibility of **129** was addressed and different synthesis strategies were tested to introduce rigidity in the four-membered flexible linker.

7.2.2 Introducing rigidity in elongated active-site binder

To reduce the number of rotatable bonds in **129**, it was aimed to increase the rigidity by a bicyclic system on the flexible linker between the para-substituted phenyl ring and the imidazole moiety (Figure 69). To this end, a small library of different para-substituted analogues of compound **129** based on an indazole and isoindole scaffold were attempted to be synthesised. The chlorine atom in **129** is positioned in 3.4 Å distance to the backbone oxygen of Phe161 at an angle of 135 °. Halogen interactions form in a "head on"-type of 180 °, and below 140 ° no significant interactions are formed.[538] Therefore, the binding affinity could significantly increase with a different scaffold to optimise the interaction angle between the terminal halogen and the backbone carboxyl (Figure 69).

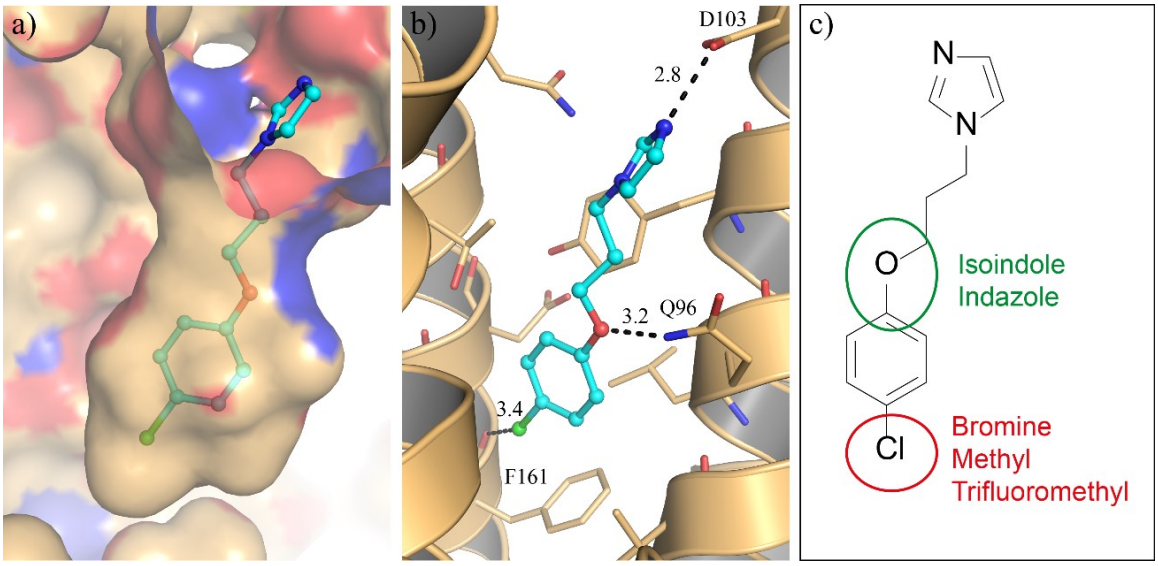
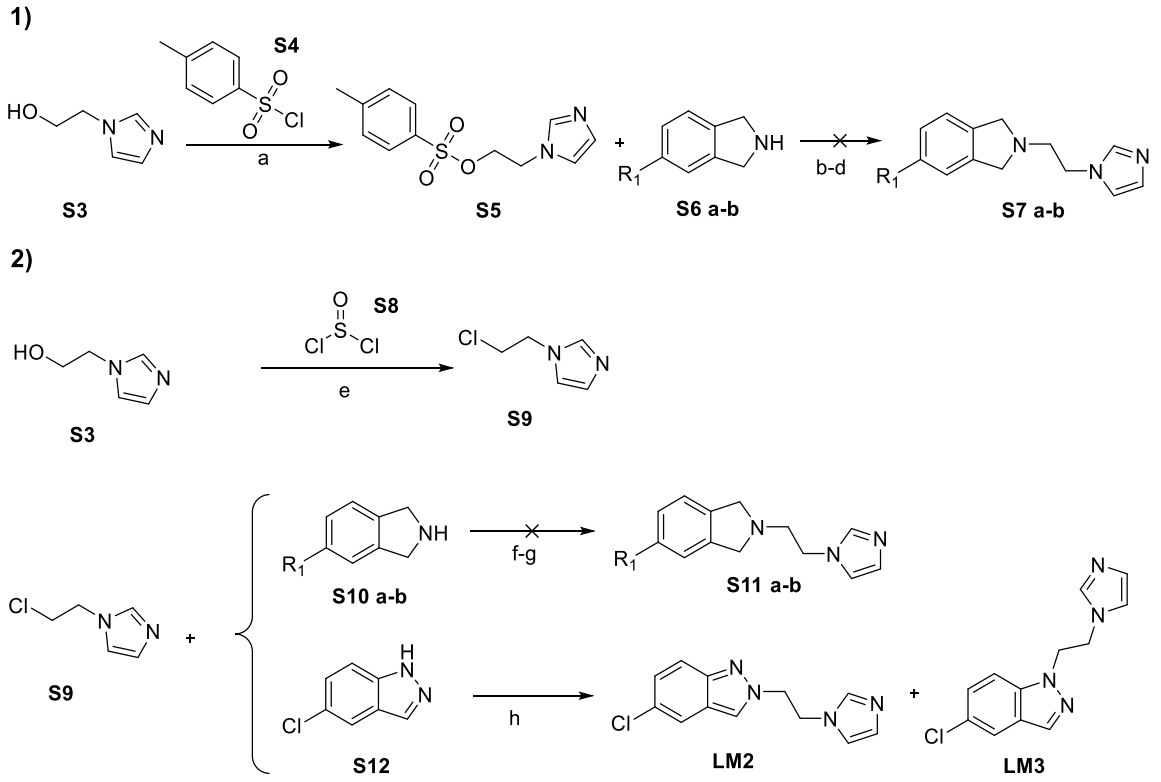


Figure 69: Medicinal chemistry optimisation by introducing rigidity on 129. The figure shows the structure of *T. brucei* FPPS in complex with **129**. Compound **129** is depicted as stick and ball representation (cyan). **(a)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) of *T. brucei* FPPS and view on the protein surface around **129**. **(b)** Cartoon and stick representation of residues in the binding site. Key atomic distances between the protein backbone and ligand are indicated by black dashed lines. Residues and atomic distances are labelled. **(c)** Chemical structure of **129** with modifications of the chlorine labelled and marked in red. Modifications on the ring system are labelled and marked in green.

Additionally, the presence of larger halogen atoms has shown to remarkably enhance protein-ligand interactions and was recently investigated in projects on human cathepsin L and MEK1 kinase.[541] More importantly, a close analogue of **129** with a para-substituted bromine was identified to retain *T. brucei* FPPS binding as detected by protein-observed NMR (see Chapter 5.2.8). Consequently, the chlorine was replaced by a bromine on an isoindoline scaffold and chloro-, methyl-, trifluoromethyl-isoindoles were attempted to be synthesised for initial SAR studies.

Different synthesis strategies for the alkylation of isoindoles were tested (Scheme 2). In a first attempt, the imidazole with the terminal hydroxyl group was activated by the formation of the tosyl-imidazole derivative, but further reaction with **S6** did not yield the desired target compounds. Moreover, as starting material the analogous chlorine imidazole **S9** was prepared and added in excess to the substituted isoindoles under basic conditions (Scheme 2).[542-544]

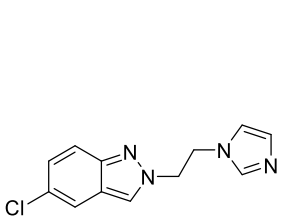


Scheme 2: Reagents and conditions for the synthesis of isoindoles, indazole, and indazolones. 1) Synthesis scheme from 2-(1*H*-imidazol-1-yl)ethan-1-ol (**S3**) to 2-(2-(1*H*-imidazol-1-yl)ethyl)-5-bromoisoindoline (**S7a**) and 2-(2-(1*H*-imidazol-1-yl)ethyl)-5-methylisoindoline (**S7a**) via 2-(1*H*-imidazol-1-yl)ethyl 4-methylbenzenesulfonate (**S5**) and the respective isoindoline (**S6a-b**). **2)** Synthesis scheme from 2-(1*H*-imidazol-1-yl)ethan-1-ol (**S3**) to 2-(2-(1*H*-imidazol-1-yl)ethyl)-5-methylisoindoline (**S11a**) and 2-(2-(1*H*-imidazol-1-yl)ethyl)-5-(trifluoromethyl)isoindoline (**S11b**) via 1-(2-chloroethyl)-1*H*-imidazole (**S9**) and the respective isoindole (**S10a-b**). Synthesis scheme from 2-(1*H*-imidazol-1-yl)ethan-1-ol (**S3**) to 2-(2-(1H-imidazol-1-yl)ethyl)-5-chloro-2H-indazole (**LM2**) and 1-(2-(1H-imidazol-1-yl)ethyl)-5-chloro-1H-indazole (**LM3**) via 1-(2-chloroethyl)-1*H*-imidazole (**S9**) and 5-chloro-1*H*-indazole (**S12**). **(a)** 1.5 eq TEA, THF 0 °C, 1 eq tosyl-Cl, 0 °C, 1 h -> rt 2 h; **(b)** R_1= Br: 3 eq K_2CO_3, DMF, 1:30 h, rt – 60 ° C, o/n; **(c)** R_1 = Br: 2.5 eq NaH, 1:30 h, 0 °C – rt, o/n; **(d)** R_1 = CH_3: 3 eq TEA, DCM, 0 ° C rt, 1:30 h – rt, o/n – microwave 100 °C, 1 h; **(e)** DCM, 0 °C - rt, 4 h; (f) R_1= Cl, 2.5 eq imidazole **S10**, 3 eq K_2CO_3, DMF, 1:30 h, rt – microwave 100 °C 2 x 1 h; **(g)** R_1 = CF_3, 3 eq imidazole **S9**, 1.1 eq. NaH, DMF, 0 °C, 30 min – 0 °C o/n – 60 °C, 2 days; **(h)** 2 eq imidazole **S10**, 3 eq K_2CO_3, DMF, 60 °C, o/n.

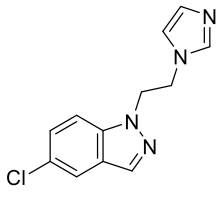
Potassium carbonate, sodium hydride and trimethylamine were investigated as bases but major side products were formed in all syntheses with only negligible product formation insufficient for any further analysis. Interestingly, the reaction of imidazole chlorine with methyl isoindole or chloro-isoindole yielded a product with a mass + 44 higher than the anticipated mass of the desired product. The structure of this side product could not be determined and thus, synthesis efforts were focused on indazole scaffolds.

The similar reactivity of the two nitrogens for alkylation in indazole resulted in a 48 % mixture of the two alkylated indazoles **LM2** and **LM3** using chlorine-imidazole **S9** under basic conditions with potassium carbonate (Scheme 2). Initially, the alkylation product at the nitrogen in position 2 was desired but both products **LM2** and **LM3** were separated and obtained in 15.8 % and 7.75 % yield, respectively (Table 32). The separation was performed using preparative SCF purification as the two compounds had very similar chemical properties with overlapping peaks in UPLC-MS. The extensive purification resulted in loss of product and thus, explains the low yield. Interestingly, the two alkylated indazoles **LM2** and **LM3** were both medium binders in NMR and crystallization experiments to determine the binding mode were performed. However, crystals immediately dissolved upon addition of the compound and crystal soaking experiments to reveal the binding mode failed.

Table 32: Synthesised compounds with an indazole scaffold.



LM2



LM3

Compound	MW	Yield	Protein-observed NMR[1]
LM2	246.7	15.8 %	medium
LM3	246.7	7.6 %	medium

[1] Classification of compounds in weak, medium and strong shifter on *T. brucei* FPPS according to [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K.

7.2.3 Fragment merging strategy with *T. cruzi* FPPS active-site binder

In a parallel project of Joy Petrick, NIBR, Basel, Switzerland, several active-site binders were identified on *T. cruzi* FPPS by an X-ray fragment screen. As FPPS of the two parasites shows a high structural similarity, active-site binders on *T. cruzi* and *T. brucei* FPPS were overlaid and the potential for fragment optimisation using fragment merging was studied (Figure 70).

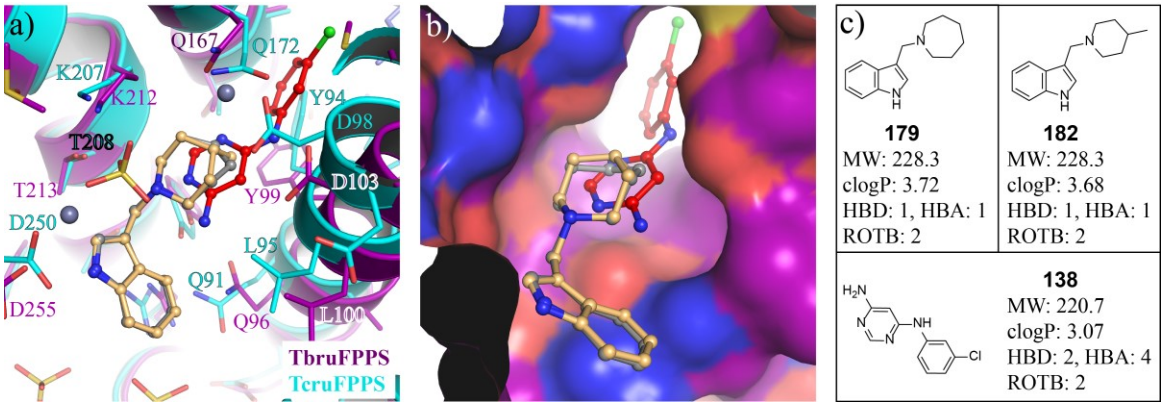


Figure 70: Overlay of *T. brucei* and *T. cruzi* FPPS active-site binders for fragment merging approach. (a) Cartoon representation of *T. brucei* FPPS in complex with **138** (protein backbone violet, compound red) overlaid with *T. cruzi* FPPS in complex with **179** (protein backbone cyan, compound grey). Compound **182** is superimposed (protein backbone not shown). **(b)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, violet: carbon) of *T. brucei* FPPS in complex with **138** overlaid with **178** and **182** (protein backbone not shown). **(c)** Chemical structures of compounds **138**, **179**, **182** with their respective molecular weight, clogP value, number of hydrogen-bond donors (HBD), hydrogen-bond acceptor (HBA) and rotatable bonds (ROTB).

Compounds **179** and **182** were identified in the active site on *T. brucei* FPPS (Figure 70). The azepane and piperidine ring of **179** and **182** attached to an indole moiety are positioned on top of the pyrimidine ring of **138**. However, **179** and **182** were previously determined as weak shifters by protein-observed NMR but no crystal structure could be determined for **179** and **182** on *T. brucei* FPPS. Although the residues around the two binders are identical in *T. brucei* and *T. cruzi* FPPS, the side chains are in different conformations and the *T. cruzi* FPPS helix α_D is positioned slightly shifted towards the bound compounds. Nevertheless, the conformation of *T. brucei* FPPS in complex with **138** could support a fragment merging of the *T. brucei* and *T. cruzi* FPPS active-site binders (Figure 70).

Fragment merging is one of the three main strategies of fragment optimisation and success stories of drug discovery programs are described in the literature[363, 486-489]. A small series of fragment mergers with a chlorine or methoxy substitution at the indole ring or compounds containing a pyrrolopyridine scaffold were synthesised to investigate SAR. **LM4** – **LM9** were synthesised of the corresponding aldehyde and chloro-substituted benzylpiperidine in a reductive amination following a synthesis protocol inspired by Jeankumar *et al*. (Scheme 3).[545]

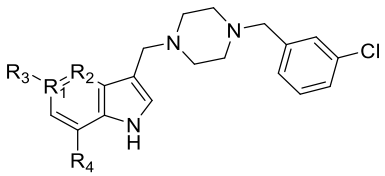
An amine is formed in a one-pot reductive amination by two telescoped steps. First, an imine is formed as intermediate, which is subsequently reduced by utilizing sodium triacetoxyborohydride as a mild hydride reagent to give the final amine. Interestingly, after 2 hours no or limited formation of the imine was detected in UPLC-MS. Possibly the reaction proceeds via direct reduction of the hemiaminal species formed by the reaction of the amine with the carbonyl group without the necessity for an intermediate imine. This is also supported by the possibility of performing the reaction in aqueous solutions.[546-547]

Scheme 3: General synthesis scheme of merged compound based on piperazine-linker. Final product 3-((4-(3-chlorobenzyl)piperazin-1-yl)methyl)-1H-indole (**LM4**) is formed in a reductive amination of 1*H*-indole-3-carbaldehyde (**S13**) and 1-(3-chlorobenzyl)piperazine (**S14**).

The complete formation of final products **LM4** – **LM9** was achieved overnight. The crude products were extracted with DCM and purified by prepLC-MS. The reaction yielded **LM4** – **LM9** in isolated yields ranging from 6.9 % to 46.3 % with purities >95 % as determined by UPLC-MS or NMR. Solubility of the synthesised fragment mergers was determined by 1D-zgegsp NMR experiments using DSS as an internal reference with nominal concentration. Product solubility was diminished in **LM9** exhibiting a chloro-substitution on the indole scaffold and ranged up to above 1000 µM in compounds **LM4** and **LM7** (Table 33). By subjecting synthesised compounds to protein-observed NMR, **LM9** was identified to precipitate *T. brucei* FPPS despite the low solubility.

In contrast, indole **LM4** and substituted indoles **LM7** and **LM8** were classified as medium shifters. Crystal soaking experiments were performed with all fragments but no protein-ligand complex structure with a bound fragment merge could be solved. Therefore, it was impossible to determine if the fragment mergers retained the binding mode in the active site.

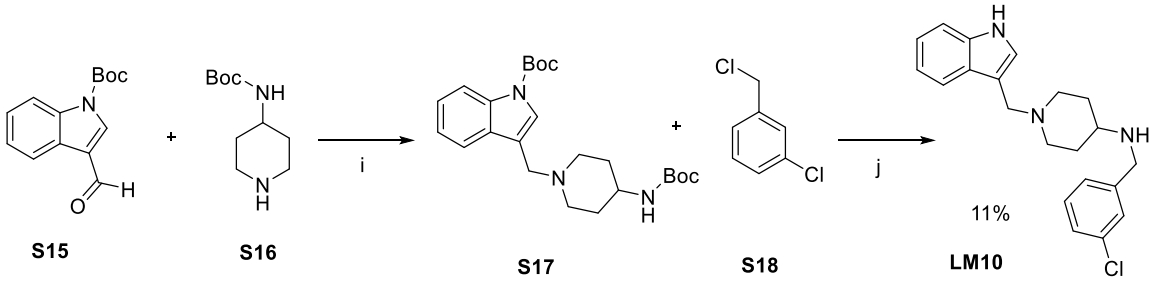
Table 33: Fragment mergers with a piperazine-linker.



Compound	R_1	R_2	R_3	R_4	MW	Yield [%]	Solubility[1] [µM]	Protein-observed NMR[2]
LM4	C	CH_2	H	H	339.9	46.3	1120	medium
LM5	N	CH	-	H	340.9	22.9	510	weak
LM6	C	N	H	H	340.9	45.3	510	weak
LM7	C	CH	OH	H	355.9	6.9	1160	medium
LM8	C	CH	H	OCH_3	369.9	46.1	100	medium
LM9	C	CH	H	Cl	374.3	27.1	<10	precipitation

[1] Solubility determined at 1 mM compound concentration in 1D-zgegsp NMR experiments at 296 K.
[2] Classification of compounds in weak, medium, strong shifter and compound resulting in *T. brucei* FPPS precipitation according to [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K.

In addition to fragment merging based on a piperazine linker containing indole, substituted indole, and pyrrolopyridine scaffolds, **LM10** was synthesised exhibiting an aminopiperidine linker. First, intermediate **S17** was formed from the boc-protected indole aldehyde and the boc-protected piperidine was subjected to a reductive amination in DCM. Boc protection of the primary amine in piperidine enabled selective amine formation with the secondary amine of the piperidine scaffold. After isolation of **S17**, **LM10** was formed of **S17** and **S18** in a nucleophilic substitution reaction under basic conditions utilizing sodium hydride in DMF. The synthesis protocol was inspired by Ombrato *et al.*[548] Subsequently, deprotection was carried out in an acidic mixture of TFA and DCM and after purification by prepLC-MS, **LM10** was obtained in 8.8 % yield.



Scheme 4: Synthesis of fragment merging based on an aminopiperidine linker. In a reductive amination of *tert*-butyl 3-formyl-1*H*-indole-1-carboxylate (**S15**) and tert-butyl piperidin-4-ylcarbamate (**S16**), tert-butyl 3-((4-((tert-butoxycarbonyl)amino)piperidin-1-yl)methyl)-1H-indole-1-carboxylate (**S17**) was formed. **S17** and 1-chloro-3-(chloromethyl)benzene (**S18**) were reacted to give the product 1-((1*H*-indol-3-yl)methyl)-N-(3-chlorobenzyl)piperidin-4-amine (**LM10**). **(i)** cat. AcOH, DCM, rt, 6 h; 1.5 eq $NaBH(OAc)_3$, rt, o/n; **(j)** NaH, DMF, 0 °C 1 h – rt, o/n; TFA/DCM, rt, 1h

LM10 binding on *T. brucei* FPPS was studied in protein-observed NMR and strong chemical shift perturbations were observed, which classified **LM10** as a strong shifter (Figure 71).

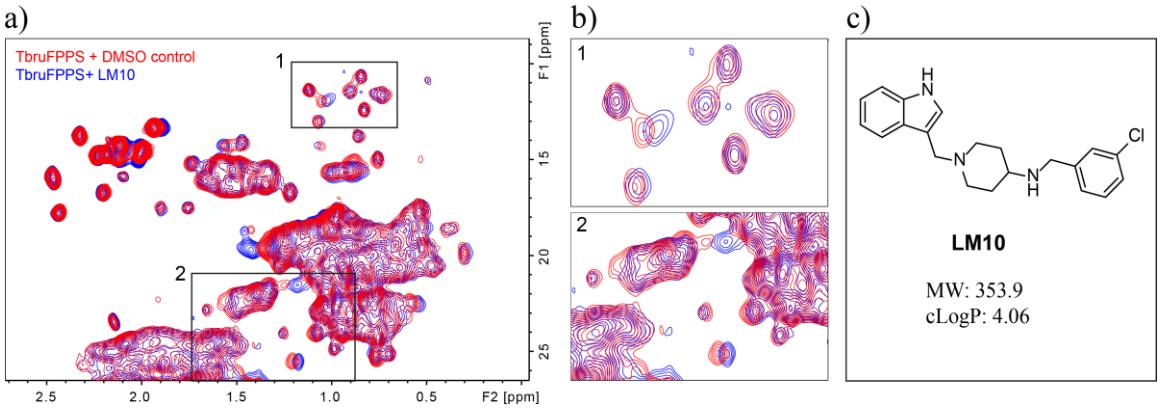


Figure 71: Protein-observed NMR of *T. brucei* FPPS with aminopiperidine LM10. (a) The [^{13}C, ^{1}H]-SOFAST-HMQC spectrum (32 scans) of 30 µM *T. brucei* FPPS with 1 mM compound measured at 310 K is shown. **(b)** Zoom-in on two areas of the [^{13}C, ^{1}H]-SOFAST-HMQC spectrum labelled with **1** and **2** in **a)**. **(c)** Chemical structure of **LM10**. Molecular weight and clogP value are noted.

Encouraged by the protein-observed NMR results, crystal soaking experiments were carried out at 10 mM compound concentration for 1 day and 50 mM compound concentration for 6 hours

or 1 day. Five diffraction data sets were collected with a diffraction limit above 2.8 Å. Diffraction data sets were analysed as part of the PanDDA run described in Chapter 5.2.4 but identified PanDDA events did not fit the ligand. The obtained X-ray diffraction limit was not sufficient which is likely why the PanDDA maps did not indicate a binding event. Conventional difference Fo-Fc map at 3 σ revealed positive electron density in a new binding site below loop α_H-α_I. For initial mFo-DFc difference electron density maps, 2 mFo-DFc electron density maps, and refined 2Fo-Fc and Fo-Fc electron density maps, see Appendix, Chapter 10.5.3, Figure 80.

LM10 was fitted and refined using BUSTER. Upon addition of **LM10**, the loop was shifted to make space for compound binding. The binding site is formed by the SARM on helix α_H and extends towards helix α_G (Figure 72). This binding site is near the active site and to the binding location of previously identified fragments as part of the NMR and X-ray fragment screen (see Chapter 5.2.5 and 6.2.2). The aminopiperidine linker forms three H-bonds to Asp256, Asp259 and Asp273. Additionally, the indole is buried below Lys269 and shows hydrophobic interactions with Met258. The chlorine-substituted phenyl ring points towards helix α_G and is surrounded by Tyr205, Val209, Gln252.

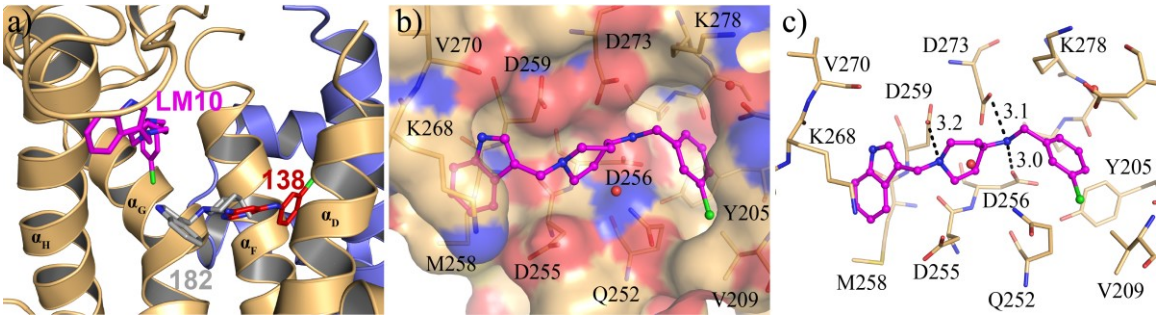


Figure 72: Fragment merger of *T. brucei* and *T. cruzi* FPPS active-site binders. *T. brucei* FPPS in complex with **LM10** (protein backbone, light orange, compound magenta). **(a)** Cartoon representation of *T. brucei* FPPS in complex with **LM10**. Compound **182** and **138** are overlaid (protein backbone not shown). **(b)** Surface representation colour-coded by atoms (blue: nitrogens, red: oxygens, light orange: carbon) and **(c)** Stick representation of **LM10** binding site. Atomic distances are indicated by black dashed lines. Residues and atomic distances are labelled.

LM10 was synthesised based on two fragments in the active site and did not adopt the expected binding mode. As soaking of larger ligands can be misleading and a ligand is enforced to adopt a geometry restrained by the packing of the protein in pre-manufactured crystals, a broad co-crystallization screen was performed following co-crystallization protocol 3 (see Methods, Chapter 2.9.7). Co-crystallization has the advantage that the protein-ligand complex forms in solution, any conformational adaption required to accommodate the ligand can take place and the formed complex crystallizes with its possible rearranged geometry. Twenty X-ray structures with diffraction limits ranging between 2.1 Å and 3.1 Å were solved. X-ray diffraction data were analysed but either the diffraction limit was not sufficient for fragment binding identification or no positive electron density was identified at 3 σ in the Fo-Fc map to fit the ligand. Therefore,

co-crystallization experiments failed to identify compound binding and potentially further co-crystallization experiments would be needed for additional protein-ligand complex structures. The hypothesis that **LM10** occupied a new binding site as a consequence that crystal soaking experiments did not allow for spatial adaptations of the protein required to accommodate the merged compound properly.

7.3 Discussion

After initial fragment screening, fragment-to-lead optimisation is performed to optimise binding affinities of initially weak fragment binders. In this chapter, obtained *T. brucei* FPPS structures were studied to identify an appropriate structure-based fragment optimisation strategy using medicinal chemistry. Fragment optimisation was focused on active-site binders as identified fragments showed high structural diversity, and orthogonal binding modes implicated potential for fragment merging. Encouraged by the identification of the first non-bisphosphonate active-site binders, the aim was to optimise these fragments further and investigate their SAR.

In total, ten compounds were synthesised based on three different fragment optimisation strategies. Compound scaffolds chosen for synthesis had to be easily accessible in an up to three step reaction and should allow for a diverse investigation of SAR. First, synthesis efforts were targeted on orthogonal binders **129** and **138** using the fragment merging approach. Fragment merging was straightforward in a one-step reaction based on a nucleophilic substitution reaction. However, the finally obtained fragment merger **LM1** exhibited low solubility and hence, no protein-ligand structure could be solved. For further fragment optimisation based on the two orthogonal active-site binders, the solubility issue would need to be addressed in follow-up compounds by e.g. introducing polar functional groups, such as amines or carboxylic acids. Additionally, both starting fragments require conformational adaptions on the protein, which could also require co-crystallization experiments.

Considering that **129** contains a four atom long flexible linker, further more rigid compounds which freeze the ligand in the geometry required at the binding site could guide fragment optimisation. Therefore, the next synthesis efforts targeted **129**. The compound's high conformational flexibility was attempted to be addressed by different bicyclic scaffolds, such as benzofuran, indazole, and isoindoles to introduce rigidity. A straightforward synthesis strategy for benzofuran could not be developed and therefore, only the synthesis of nitrogen-containing bicyclic scaffolds was performed. However, alkylation of the isoindole amine was unsuccessful but alkylation on both nitrogens of an indazole scaffold could be obtained and products **LM3** and **LM4** were isolated. Strikingly, both products were medium shifters in protein-observed NMR. Crystal soaking experiments failed as crystals dissolved immediately upon compound addition. Possibly, a different crystal system would need to be optimized or co-crystallization would need to be

performed as possible the crystal packing interferes with compound binding. Alternatively, compound binding could induce significant conformational changes, which break the crystal lattice. At this stage, no additional crystallization trials were performed and instead, crystal structures with active-site binders found for the homologous *T. cruzi* FPPS were studied. Synthesis based on fragments bound to a homologue FPPS can accelerate the fragment optimisation process as compounds with additional structural diversity are included but at the same time, compounds identified on *T. cruzi* FPPS do not necessarily bind at the same binding site on *T. brucei* FPPS since both proteins exhibit structural diversities. However, residues around the active-site binders **179** and **182** (see Chapter 7.2.3, Figure 70) present in *T. cruzi* FPPS overlaid convincingly well with the same residues in *T. brucei* FPPS. Additionally, **179** and **182** overlaid with **138** bound to *T. brucei* FPPS and thus, showed potential for fragment merging. Synthesis of fragment mergers **LM4**-**LM9** was straightforward via reductive amination of the corresponding aldehyde and aminopiperidine. In this work, synthesis yielded compounds in sufficient to good yields for biophysical characterization. Compound binding was only observed by protein-observed NMR and no protein-ligand crystal structure was solved with a bound fragment merger. Therefore, the binding mode and effect of the different scaffolds could not be studied and additional crystallization trials, possibly including co-crystallization experiments would be needed.

Alternatively to the compounds with a piperazine-linker, **LM10** with an aminopiperidine scaffold was synthesised in a two-step reaction. Interestingly, the synthesised fragment merger **LM10** was a medium shifter in protein-observed NMR and crystal soaking experiments revealed that the compound occupied a new binding site close to the SARM. Binding in this area would prevent conformational changes from open-to-closed conformation occurring during condensation reactions of IPP and DMAPP to form GPP and further GPP and IPP to release FPP. Therefore, this binding site could also be biologically relevant to block enzyme function but was not further investigated as part of this thesis. However, compound binding at this binding site could have occurred only for the fact that required conformational changes were prevented in crystal soaking experiments. In consequence, the binding pose of the merged compound cannot occupy the active site thus superimposing with the starting fragments. Therefore, co-crystallization trials were performed but no protein-ligand complex structure with the bound merged compound was obtained.

8. Conclusion and Outlook

In this work, the structural superimposition of the crystal structures of human FPPS on *T. brucei* FPPS was performed and revealed the same allosteric site in both proteins. Furthermore, sequence alignment showed that the *T. brucei* FPPS allosteric site has a 70 % amino acid identity to the reported allosteric site on human FPPS. Fragment screening by NMR and X-ray crystallography was combined to test whether the reported human FPPS allosteric binders can also be identified as binding compounds on *T. brucei* FPPS. Therefore, *T. brucei* FPPS was expressed unlabelled, $^{13}C^{15}N$-labelled and avi-tagged biotinylated in *E. coli* cells and the protein was obtained in yields up to 280 mg at >95 % purity. For the first time, allosteric site binders on *T. brucei* FPPS were detected, and moreover, the first non-bisphosphonate ligands in the active site were identified. Protein-ligand complex crystal structures were analysed using PanDDA to detect even weak fragment binders and also obtain structures with fragments bound at low occupancy in order to guide a subsequent fragment optimisation strategy. The combination of NMR and X-ray fragment screening and testing of human allosteric binders on *T. brucei* FPPS revealed five fragments in the active site, seven fragments in the previously unknown allosteric site and five fragments at additional binding sites. Remarkably, one active-site binder introduced conformational changes upon binding to open up a binding site along helix a_D and the compound was found to adopt a binding mode along helix a_D which was previously not seen in protein-bisphosphonate structures.

Initial fragment hits were followed-up by SAR using compounds from the archive and catalogue. For a fragment analogue of a previously detected active-site binder, pocket hopping into the allosteric site was revealed by crystal soaking experiments. Starting points for structure-based fragment optimisation were identified and ten compounds were synthesised following the fragment merging and growing approach of fragments bound in the active site. To this end, also *T. cruzi* FPPS active-site binders were analysed for potential binding on *T. brucei* FPPS and conceptionally merged compounds were synthesised based on two spatially overlapping *T. brucei* and *T.cruzi* active-site binders. Synthesised compounds showed weak to strong chemical shift perturbations in protein-observed NMR and were subjected to crystallization experiments. Strikingly, crystal soaking experiments revealed binding of one merged compound based on a *T. brucei* and *T. cruzi* FPPS active-site binder in a new binding site close to the second aspartate-rich motif remote from the binding site where the starting fragments have been located in the active site.

In total, 19 protein-ligand structures with fragments in seven different binding sites were determined. The overlay of the different binders paves the way for fragment-to-lead optimisation and further design hypotheses include the synthesis of follow-up compounds of fragments identified in the allosteric site. Structural scaffolds and binding modes of identified allosteric binders show potential for fragment merging. Up to now, the discovery on *T. brucei* FPPS was limited to

bisphosphonate analogues in the active site and restricted by the pharmacokinetic properties thereof. The discovery of binders with a new chemotype sets the framework for a fragment-to-lead optimisation towards compounds with pharmacokinetic properties suitable for parasitic indications. Investigating the biological relevance of the novel binding sites discovered as part of this thesis could result in new druggable targets to inhibit FPPS activity in *T. brucei* parasites.

9. References

1. Welburn, S. C., Fèvre, E. M., Coleman, P. G.*, et al.*, Sleeping sickness: a tale of two diseases. *Trends Parasitol* **2001**, *17*, 19-24.
2. Steverding, D., The history of African trypanosomiasis. *Parasit Vectors* **2008**, *1*, 3.
3. Langousis, G., Hill, K. L., Motility and more: the flagellum of Trypanosoma brucei. *Nat Rev Microbiol* **2014**, *12*, 505-18.
4. Büscher, P., Cecchi, G., Jamonneau, V.*, et al.*, Human African trypanosomiasis. *The Lancet* **2017**, *390*, 2397-2409.
5. Wamwiri, F. N., Changasi, R. E., Tsetse Flies (Glossina) as Vectors of Human African Trypanosomiasis: A Review. *Biomed Res Int* **2016**, *2016*, 1-8.
6. Franco, J. R., Simarro, P. P., Diarra, A.*, et al.*, Epidemiology of human African trypanosomiasis. *Clin Epidemiol* **2014**, *6*, 257-275.
7. Ooi, C. P., Bastin, P., More than meets the eye: understanding Trypanosoma brucei morphology in the tsetse. *Front Cell Infect Microbiol* **2013**, *3*, 71.
8. Fenn, K., Matthews, K. R., The cell biology of Trypanosoma brucei differentiation. *Curr Opin Microbiol* **2007**, *10*, 539-46.
9. Sharma, R., Gluenz, E., Peacock, L.*, et al.*, The heart of darkness: growth and form of Trypanosoma brucei in the tsetse fly. *Trends Parasitol* **2009**, *25*, 517-24.
10. Kennedy, P. G., Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). *Lancet Neurol* **2013**, *12*, 186-94.
11. Gibson, W., Bailey, M., The development of Trypanosoma brucei within the tsetse fly midgut observed using green fluorescent trypanosomes. *Kinetoplastid Biol Dis* **2003**, *2*, 1.
12. Dyer, N. A., Rose, C., Ejeh, N. O.*, et al.*, Flying tryps: survival and maturation of trypanosomes in tsetse flies. *Trends Parasitol* **2013**, *29*, 188-96.
13. Brun, R., Blum, J., Chappuis, F.*, et al.*, Human African trypanosomiasis. *Lancet* **2010**, *375*, 148-59.
14. MacGregor, P., Matthews, K. R., New discoveries in the transmission biology of sleeping sickness parasites: applying the basics. *J Mol Med (Berl)* **2010**, *88*, 865-71.
15. World Health Organization, Control and surveillance of human African trypanosomiasis Report of a WHO Expert Committee; *WHO Technical Report Series 984*. World Health Organization Geneva, Switzerland, 2013.
16. da Silva, A. J., Moser, M., Trypanosomiasis, African. https://www.cdc.gov/dpdx/trypanosomiasisafrican/index.html (accessed 26th of February 2019).
17. Berriman, M., Ghedin, E., Hertz-Fowler, C.*, et al.*, The genome of the African trypanosome Trypanosoma brucei. *Science* **2005**, *309*, 416-22.
18. Barry, J. D., Marcello, L., Morrison, L. J.*, et al.*, What the genome sequence is revealing about trypanosome antigenic variation. *Biochem Soc Trans* **2005**, *33*, 986-989.
19. Taylor, J. E., Rudenko, G., Switching trypanosome coats: what's in the wardrobe? *Trends Genet* **2006**, *22*, 614-20.
20. Donelson, J. E., Antigenic variation and the African trypanosome genome. *Acta Trop* **2003**, *85*, 391-404.
21. Raper, J., Fung, R., Ghiso, J.*, et al.*, Characterization of a novel trypanosome lytic factor from human serum. *Infect immun* **1999**, *67*, 1910-1916.
22. Capewell, P., Veitch, N. J., Turner, C. M.*, et al.*, Differences between Trypanosoma brucei gambiense groups 1 and 2 in their resistance to killing by trypanolytic factor 1. *PLoS Negl Trop Dis* **2011**, *5*, e1287.
23. Stephens, N. A., Hajduk, S. L., Endosomal localization of the serum resistance-associated protein in African trypanosomes confers human infectivity. *Eukaryot cell* **2011**, *10*, 1023-1033.
24. Van Xong, H., Vanhamme, L., Chamekh, M.*, et al.*, A VSG Expression Site–Associated Gene Confers Resistance to Human Serum in Trypanosoma rhodesiense. *Cell* **1998**, *95*, 839-846.
25. DeJesus, E., Kieft, R., Albright, B.*, et al.*, A single amino acid substitution in the group 1 Trypanosoma brucei gambiense haptoglobin-hemoglobin receptor abolishes TLF-1 binding. *PLoS Pathog* **2013**, *9*, e1003317.
26. Capewell, P., Clucas, C., DeJesus, E.*, et al.*, The TgsGP gene is essential for resistance to human serum in Trypanosoma brucei gambiense. *PLoS Pathog* **2013**, *9*, e1003686.
27. Franco, J. R., Cecchi, G., Priotto, G.*, et al.*, Monitoring the elimination of human African trypanosomiasis: Update to 2016. *PLoS Negl Trop Dis* **2018**, *12*, e0006890.

28. Franco, J. R., Cecchi, G., Priotto, G.*, et al.*, Monitoring the elimination of human African trypanosomiasis: Update to 2014. *PLoS Negl Trop Dis* **2017**, *11*, e0005585.
29. WHO/CDS/CSR/ISR/2000.1, WHO/CDS/CSR/ISR/2000.1, *Chapter 8 - African trypanosomiasis.* . World Health Organization 2000.
30. Fevre, E. M., Wissmann, B. V., Welburn, S. C.*, et al.*, The burden of human African trypanosomiasis. *PLoS Negl Trop Dis* **2008**, *2*, e333.
31. Simarro, P. P., Franco, J. R., Cecchi, G.*, et al.*, Human African Trypanosomiasis in Non-Endemic Countries (2000–2010). *Journal of Travel Medicine* **2012**, *19*, 44-53.
32. Willett, K. C., The "trypanosome chancre" in Rhodesian sleeping sickness. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **1966**, *60*, 689-690.
33. Apted, F. I. C., Clinical manifestations and diagnosis of sleeping sickness. In *The African trypanosomiasis*, Mulligan, H. W., Ed. George Allen & Unwin: 1970.
34. Kennedy, P. G., Human African trypanosomiasis of the CNS: current issues and challenges. *J Clin Invest* **2004**, *113*, 496-504.
35. MacLean, L. M., Odiit, M., Chisi, J. E.*, et al.*, Focus-specific clinical profiles in human African Trypanosomiasis caused by Trypanosoma brucei rhodesiense. *PLoS Negl Trop Dis* **2010**, *4*, e906.
36. Odiit, M., Kansiime, F., Enyaru, J. C. K., Duration of symptoms and case fatality of sleeping sickness caused by Trypanosoma brucei rhodesiense in Tororo, Uganda. **1998**, *74*, 792-5.
37. Checchi, F., Filipe, J. A., Barrett, M. P.*, et al.*, The natural progression of Gambiense sleeping sickness: what is the evidence? *PLoS Negl Trop Dis* **2008**, *2*, e303.
38. Checchi, F., Filipe, J. A., Haydon, D. T.*, et al.*, Estimates of the duration of the early and late stage of gambiense sleeping sickness. *BMC Infectious Diseases* **2008**, *8*, 16.
39. Buguet, A., Bourdon, L., Bouteille, B.*, et al.*, The duality of sleeping sickness: focusing on sleep. *Sleep Medicine Reviews* **2001**, *5*, 139-53.
40. Blum, J., Schmid, C., Burri, C., Clinical aspects of 2541 patients with second stage human African trypanosomiasis. *Acta Trop* **2006**, *97*, 55-64.
41. Kato, C. D., Nanteza, A., Mugasa, C.*, et al.*, Clinical profiles, disease outcome and co-morbidities among T. b. rhodesiense sleeping sickness patients in Uganda. *PLoS One* **2015**, *10*, e0118370.
42. Baker, C. H., Welburn, S. C., The Long Wait for a New Drug for Human African Trypanosomiasis. *Trends in Parasitology* **2018**, *34*, 818-827.
43. Kennedy, P. G., Diagnostic and neuropathogenesis issues in human African trypanosomiasis. *Int J Parasitol* **2006**, *36*, 505-12.
44. Wastling, S. L., Welburn, S. C., Diagnosis of human sleeping sickness: sense and sensitivity. *Trends Parasitol* **2011**, *27*, 394-402.
45. Magnus, E., Vervoort, T., Van Meirvenne, N., A card-agglutination test with stained trypanosomes (C.A.T.T.) for the serological diagnosis of T. B. gambiense trypanosomiasis. *Ann Soc Belg Med Trop.* **1978**, *58*, 169-76.
46. Inojosa, W. O., Augusto, I., Bisoffi, Z.*, et al.*, Diagnosing human African trypanosomiasis in Angola using a card agglutination test: observational study of active and passive case finding strategies. *BMJ* **2006**, *332*, 1479.
47. Stich, A., Ponte-Sucre, A., Holzgrabe, U., Do we need new drugs against human African trypanosomiasis? *Lancet Infect Dis* **2013**, *13*, 733-34.
48. Sands, M., Kron, M. A., Brown, R. B., Pentamidine: a review. *Rev Infect Dis* **1985**, *7*, 625-34.
49. Zhang, Y., Li, Z. J., Pilch, D. S.*, et al.*, Pentamidine inhibits catalytic activity of group I intron Ca.LSU by altering RNA folding. *Nucleic Acids Res* **2002**, *30*, 2961-71.
50. Munday, J. C., Eze, A. A., Baker, N.*, et al.*, Trypanosoma brucei aquaglyceroporin 2 is a high-affinity transporter for pentamidine and melaminophenyl arsenic drugs and the main genetic determinant of resistance to these drugs. *J Antimicrob Chemother* **2014**, *69*, 651-63.
51. Nagle, A. S., Khare, S., Kumar, A. B.*, et al.*, Recent developments in drug discovery for leishmaniasis and human African trypanosomiasis. *Chem Rev* **2014**, *114*, 11305-47.
52. Pohlig, G., Bernhard, S. C., Blum, J.*, et al.*, Efficacy and Safety of Pafuramidine versus Pentamidine Maleate for Treatment of First Stage Sleeping Sickness in a Randomized, Comparator-Controlled, International Phase 3 Clinical Trial. *PLoS Neglect Trop D* **2016**, *10*, e0004363.
53. Doua, F., Miezan, T. W., Sanon Singaro, J. R.*, et al.*, The efficacy of pentamidine in the treatment of early-late stage Trypanosoma brucei gambiense trypanosomiasis. *Am J Trop Med Hyg* **1996**, *55*, 586-8.
54. Dressel, J., The discovery of Germanin by Oskar Dressel and Richard Kothe. *J Chem Educ* **1961**, *38*, 620.
55. Babokhov, P., Sanyaolu, A. O., Oyibo, W. A.*, et al.*, A current analysis of chemotherapy strategies for the treatment of human African trypanosomiasis. *Pathog Glob Health* **2013**, *107*, 242-52.

56. Brun, R., Don, R., Jacobs, R. T., *et al.*, Development of novel drugs for human African trypanosomiasis. *Future Microbiol* **2011**, *6*, 677-91.
57. Steverding, D., The development of drugs for treatment of sleeping sickness: a historical review. *Parasit Vectors* **2010**, *3*, 15.
58. Denise, H., Barrett, M. P., Uptake and mode of action of drugs used against sleeping sickness. *Biochem Pharmacol* **2001**, *61*, 1-5.
59. Barrett, M. P., Boykin, D. W., Brun, R., *et al.*, Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. *Br J Pharmacol* **2007**, *152*, 1155-71.
60. Bacchi, C. J., Garofalo, J., Mockenhaupt, D., *et al.*, In vivo effects of alpha-DL-difluoromethylornithine on the metabolism and morphology of Trypanosoma brucei brucei. *Mol Biochem Parasitol* **1983**, *7*, 209-25.
61. Phillips, M. A., Coffino, P., Wang, C. C., Cloning and sequencing of the ornithine decarboxylase gene from Trypanosoma brucei. Implications for enzyme turnover and selective difluoromethylornithine inhibition. *J Biol Chem* **1987**, *262*, 8721-7.
62. Vincent, I. M., Creek, D., Watson, D. G., *et al.*, A molecular mechanism for eflornithine resistance in African trypanosomes. *PLoS Pathog* **2010**, *6*, e1001204.
63. Priotto, G., Kasparian, S., Ngouama, D., *et al.*, Nifurtimox-Eflornithine Combination Therapy for Second-Stage Trypanosoma brucei gambiense Sleeping Sickness: A Randomized Clinical Trial in Congo. *Clinical Infectious Diseases* **2007**, *45*, 1435-42.
64. Priotto, G., Kasparian, S., Mutombo, W., *et al.*, Nifurtimox-eflornithine combination therapy for second-stage African Trypanosoma brucei gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *The Lancet* **2009**, *374*, 56-64.
65. Lutje, V., Seixas, J., Kennedy, A., Chemotherapy for second-stage Human African trypanosomiasis. *Cochrane Database Syst Rev* **2010**, CD006201.
66. Alirol, E., Schrumpf, D., Amici Heradi, J., *et al.*, Nifurtimox-eflornithine combination therapy for second-stage gambiense human African trypanosomiasis: Medecins Sans Frontieres experience in the Democratic Republic of the Congo. *Clin Infect Dis* **2013**, *56*, 195-203.
67. Munoz, C., Zulantay, I., Apt, W., *et al.*, Evaluation of nifurtimox treatment of chronic Chagas disease by means of several parasitological methods. *Antimicrob Agents Chemother* **2013**, *57*, 4518-23.
68. World Health Organization, WHO Model List of Essential Medicines https://www.who.int/medicines/publications/essentialmedicines/en/ (accessed 21st January 2019).
69. Yun, O., Priotto, G., Tong, J., *et al.*, NECT is next: implementing the new drug combination therapy for Trypanosoma brucei gambiense sleeping sickness. *PLoS Negl Trop Dis* **2010**, *4*, e720.
70. Schmid, C., Kuemmerle, A., Blum, J., *et al.*, In-hospital safety in field conditions of nifurtimox eflornithine combination therapy (NECT) for T. b. gambiense sleeping sickness. *PLoS Negl Trop Dis* **2012**, *6*, e1920.
71. Schmid, C., Nkunku, S., Merolle, A., *et al.*, Efficacy of 10-day melarsoprol schedule 2 years after treatment for late-stage gambiense sleeping sickness. *The Lancet* **2004**, *364*, 789-90.
72. Schmid, C., Richer, M., Bilenge, C. M., *et al.*, Effectiveness of a 10-day melarsoprol schedule for the treatment of late-stage human African trypanosomiasis: confirmation from a multinational study (IMPAMEL II). *J Infect Dis* **2005**, *191*, 1922-31.
73. Blum, J., Nkunku, S., Burri, C., Clinical description of encephalopathic syndromes and risk factors for their occurrence and outcome during melarsoprol treatment of human African trypanosomiasis. *Trop Med Int Health* **2001**, *6*, 390-400.
74. Smilkstein, M., Sriwilaijaroen, N., Kelly, J. X., *et al.*, Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. *Antimicrob Agents Chemother* **2004**, *48*, 1803-6.
75. Bhatia, R., Gautam, A., Gautam, S. K., *et al.*, Assessment of SYBR green I dye-based fluorescence assay for screening antimalarial activity of cationic peptides and DNA intercalating agents. *Antimicrob Agents Chemother* **2015**, *59*, 2886-9.
76. Chatelain, E., Ioset, J. R., Phenotypic screening approaches for Chagas disease drug discovery. *Expert Opin Drug Discov* **2018**, *13*, 141-153.
77. Ortiz, D., Guiguemde, W. A., Hammill, J. T., *et al.*, Discovery of novel, orally bioavailable, antileishmanial compounds using phenotypic screening. *PLoS Negl Trop Dis* **2017**, *11*, e0006157.
78. Haasen, D., Schopfer, U., Antczak, C., *et al.*, How Phenotypic Screening Influenced Drug Discovery: Lessons from Five Years of Practice. *Assay Drug Dev Technol* **2017**, *15*, 239-46.
79. Räz, B., Iten, M., Grether-Bühler, Y., *et al.*, The Alamar Blue® assay to determine drug sensitivity of African trypanosomes (T.b. rhodesiense and T.b. gambiense) in vitro. *Acta Tropica* **1997**, *68*, 139-47.

80. Lanteri, C. A., Stewart, M. L., Brock, J. M.*, et al.*, Roles for the Trypanosoma brucei P2 transporter in DB75 uptake and resistance. *Mol Pharmacol* **2006**, *70*, 1585-92.
81. Merschjohann, K., Steverding, D., In vitro growth inhibition of bloodstream forms of Trypanosoma brucei and Trypanosoma congolense by iron chelators. *Kinetoplastid Biol Dis* **2006**, *5*, 3.
82. Sola, I., Castella, S., Viayna, E.*, et al.*, Synthesis, biological profiling and mechanistic studies of 4-aminoquinoline-based heterodimeric compounds with dual trypanocidal-antiplasmodial activity. *Bioorg Med Chem* **2015**, *23*, 5156-67.
83. Baell, J. B., Holloway, G. A., New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J Med Chem* **2010**, *53*, 2719-2740.
84. Sykes, M. L., Avery, V. M., Development of an Alamar Blue viability assay in 384-well format for high throughput whole cell screening of Trypanosoma brucei brucei bloodstream form strain 427. *Am J Trop Med Hyg* **2009**, *81*, 665-74.
85. Sykes, M. L., Baell, J. B., Kaiser, M.*, et al.*, Identification of compounds with anti-proliferative activity against Trypanosoma brucei brucei strain 427 by a whole cell viability based HTS campaign. *PLoS Negl Trop Dis* **2012**, *6*, e1896.
86. Mackey, Z. B., Baca, A. M., Mallari, J. P.*, et al.*, Discovery of trypanocidal compounds by whole cell HTS of Trypanosoma brucei. *Chem Biol Drug Des* **2006**, *67*, 355-63.
87. Sykes, M. L., Avery, V. M., A luciferase based viability assay for ATP detection in 384-well format for high throughput whole cell screening of Trypanosoma brucei brucei bloodstream form strain 427. *Parasit Vectors* **2009**, *2*, 54.
88. Tatipaka, H. B., Gillespie, J. R., Chatterjee, A. K.*, et al.*, Substituted 2-phenylimidazopyridines: a new class of drug leads for human African trypanosomiasis. *J Med Chem* **2014**, *57*, 828-35.
89. Faria, J., Moraes, C. B., Song, R.*, et al.*, Drug discovery for human African trypanosomiasis: identification of novel scaffolds by the newly developed HTS SYBR Green assay for Trypanosoma brucei. *J Biomol Screen* **2015**, *20*, 70-81.
90. Cullen, D. R., Mocerino, M., A Brief Review of Drug Discovery Research for Human African Trypanosomiasis. *Curr Med Chem* **2017**, *24*, 701-17.
91. Ferrins, L., Rahmani, R., Baell, J. B., Drug discovery and human African trypanosomiasis: a disease less neglected? *Future Med Chem* **2013**, *5*, 1801-41.
92. Scotti, L., Mendonca, F. J., da Silva, M. S.*, et al.*, Enzymatic Targets in Trypanosoma brucei. *Curr Protein Pept Sci* **2016**, *17*, 243-59.
93. Price, H. P., Menon, M. R., Panethymitaki, C.*, et al.*, Myristoyl-CoA:protein N-myristoyltransferase, an essential enzyme and potential drug target in kinetoplastid parasites. *J Biol Chem* **2003**, *278*, 7206-14.
94. Bowyer, P. W., Tate, E. W., Leatherbarrow, R. J.*, et al.*, N-myristoyltransferase: a prospective drug target for protozoan parasites. *ChemMedChem* **2008**, *3*, 402-8.
95. Brenk, R., Schipani, A., James, D.*, et al.*, Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem* **2008**, *3*, 435-44.
96. Frearson, J. A., Brand, S., McElroy, S. P.*, et al.*, N-myristoyltransferase inhibitors as new leads to treat sleeping sickness. *Nature* **2010**, *464*, 728-32.
97. Brand, S., Cleghorn, L. A., McElroy, S. P.*, et al.*, Discovery of a novel class of orally active trypanocidal N-myristoyltransferase inhibitors. *J Med Chem* **2012**, *55*, 140-52.
98. Hamilton, C. J., Saravanamuthu, A., Eggleston, I. M.*, et al.*, Ellman's-reagent-mediated regeneration of trypanothione in situ: substrate-economical microplate and time-dependent inhibition assays for trypanothione reductase. *Biochem J* **2003**, *369*, 529-37.
99. Richardson, J. L., Nett, I. R., Jones, D. C.*, et al.*, Improved tricyclic inhibitors of trypanothione reductase by screening and chemical synthesis. *ChemMedChem* **2009**, *4*, 1333-40.
100. Persch, E., Bryson, S., Todoroff, N. K.*, et al.*, Binding to large enzyme pockets: small-molecule inhibitors of trypanothione reductase. *ChemMedChem* **2014**, *9*, 1880-91.
101. Martyn, D. C., Jones, D. C., Fairlamb, A. H.*, et al.*, High-throughput screening affords novel and selective trypanothione reductase inhibitors with anti-trypanosomal activity. *Bioorg Med Chem Lett* **2007**, *17*, 1280-83.
102. Spinks, D., Torrie, L. S., Thompson, S.*, et al.*, Design, synthesis and biological evaluation of Trypanosoma brucei trypanothione synthetase inhibitors. *ChemMedChem* **2012**, *7*, 95-106.
103. Royo, S., Rodríguez, S., Schirmeister, T.*, et al.*, Dipeptidyl Enoates As Potent Rhodesain Inhibitors That Display a Dual Mode of Action. *ChemMedChem* **2015**, *10*, 1484-87.
104. Ettari, R., Tamborini, L., Angelo, I. C.*, et al.*, Inhibition of Rhodesain as a Novel Therapeutic Modality for Human African Trypanosomiasis. *J of Med Chem* **2013**, *56*, 5637-58.

105. Mallari, J. P., Shelat, A. A., Kosinski, A.*, et al.*, Structure-guided development of selective TbcatB inhibitors. *J Med Chem* **2009**, *52*, 6489-93.
106. Ochiana, S. O., Pandarinath, V., Wang, Z.*, et al.*, The human Aurora kinase inhibitor danusertib is a lead compound for anti-trypanosomal drug discovery via target repurposing. *Eur J Med Chem* **2013**, *62*, 777-84.
107. Spinks, D., Ong, H. B., Mpamhanga, C. P.*, et al.*, Design, synthesis and biological evaluation of novel inhibitors of Trypanosoma brucei pteridine reductase 1. *ChemMedChem* **2011**, *6*, 302-8.
108. Moraes, C. B., Witt, G., Kuzikov, M.*, et al.*, Accelerating Drug Discovery Efforts for Trypanosomatidic Infections Using an Integrated Transnational Academic Drug Discovery Platform. *SLAS Discov* **2019**, *24*, 346-61.
109. Dax, C., Duffieux, F., Chabot, N.*, et al.*, Selective irreversible inhibition of fructose 1,6-bisphosphate aldolase from Trypanosoma brucei. *J Med Chem* **2006**, *49*, 1499-502.
110. Ferreira, L. L. G., Ferreira, R. S., Palomino, D. L.*, et al.*, Structure-Based Virtual Screening and Biochemical Evaluation for the Identification of Novel Trypanosoma brucei Aldolase Inhibitors. *Curr Top Med Chem* **2018**, *18*, 397-405.
111. Nowicki, M. W., Tulloch, L. B., Worralll, L.*, et al.*, Design, synthesis and trypanocidal activity of lead compounds based on inhibitors of parasite glycolysis. *Bioorg Med Chem* **2008**, *16*, 5050-61.
112. Yokoyama, K., Lin, Y., Stuart, K. D.*, et al.*, Prenylation of proteins in Trypanosoma brucei. *Mol Biochem Parasit* **1997**, *87*, 61-69.
113. Yokoyama, K., Trobridge, P., Buckner, F. S.*, et al.*, The effects of protein farnesyltransferase inhibitors on trypanosomatids: inhibition of protein farnesylation and cell growth. *Mol Biochem Parasit* **1998**, *94*, 87-97.
114. Rodriguez, J. B., Falcone, B. N., Szajnman, S. H., Approaches for Designing new Potent Inhibitors of Farnesyl Pyrophosphate Synthase. *Expert Opin Drug Discov* **2016**, *11*, 307-20.
115. Martin, M. B., Grimley, J. S., Lewis, J. C.*, et al.*, Bisphosphonates Inhibit the Growth of Trypanosoma brucei, Trypanosoma cruzi, Leishmania donovani, Toxoplasma gondii, and Plasmodium falciparum: A Potential Route to Chemotherapy. *J Med Chem* **2001**, *44*, 909-16.
116. Martin, M. B., Sanders, J. M., Kendrick, H.*, et al.*, Activity of Bisphosphonates against Trypanosoma brucei rhodesiense. *J Med Chem* **2002**, *45*, 2904-14.
117. Jordao, F. M., Saito, A. Y., Miguel, D. C.*, et al.*, In vitro and in vivo antiplasmodial activities of risedronate and its interference with protein prenylation in Plasmodium falciparum. *Antimicrob Agents Chemother* **2011**, *55*, 2026-31.
118. Demoro, B., Caruso, F., Rossi, M.*, et al.*, Risedronate metal complexes potentially active against Chagas disease. *J Inorg Biochem* **2010**, *104*, 1252-8.
119. Arafa, R. K., Wenzler, T., Brun, R.*, et al.*, Molecular modeling study and synthesis of novel dicationic flexible triaryl guanidines and imidamides as antiprotozoal agents. *Eur J Med Chem* **2011**, *46*, 5852-60.
120. Huang, T. L., Vanden Eynde, J. J., Mayence, A.*, et al.*, Synthesis and SAR of alkanediamide-linked bisbenzamidines with anti-trypanosomal and anti-pneumocystis activity. *Bioorg Med Chem Lett* **2009**, *19*, 5884-6.
121. Bakunov, S. A., Bakunova, S. M., Bridges, A. S.*, et al.*, Synthesis and antiprotozoal properties of pentamidine congeners bearing the benzofuran motif. *J Med Chem* **2009**, *52*, 5763-7.
122. Bakunova, S. M., Bakunov, S. A., Patrick, D. A.*, et al.*, Structure−Activity Study of Pentamidine Analogues as Antiprotozoal Agents. *J Med Chem* **2009**, *52*, 2016-35.
123. Yang, G., Zhu, W., Wang, Y.*, et al.*, In Vitro and in Vivo Activity of Multitarget Inhibitors against Trypanosoma brucei. *ACS Infect Dis* **2015**, *1*, 388-98.
124. Paine, M. F., Wang, M. Z., Generaux, C. N.*, et al.*, Diamidines for human African trypanosomiasis. *Curr Opin Invest Dr* **2010**, *11*, 876-83.
125. Wenzler, T., Yang, S., Braissant, O.*, et al.*, Pharmacokinetics, Trypanosoma brucei gambiense efficacy, and time of drug action of DB829, a preclinical candidate for treatment of second-stage human African trypanosomiasis. *Antimicrob Agents Chemother* **2013**, *57*, 5330-43.
126. Wenzler, T., Yang, S., Patrick, D. A.*, et al.*, In vitro and in vivo evaluation of 28DAP010, a novel diamidine for treatment of second-stage African sleeping sickness. *Antimicrob Agents Chemother* **2014**, *58*, 4452-63.
127. Thuita, J. K. Biological and pharmacological investigations of novel diamidines in animal models of human African trypanosomiasis. Basel, Switzerland, 2013.
128. Walochnik, J., Duchêne, M., From Molecule to Drug. In *Molecular Parasitology: Protozoan Parasites and their Molecules*, Walochnik, J.; Duchêne, M., Eds. Springer: 2016.

129. Tarral, A., Blesson, S., Mordt, O. V.*, et al.*, Determination of an optimal dosing regimen for fexinidazole, a novel oral drug for the treatment of human African trypanosomiasis: first-in-human studies. *Clin Pharmacokinet* **2014**, *53*, 565-80.
130. Raether, W., Seidenath, H., The activity of fexinidazole (HOE 239) against experimental infections with Trypanosoma cruzi, trichomonads and Entamoeba histolytica. *Ann Trop Med Parasitol* **1983**, *77*, 13-26.
131. Mesu, V. K. B. K., Kalonji, W. M., Bardonneau, C.*, et al.*, Oral fexinidazole for late-stage African Trypanosoma brucei gambiense trypanosomiasis: a pivotal multicentre, randomised, non-inferiority trial. *The Lancet* **2018**, *391*, 144-154.
132. DNDi, Sanofi, DNDi seek European Medicines Agency review for sleeping sickness treatment. https://www.dndi.org/2018/media-centre/press-releases/sanofi-dndi-seek-ema-review-for-sleeping-sickness-treatment/ (accessed 7th January 2019).
133. DNDi, European Medicines Agency recommends fexinidazole, the first all-oral treatment for sleeping sickness. https://www.dndi.org/2018/media-centre/press-releases/ema-recommends-fexinidazole-first-all-oral-treatment-sleeping-sickness/ (accessed 7th January 2019).
134. DNDi, Fexinidazole (HAT). https://www.dndi.org/diseases-projects/portfolio/fexinidazole (accessed 3rd of March 2019).
135. Ding, D., Zhao, Y., Meng, Q.*, et al.*, Discovery of novel benzoxaborole-based potent antitrypanosomal agents. *ACS Med Chem Lett* **2010**, *1*, 165-9.
136. Eperon, G., Balasegaram, M., Potet, J.*, et al.*, Treatment options for second-stage gambiense human African trypanosomiasis. *Expert Rev Anti-Infect Therap* **2014**, *12*, 1407-1417.
137. DNDi, SCYX-1330682 & SCYX-1608210 Oxaboroles. https://www.dndi.org/diseases-projects/portfolio/scyx-1330682/ (accessed 3rd March 2019).
138. DNDi, Pivotal clinical trial to begin for first oral drug candidate specifically developed for sleeping sickness. https://www.dndi.org/2015/media-centre/press-releases/pr-scyx-7158/.
139. DNDi, Acoziborole. https://www.dndi.org/diseases-projects/portfolio/scyx-7158/ (accessed 7th January 2019).
140. Wei, J. H., Yin, X., Welander, P. V., Sterol Synthesis in Diverse Bacteria. *Front Microbiol* **2016**, *7*, 990-990.
141. Rodrigues, M. L., The Multifunctional Fungal Ergosterol. *MBio* **2018**, *9*, e01755-18.
142. Joffrion, T. M., Cushion, M. T., Sterol biosynthesis and sterol uptake in the fungal pathogen Pneumocystis carinii. *FEMS Microbiol Lett* **2010**, *311*, 1-9.
143. Desmond, E., Gribaldo, S., Phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature. *Genome Biol Evol* **2009**, *1*, 364-81.
144. Goldstein, J. L., Brown, M. S., Regulation of the mevalonate pathway. *Nature* **1990**, *343*, 425-30.
145. Berg , J. M., Tymoczko, J. L., Stryer, L., Important Derivatives of Cholesterol Include Bile Salts and Steroid Hormones. In *Biochemistry. 5th edition*, W H Freeman: 2002.
146. Ikonen, E., Cellular cholesterol trafficking and compartmentalization. *Nat Rev Mol Cell Biol* **2008**, *9*, 125-38.
147. Wang, Y., Hekimi, S., Understanding Ubiquinone. *Trends Cell Biol* **2016**, *26*, 367-378.
148. Chojnacki, T., Dallner, G., The biological role of dolichol. *Biochemical J* **1988**, *251*, 1-9.
149. Zhang, F. L., Casey, P. J., Protein prenylation: molecular mechanisms and functional consequences. *Annu Rev Biochem* **1996**, *65*, 241-69.
150. Wang, M., Casey, P. J., Protein prenylation: unique fats make their mark on biology. *Nat Rev Mol Cell Biol* **2016**, *17*, 110-22.
151. Hartmann, M., Plant sterols and the membrane environment. *Trends in Plant Science* **1998**, *3*, 170-75.
152. Dhingra, S., Cramer, R. A., Regulation of Sterol Biosynthesis in the Human Fungal Pathogen Aspergillus fumigatus: Opportunities for Therapeutic Development. *Front Microbiol* **2017**, *8*, 92.
153. de Souza, W., Rodrigues, J. C., Sterol Biosynthesis Pathway as Target for Anti-trypanosomatid Drugs. *Interdiscip Perspect Infect Dis* **2009**, *2009*, 642502.
154. Heyden, N., Docampo, R., Significant Differences Between Procyclic and Bloodstream Forms of Trypanosoma brucei in the Maintenance of their Plasma Membrane Potential. *J Eukaryot Microbiol* **2002**, *49*, 407-13.
155. Nes, C. R., Singha, U. K., Liu, J.*, et al.*, Novel sterol metabolic network of Trypanosoma brucei procyclic and bloodstream forms. *Biochem J* **2012**, *443*, 267-77.
156. Millerioux, Y., Mazet, M., Bouyssou, G.*, et al.*, De novo biosynthesis of sterols and fatty acids in the Trypanosoma brucei procyclic form: Carbon source preferences and metabolic flux redistributions. *PLoS Pathog* **2018**, *14*, e1007116.

157. Mazet, M., Harijan, R. K., Kiema, T. R.*, et al.*, The characterization and evolutionary relationships of a trypanosomal thiolase. *Int J Parasitol* **2011**, *41*, 1273-83.
158. Harijan, R. K., Kiema, T. R., Karjalainen, M. P.*, et al.*, Crystal structures of SCP2-thiolases of Trypanosomatidae, human pathogens causing widespread tropical diseases: the importance for catalysis of the cysteine of the unique HDCF loop. *Biochem J* **2013**, *455*, 119-30.
159. Misra, I., Wang, C. Z., Miziorko, H. M., The influence of conserved aromatic residues in 3-hydroxy-3-methylglutaryl-CoA synthase. *J Biol Chem* **2003**, *278*, 26443-9.
160. Carrero-Lerida, J., Perez-Moreno, G., Castillo-Acosta, V. M.*, et al.*, Intracellular location of the early steps of the isoprenoid biosynthetic pathway in the trypanosomatids Leishmania major and Trypanosoma brucei. *Int J Parasitol* **2009**, *39*, 307-14.
161. Coppens, I., Courtoy, P. J., The mevalonate pathway in parasitic protozoa and helminths. *Exp Parasitol* **1996**, *82*, 76-85.
162. Heise, N., Opperdoes, F. R., Localisation of a 3-hydroxy-3-methylglutaryl-coenzyme A reductase in the mitochondrial matrix of Trypanosoma brucei procyclics. *Z Naturforsch C* **2000**, *55*, 473-7.
163. Pena-Diaz, J., Montalvetti, A., Flores, C. L.*, et al.*, Mitochondrial localization of the mevalonate pathway enzyme 3-Hydroxy-3-methyl-glutaryl-CoA reductase in the Trypanosomatidae. *Mol Biol Cell* **2004**, *15*, 1356-63.
164. Cosentino, R. O., Agüero, F., Genetic Profiling of the Isoprenoid and Sterol Biosynthesis Pathway Genes of Trypanosoma cruzi. *PLoS One* **2014**, *9*, e96762.
165. Miziorko, H. M., Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch Biochem Biophys* **2011**, *505*, 131-43.
166. Subramani, S., Protein import into peroxisomes and biogenesis of the organelle. *Annu Rev Cell Biol* **1993**, *9*, 445-78.
167. Ferreira, E. R., Horjales, E., Bonfim-Melo, A.*, et al.*, Unique behavior of Trypanosoma cruzi mevalonate kinase: A conserved glycosomal enzyme involved in host cell invasion and signaling. *Sci Rep* **2016**, *6*, 24610.
168. Colasante, C., Ellis, M., Ruppert, T.*, et al.*, Comparative proteomics of glycosomes from bloodstream form and procyclic culture form Trypanosoma brucei brucei. *Proteomics* **2006**, *6*, 3275-93.
169. Byres, E., Alphey, M., K Smith, T.*, et al.*, Crystal Structures of Trypanosoma brucei and Staphylococcus aureus Mevalonate Diphosphate Decarboxylase Inform on the Determinants of Specificity and Reactivity. *J Mol Biol* **2007**, *371*, 540-53.
170. Buhaescu, I., Izzedine, H., Mevalonate pathway: a review of clinical and therapeutical implications. *Clin Biochem* **2007**, *40*, 575-84.
171. Byres, E., Martin, D. M., Hunter, W. N., A preliminary crystallographic analysis of the putative mevalonate diphosphate decarboxylase from Trypanosoma brucei. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2005**, *61*, 581-4.
172. Gao, J., Chu, X., Qiu, Y.*, et al.*, Discovery of potent inhibitor for farnesyl pyrophosphate synthase in the mevalonate pathway. *Chem Comm* **2010**, *46*, 5340-42.
173. Ferella, M., Li, Z.-H., Andersson, B.*, et al.*, Farnesyl diphosphate synthase localizes to the cytoplasm of Trypanosoma cruzi and T. brucei. *Exp Parasitol* **2008**, *119*, 308-12.
174. Buckner, F. S., Yokoyama, K., Nguyen, L.*, et al.*, Cloning, heterologous expression, and distinct substrate specificity of protein farnesyltransferase from Trypanosoma brucei. *J Biol Chem* **2000**, *275*, 21870-6.
175. Low, P., Dallner, G., Mayor, S.*, et al.*, The mevalonate pathway in the bloodstream form of Trypanosoma brucei. Identification of dolichols containing 11 and 12 isoprene residues. *J Biol Chem* **1991**, *266*, 19250-7.
176. Lai, D. H., Poropat, E., Pravia, C.*, et al.*, Solanesyl diphosphate synthase, an enzyme of the ubiquinone synthetic pathway, is required throughout the life cycle of Trypanosoma brucei. *Eukaryot Cell* **2014**, *13*, 320-8.
177. Haubrich, B. A., Microbial Sterolomics as a Chemical Biology Tool. *Molecules* **2018**, *23*, 2768.
178. Perez-Moreno, G., Sealey-Cardona, M., Rodrigues-Poveda, C.*, et al.*, Endogenous sterol biosynthesis is important for mitochondrial function and cell morphology in procyclic forms of Trypanosoma brucei. *Int J Parasitol* **2012**, *42*, 975-89.
179. Berg, J. M., Stryer, L., Tymoczko, J. L., *Biochemie*. Springer Spektrum, 2013.
180. Abe, I., Rohmer, M., Prestwich, G. D., Enzymatic Cyclization of Squalene and Oxidosqualene to Sterols and Triterpenes. *Chemical Reviews* **1993**, *93*, 2189-206.
181. Buckner, F. S., Nguyen, L. N., Joubert, B. M.*, et al.*, Cloning and heterologous expression of the Trypanosoma brucei lanosterol synthase gene. *Mol Biochem Parasitol* **2000**, *110*, 399-403.

182. Zhou, W., Cross, G. A., Nes, W. D., Cholesterol import fails to prevent catalyst-based inhibition of ergosterol synthesis and cell proliferation of Trypanosoma brucei. *J Lipid Res* **2007**, *48*, 665-73.
183. Leaver, D. J., Synthesis and Biological Activity of Sterol 14alpha-Demethylase and Sterol C24-Methyltransferase Inhibitors. *Molecules* **2018**, *23*, 1753.
184. Chen, C. K., Leung, S. S., Guilbert, C.*, et al.*, Structural characterization of CYP51 from Trypanosoma cruzi and Trypanosoma brucei bound to the antifungal drugs posaconazole and fluconazole. *PLoS Negl Trop Dis* **2010**, *4*, e651.
185. Lepesheva, G. I., Park, H. W., Hargrove, T. Y.*, et al.*, Crystal structures of Trypanosoma brucei sterol 14alpha-demethylase and implications for selective treatment of human infections. *J Biol Chem* **2010**, *285*, 1773-80.
186. Yao, C., Wilson, M. E., Dynamics of sterol synthesis during development of Leishmania spp. parasites to their virulent form. *Parasit Vectors* **2016**, *9*, 200.
187. Hargrove, T. Y., Kim, K., de Nazare Correia Soeiro, M.*, et al.*, CYP51 structures and structure-based development of novel, pathogen-specific inhibitory scaffolds. *Int J Parasitol Drugs Drug Resist* **2012**, *2*, 178-86.
188. Haubrich, B. A., Singha, U. K., Miller, M. B.*, et al.*, Discovery of an ergosterol-signaling factor that regulates Trypanosoma brucei growth. *Journal of Lipid Research* **2015**, *56*, 331-41.
189. Gros, L., Castillo-Acosta, V. M., Jimenez Jimenez, C.*, et al.*, New azasterols against Trypanosoma brucei: role of 24-sterol methyltransferase in inhibitor action. *Antimicrob Agents Chemother* **2006**, *50*, 2595-601.
190. Coppens, I., Courtoy, P. J., The adaptative mechanisms of Trypanosoma brucei for sterol homeostasis in its different life-cycle environments. *Annu Rev Microbiol* **2000**, *54*, 129-56.
191. Coppens, I., Baudhuin, P., Opperdoes, F. R.*, et al.*, Receptors for the host low density lipoproteins on the hemoflagellate Trypanosoma brucei: purification and involvement in the growth of the parasite. *Proc Natl Acad Sci U S A* **1988**, *85*, 6753-7.
192. Coppens, I., Bastin, P., Levade, T.*, et al.*, Activity, pharmacological inhibition and biological regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in Trypanosoma brucei. *Mol Biochem Parasitol* **1995**, *69*, 29-40.
193. Lepesheva, G. I., Ott, R. D., Hargrove, T. Y.*, et al.*, Sterol 14alpha-demethylase as a potential target for antitrypanosomal therapy: enzyme inhibition and parasite cell growth. *Chem Biol* **2007**, *14*, 1283-93.
194. Montalvetti, A., Fernandez, A., Sanders, J. M.*, et al.*, Farnesyl pyrophosphate synthase is an essential enzyme in Trypanosoma brucei. In vitro RNA interference and in vivo inhibition studies. *J Biol Chem* **2003**, *278*, 17075-83.
195. Rodriguez, R. J., Low, C., Bottema, C. D.*, et al.*, Multiple functions for sterols in Saccharomyces cerevisiae. *Biochim Biophys Acta* **1985**, *837*, 336-43.
196. Aguilar, P. S., Heiman, M. G., Walther, T. C.*, et al.*, Structure of sterol aliphatic chains affects yeast cell shape and cell fusion during mating. *Proc Natl Acad Sci U S A* **2010**, *107*, 4170-5.
197. Istvan, E. S., Deisenhofer, J., Structural mechanism for statin inhibition of HMG-CoA reductase. *Science* **2001**, *292*, 1160-4.
198. Istvan, E., Statin inhibition of HMG-CoA reductase: a 3-dimensional view. *Atheroscler Suppl* **2003**, *4*, 3-8.
199. Davies, J. T., Delfino, S. F., Feinberg, C. E.*, et al.*, Current and Emerging Uses of Statins in Clinical Therapeutics: A Review. *Lipid Insights* **2016**, *9*, 13-29.
200. Markovic-Plese, S., Singh, A. K., Singh, I., Therapeutic potential of statins in multiple sclerosis: immune modulation, neuroprotection and neurorepair. *Future Neurol* **2008**, *3*, 153.
201. Weber, M. S., Youssef, S., Dunn, S. E.*, et al.*, Statins in the treatment of central nervous system autoimmune disease. *J Neuroimmunol* **2006**, *178*, 140-8.
202. Chow, S. C., Immunomodulation by statins: mechanisms and potential impact on autoimmune diseases. *Arch Immunol Ther Exp (Warsz)* **2009**, *57*, 243-51.
203. Andersson, M., Low, P., Bakhiet, M., Lovastatin inhibits interferon-gamma-induced Trypanosoma brucei brucei proliferation: evidence for mevalonate pathway involvement. *J Interferon Cytokine Res* **1996**, *16*, 435-9.
204. Field, H., Blench, I., Croft, S.*, et al.*, Characterisation of protein isoprenylation in procyclic form Trypanosoma brucei. *Mol Biochem Parasit* **1996**, *82*, 67-80.
205. Field, M. C., Signalling the genome: the Ras-like small GTPase family of trypanosomatids. *Trends Parasitol* **2005**, *21*, 447-50.
206. Qian, Y. M., Sebti, S. M., Hamilton, A. D., Farnesyltransferase as a target for anticancer drug design. *Biopolymers* **1997**, *43*, 25-41.

207. Fiordalisi, J. J., Cox, A. D., Chapter 222 - Farnesyltransferase Inhibitors. In *Handbook of Cell Signaling (Second Edition)*, Bradshaw, R. A.; Dennis, E. A., Eds. Academic Press: San Diego, 2010.
208. Ayral-Kaloustian, S., Salaski, E., Protein Farnesyltransferase Inhibitors. *Curr Med Chem* **2002**, *9*, 1003-32.
209. Yokoyama, K., Trobridge, P., Buckner, F. S.*, et al.*, Protein farnesyltransferase from Trypanosoma brucei - A heterodimer of 61- and 65-kDa subunits as a new target for antiparasite therapeutics. *Journal of Biological Chemistry* **1998**, *273*, 26497-505.
210. Clerici, F., Gelmi, M. L., Yokoyama, K.*, et al.*, Isothiazole dioxides: synthesis and inhibition of Trypanosoma brucei protein farnesyltransferase. *Bioorg Med Chem Lett* **2002**, *12*, 2217-20.
211. Gelb, M. H., Van Voorhis, W. C., Buckner, F. S.*, et al.*, Protein farnesyl and N-myristoyl transferases: piggy-back medicinal chemistry targets for the development of antitrypanosomatid and antimalarial therapeutics. *Mol Biochem Parasit* **2003**, *126*, 155-163.
212. Kourounakis, A. P., Katselou, M. G., Matralis, A. N.*, et al.*, Squalene synthase inhibitors: An update on the search for new antihyperlipidemic and antiatherosclerotic agents. *Curr Med Chem* **2011**, *18*, 4418-39.
213. Ebihara, T., Teshima, K., Kondo, T.*, et al.*, Pharmacokinetics of TAK-475, a Squalene Synthase Inhibitor, in Rats and Dogs. *Drug Res* **2016**, *66*, 287-92.
214. Lorente, S. O., Gomez, R., Jimenez, C.*, et al.*, Biphenylquinuclidines as inhibitors of squalene synthase and growth of parasitic protozoa. *Bioorgan Med Chem* **2005**, *13*, 3519-29.
215. Cammerer, S. B., Jimenez, C., Jones, S.*, et al.*, Quinuclidine derivatives as potential antiparasitics. *Antimicrob Agents Chemother* **2007**, *51*, 4049-61.
216. Pierce, H. D., Pierce, A. M., Srinivasan, R.*, et al.*, Azasterol inhibitors in yeast. *Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism* **1978**, *529*, 429-437.
217. Magaraci, F., Jimenez, C. J., Rodrigues, C.*, et al.*, Azasterols as inhibitors of sterol 24-methyltransferase in Leishmania species and Trypanosoma cruzi. *J Med Chem* **2003**, *46*, 4714-27.
218. Urbina, J. A., Vivas, J., Visbal, G.*, et al.*, Modification of the sterol composition of Trypanosoma (Schizotrypanum) cruzi epimastigotes by Δ24(25)-sterol methyl transferase inhibitors and their combinations with ketoconazole. *Mol Biochem Parasit* **1995**, *73*, 199-210.
219. Urbina, J. A., Visbal, G., Contreras, L. M.*, et al.*, Inhibitors of delta24(25) sterol methyltransferase block sterol synthesis and cell proliferation in Pneumocystis carinii. *Antimicrob Agents Chemother* **1997**, *41*, 1428-32.
220. Lorente, S. O., Rodrigues, J. C., Jimenez Jimenez, C.*, et al.*, Novel azasterols as potential agents for treatment of leishmaniasis and trypanosomiasis. *Antimicrob Agents Chemother* **2004**, *48*, 2937-50.
221. Gigante, F., Kaiser, M., Brun, R.*, et al.*, SAR studies on azasterols as potential anti-trypanosomal and anti-leishmanial agents. *Bioorg Med Chem* **2009**, *17*, 5950-61.
222. Leaver, D. J., Patkar, P., Singha, U. K.*, et al.*, Fluorinated Sterols Are Suicide Inhibitors of Ergosterol Biosynthesis and Growth in Trypanosoma brucei. *Chem Biol* **2015**, *22*, 1374-83.
223. Miller, M. B., Patkar, P., Singha, U. K.*, et al.*, 24-Methylenecyclopropane steroidal inhibitors: A Trojan horse in ergosterol biosynthesis that prevents growth of Trypanosoma brucei. *Biochim Biophys Acta Mol Cell Biol Lipids* **2017**, *1862*, 305-13.
224. Lynen F, A. B., Eggerer H, et al., y,y-Dimethyl-ally-pyrophosphat und Geranyl-pyrophosphat, biologische Vorstufen des Squalens VI1). *Angew Chem* **1959**, *71*, 657-63.
225. Sun, S., McKenna, C. E., Farnesyl pyrophosphate synthase modulators: a patent review (2006 - 2010). *Expert Opin Ther Pat* **2011**, *21*, 1433-51.
226. Tarshis, L. C., Yan, M., Poulter, C. D.*, et al.*, Crystal Structure of Recombinant Farnesyl Diphosphate Synthase at 2.6-.ANG. Resolution. *Biochemistry* **1994**, *33*, 10871-77.
227. Gabelli, S. B., McLellan, J. S., Montalvetti, A.*, et al.*, Structure and mechanism of the farnesyl diphosphate synthase from Trypanosoma cruzi: implications for drug design. *Proteins* **2006**, *62*, 80-8.
228. Schmidberger, J. W., Schnell, R., Schneider, G., Structural characterization of substrate and inhibitor binding to farnesyl pyrophosphate synthase from Pseudomonas aeruginosa. *Acta Crystallogr D Biol Crystallogr* **2015**, *71*, 721-31.
229. Mao, J., Gao, Y. G., Odeh, S.*, et al.*, Crystallization and preliminary X-ray diffraction study of the farnesyl diphosphate synthase from Trypanosoma brucei. *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 1863-6.
230. Hosfield, D. J., Zhang, Y., Dougan, D. R.*, et al.*, Structural basis for bisphosphonate-mediated inhibition of isoprenoid biosynthesis. *J Biol Chem* **2004**, *279*, 8526-9.
231. Rondeau, J. M., Bitsch, F., Bourgier, E.*, et al.*, Structural basis for the exceptional in vivo efficacy of bisphosphonate drugs. *ChemMedChem* **2006**, *1*, 267-73.

232. Kavanagh, K. L., The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7829-34.
233. Joly, A., Edwards, P. A., Effect of Site-Directed Mutagenesis of Conserved Aspartate and Arginine Residues Upon Farnesyl Diphosphate Synthase Activity. *J Biol Chem* **1993**, *268*, 26983-89.
234. Sanchez, V. M., Crespo, A., Gutkind, J. S.*, et al.*, Investigation of the Catalytic Mechanism of Farnesyl Pyrophosphate Synthase by Computer Simulation. *J Phys Chem B* **2006**, *110*, 18052-57.
235. Marrero, P. F., Poulter, C. D., Edwards, P. A., Effects of site-directed mutagenesis of the highly conserved aspartate residues in domain II of farnesyl diphosphate synthase activity. *J Biol Chem* **1992**, *267*, 21873-8.
236. Aaron, J. A., Christianson, D. W., Trinuclear Metal Clusters in Catalysis by Terpenoid Synthases. *Pure Appl Chem* **2010**, *82*, 1585-97.
237. Tsoumpra, M. K., Muniz, J. R., Barnett, B. L.*, et al.*, The inhibition of human farnesyl pyrophosphate synthase by nitrogen-containing bisphosphonates. Elucidating the role of active site threonine 201 and tyrosine 204 residues using enzyme mutants. *Bone* **2015**, *81*, 478-86.
238. Cornforth, J. W., Cornforth, R. H., Popjak, G.*, et al.*, Studies on the biosynthesis of cholesterol. XX. Steric course of decarboxylation of 5-pyrophosphomevalonate and of the carbon to carbon bond formation in the biosynthesis of farnesyl pyrophosphate. *J Biol Chem* **1966**, *241*, 3970-87.
239. Ramamoorthy, G., Pugh, M. L., Tian, B. X.*, et al.*, Synthesis and enzymatic studies of bisubstrate analogues for farnesyl diphosphate synthase. *J Org Chem* **2015**, *80*, 3902-13.
240. Tarshis, L. C., Proteau, P. J., Kellogg, B. A.*, et al.*, Regulation of product chain length by isoprenyl diphosphate synthases. *Proc Natl Acad Sci U S A* **1996**, *93*, 15018-23.
241. Menschutkin, N., Ueber die Einwirkung des Chloracetyls auf phosphorige Säure. *Annalen der Chemie und Pharmacie* **1865**, *133*, 317-20.
242. Fleisch, H., Russell, R. G., Francis, M. D., Diphosphonates inhibit hydroxyapatite dissolution in vitro and bone resorption in tissue culture and in vivo. *Science* **1969**, *165*, 1262-4.
243. Russell, R. G. G., Mühlbauer, R. C., Bisaz, S.*, et al.*, The influence of pyrophosphate, condensed phosphates, phosphonates and other phosphate compounds on the dissolution of hydroxyapatitein vitro and on bone resorption induced by parathyroid hormone in tissue culture and in thyroparathyroidectomised rats. *Calcified Tissue Research* **1970**, *6*, 183-96.
244. Watts, N. B., Diab, D. L., Long-term use of bisphosphonates in osteoporosis. *J Clin Endocrinol Metab* **2010**, *95*, 1555-65.
245. Reid, I. R., Hosking, D. J., Bisphosphonates in Paget's disease. *Bone* **2011**, *49*, 89-94.
246. Geusens, P., McClung, M., Review of risedronate in the treatment of osteoporosis. *Expert Opin Pharmacother* **2001**, *2*, 2011-25.
247. Drake, M. T., Clarke, B. L., Khosla, S., Bisphosphonates: mechanism of action and role in clinical practice. *Mayo Clin Proc* **2008**, *83*, 1032-45.
248. Toller, C. S., Charlesworth, S., Mihalyo, M.*, et al.*, Bisphosphonates: AHFS 92:24. *J Pain Symptom Manage* **2019**, *57*, 1018-1030.
249. Hanley, D. A., Ioannidis, G., Adachi, J. D., Etidronate therapy in the treatment and prevention of osteoporosis. *J Clin Densitom* **2000**, *3*, 79-95.
250. Frediani, B., Baraldi, E., Cremonesi, G., Effect of clodronate treatment on risk of fracture: a systematic review and meta-analysis. *Calcif Tissue Int* **2014**, *95*, 295-307.
251. Reginster, J. Y., Christiansen, C., Roux, C.*, et al.*, Intermittent cyclic tiludronate in the treatment of osteoporosis. *Osteoporos Int* **2001**, *12*, 169-77.
252. Coukell, A. J., Markham, A., Pamidronate. A review of its use in the management of osteolytic bone metastases, tumour-induced hypercalcaemia and Paget's disease of bone. *Drugs Aging* **1998**, *12*, 149-68.
253. John Yates, A., Rodan, G. A., Alendronate and osteoporosis. *Drug Discov Today* **1998**, *3*, 69-78.
254. Frampton, J. E., Perry, C. M., Ibandronate: a review of its use in the management of postmenopausal osteoporosis. *Drugs* **2008**, *68*, 2683-707.
255. Dhillon, S., Zoledronic Acid (Reclast®, Aclasta®): A Review in Osteoporosis. *Drugs* **2016**, *76*, 1683-97.
256. Fernandez, D., Ramis, R., Ortega-Castro, J.*, et al.*, New insights into human farnesyl pyrophosphate synthase inhibition by second-generation bisphosphonate drugs. *J Comput Aided Mol Des* **2017**, *31*, 675-688.
257. Lewiecki, E. M., Bisphosphonates for the treatment of osteoporosis: insights for clinicians. *Therapeutic advances in chronic disease* **2010**, *1*, 115-128.
258. Russell, R. G., Bisphosphonates: from bench to bedside. *Ann N Y Acad Sci* **2006**, *1068*, 367-401.
259. Muhlbauer, R. C., Bauss, F., Schenk, R.*, et al.*, BM 21.0955, a potent new bisphosphonate to inhibit bone resorption. *J Bone Miner Res* **1991**, *6*, 1003-11.

260. Widler, L., Jaeggi, K. A., Green, J. R., Amino-substituted GEM-bisphosphonates. *Phosphorus Sulfur* **1999**, *144*, 5-8.
261. Russell, R. G. G., Watts, N. B., Ebetino, F. H.*, et al.*, Mechanisms of action of bisphosphonates: similarities and differences and their potential influence on clinical efficacy. *Osteoporosis Int* **2008**, *19*, 733-759.
262. Ebetino, F. H., The discovery of the bone-active agent risedronate, and bisphosphonate structure-activity considerations including the aminophenylethane phosphonate series. *Phosphorus Sulfur* **1999**, *144*, 9-12.
263. Ebetino, F. H., Hogan, A. M., Sun, S.*, et al.*, The relationship between the chemistry and biological activity of the bisphosphonates. *Bone* **2011**, *49*, 20-33.
264. Dunford, J. E., Thompson, K., Coxon, F. P.*, et al.*, Structure-Activity Relationships for Inhibition of Farnesyl Diphosphate Synthase in Vitro and Inhibition of Bone Resorption in Vivo by Nitrogen-Containing Bisphosphonates. *J Pharm Exp Ther* **2001**, *296*, 235-242.
265. Frith, J. C., Monkkonen, J., Auriola, S.*, et al.*, The molecular mechanism of action of the antiresorptive and antiinflammatory drug clodronate - Evidence for the formation in vivo of a metabolite that inhibits bone resorption and causes osteoclast and macrophage apoptosis. *Arthritis Rheum* **2001**, *44*, 2201-10.
266. Luckman, S. P., Hughes, D. E., Coxon, F. P.*, et al.*, Nitrogen-containing bisphosphonates inhibit the mevalonate pathway and prevent post-translational prenylation of GTP-binding proteins, including Ras. *J Bone Miner Res* **1998**, *13*, 581-9.
267. van Beek, E., Lowik, C., van der Pluijm, G.*, et al.*, The role of geranylgeranylation in bone resorption and its suppression by bisphosphonates in fetal bone explants in vitro: A clue to the mechanism of action of nitrogen-containing bisphosphonates. *J Bone Miner Res* **1999**, *14*, 722-9.
268. van Beek, E., Pieterman, E., Cohen, L.*, et al.*, Nitrogen-containing bisphosphonates inhibit isopentenyl pyrophosphate isomerase/farnesyl pyrophosphate synthase activity with relative potencies corresponding to their antiresorptive potencies in vitro and in vivo. *Biochem Biophys Res Commun* **1999**, *255*, 491-4.
269. van Beek, E., Pieterman, E., Cohen, L.*, et al.*, Farnesyl Pyrophosphate Synthase Is the Molecular Target of Nitrogen-Containing Bisphosphonates. *Biochem Biophys Res Commun* **1999**, *264*, 108-11.
270. Rogers, M. J., New insights into the molecular mechanisms of action of bisphosphonates. *Curr. Pharm. Des.* **2003**, *9*, 2643-58.
271. Keller, R. K., Fliesler, S. J., Mechanism of Aminobisphosphonate Action: Characterization of Alendronate Inhibition of the Isoprenoid Pathway. *Biochem Biophys Res Commun* **1999**, *266*, 560-563.
272. Martin, M. B., Arnold, W., Heath, H. T., 3rd*, et al.*, Nitrogen-containing bisphosphonates as carbocation transition state analogs for isoprenoid biosynthesis. *Biochem Biophys Res Commun* **1999**, *263*, 754-8.
273. Coleman, R. E., Bisphosphonates: clinical experience. *Oncologist* **2004**, *9 Suppl 4*, 14-27.
274. Boonen, S., Bisphosphonate efficacy and clinical trials for postmenopausal osteoporosis: Similarities and differences. *Bone* **2007**, *40*, 26-31.
275. Traina, A. N., Kane, M. P., Bakst, G.*, et al.*, Pharmacist-run zoledronic acid clinic. *Am J Health Syst Pharm* **2011**, *68*, 1399-403.
276. Stresing, V., Daubine, F., Benzaid, I.*, et al.*, Bisphosphonates in cancer therapy. *Cancer Lett* **2007**, *257*, 16-35.
277. Green, J., Lipton, A., Anticancer properties of zoledronic acid. *Cancer Invest* **2010**, *28*, 944-57.
278. Tamura, T., Shomori, K., Nakabayashi, M.*, et al.*, Zoledronic acid, a third-generation bisphosphonate, inhibits cellular growth and induces apoptosis in oral carcinoma cell lines. *Oncol Rep* **2011**, *25*, 1139-43.
279. Van Acker, H. H., Anguille, S., Willemen, Y.*, et al.*, Bisphosphonates for cancer treatment: Mechanisms of action and lessons from clinical trials. *Pharmacol Ther* **2016**, *158*, 24-40.
280. De Schutter, J. W., Park, J., Leung, C. Y.*, et al.*, Multistage screening reveals chameleon ligands of the human farnesyl pyrophosphate synthase: implications to drug discovery for neurodegenerative diseases. *J Med Chem* **2014**, *57*, 5764-76.
281. Zameer, S., Najmi, A. K., Vohora, D.*, et al.*, Bisphosphonates: Future perspective for neurological disorders. *Pharmacol Rep* **2018**, *70*, 900-907.
282. Urbina, J. A., Moreno, B., Vierkotter, S.*, et al.*, Trypanosoma cruzi contains major pyrophosphate stores, and its growth in vitro and in vivo is blocked by pyrophosphate analogs. *J Biolog Chem* **1999**, *274*, 33609-15.

283. Huang, C. H., Gabelli, S. B., Oldfield, E.*, et al.*, Binding of nitrogen-containing bisphosphonates (N-BPs) to the Trypanosoma cruzi farnesyl diphosphate synthase homodimer. *Proteins* **2010**, *78*, 888-99.
284. Garzoni, L. R., Waghabi, M. C., Baptista, M. M.*, et al.*, Antiparasitic activity of risedronate in a murine model of acute Chagas' disease. *Int J Antimicrob Agents* **2004**, *23*, 286-290.
285. Garzoni, L. R., Caldera, A., Meirelles Mde, N.*, et al.*, Selective in vitro effects of the farnesyl pyrophosphate synthase inhibitor risedronate on Trypanosoma cruzi. *Int J Antimicrob Agents* **2004**, *23*, 273-85.
286. Szajnman, S. H., Bailey, B. N., Docampo, R.*, et al.*, Bisphosphonates derived from fatty acids are potent growth inhibitors of Trypanosoma cruzi. *Bioorg Med Chem Lett* **2001**, *11*, 789-92.
287. Szajnman, S. H., Ravaschino, E. L., Docampo, R.*, et al.*, Synthesis and biological evaluation of 1-amino-1,1-bisphosphonates derived from fatty acids against Trypanosoma cruzi targeting farnesyl pyrophosphate synthase. *Bioorg Med Chem Lett* **2005**, *15*, 4685-90.
288. Moreno, B., Bailey, B. N., Luo, S.*, et al.*, (31)P NMR of apicomplexans and the effects of risedronate on Cryptosporidium parvum growth. *Biochem Biophys Res Commun* **2001**, *284*, 632-7.
289. Artz, J. D., Dunford, J. E., Arrowood, M. J.*, et al.*, Targeting a uniquely nonspecific prenyl synthase with bisphosphonates to combat cryptosporidiosis. *Chem Biol* **2008**, *15*, 1296-306.
290. Ghosh, S., Chan, J. M. W., Lea, C. R.*, et al.*, Effects of Bisphosphonates on the Growth of Entamoeba histolytica and Plasmodium Species in Vitro and in Vivo. *J Med Chem* **2004**, *47*, 175-87.
291. Rodriguez, N., Bailey, B. N., Martin, M. B.*, et al.*, Radical cure of experimental cutaneous leishmaniasis by the bisphosphonate pamidronate. *J Infect Dis* **2002**, *186*, 138-40.
292. Sanders, J. M., Gomez, A. O., Mao, J.*, et al.*, 3-D QSAR investigations of the inhibition of Leishmania major farnesyl pyrophosphate synthase by bisphosphonates. *J Med Chem* **2003**, *46*, 5171-83.
293. Sanders, J. M., Song, Y., Chan, J. M.*, et al.*, Pyridinium-1-yl bisphosphonates are potent inhibitors of farnesyl diphosphate synthase and bone resorption. *J Med Chem* **2005**, *48*, 2957-63.
294. Rodrigues, C. O., Scott, D. A., Bailey, B. N.*, et al.*, Vacuolar proton pyrophosphatase activity and pyrophosphate (PPi) in Toxoplasma gondii as possible chemotherapeutic targets. *Biochem J* **2000**, *349*, 737-45.
295. Ling, Y., Sahota, G., Odeh, S.*, et al.*, Bisphosphonate inhibitors of Toxoplasma gondi growth: in vitro, QSAR, and in vivo investigations. *J Med Chem* **2005**, *48*, 3130-40.
296. World Health Organization, World malaria report 2018. **2018**.
297. Singh, A. P., Zhang, Y., No, J.-H.*, et al.*, Lipophilic Bisphosphonates Are Potent Inhibitors of Plasmodium Liver-Stage Growth. *Antimicrobial Agents and Chemotherapy* **2010**, *54*, 2987-93.
298. No, J. H., de Macedo Dossin, F., Zhang, Y.*, et al.*, Lipophilic analogs of zoledronate and risedronate inhibit Plasmodium geranylgeranyl diphosphate synthase (GGPPS) and exhibit potent antimalarial activity. *Proc Natl Acad Sci U S A* **2012**, *109*, 4058-63.
299. Yang, G., Zhu, W., Kim, K.*, et al.*, In Vitro and In Vivo Investigation of the Inhibition of Trypanosoma brucei Cell Growth by Lipophilic Bisphosphonates. *Antimicrob Agents Chemother* **2015**, *59*, 7530-9.
300. Liu, Y. L., Cao, R., Wang, Y.*, et al.*, Farnesyl diphosphate synthase inhibitors with unique ligand-binding geometries. *ACS Med Chem Lett* **2015**, *6*, 349-54.
301. Zhang, Y., Cao, R., Yin, F.*, et al.*, Lipophilic bisphosphonates as dual farnesyl/geranylgeranyl diphosphate synthase inhibitors: an X-ray and NMR investigation. *J Am Chem Soc* **2009**, *131*, 5153-62.
302. Zhang, H. Z., Lin, J. M., Rasheed, S.*, et al.*, Design, synthesis, and biological evaluation of novel benzimidazole derivatives and their interaction with calf thymus DNA and synergistic effects with clinical drugs. *Sci. China: Chem.* **2014**, *57*, 807-822.
303. Cao, R., Chen, C. K. M., Guo, R.-T.*, et al.*, Structures of a potent phenylalkyl bisphosphonate inhibitor bound to farnesyl and geranylgeranyl diphosphate synthases. *Proteins* **2008**, *73*, 431-439.
304. Mao, J., Mukherjee, S., Zhang, Y.*, et al.*, Solid-state NMR, crystallographic, and computational investigation of bisphosphonates and farnesyl diphosphate synthase-bisphosphonate complexes. *J Am Chem Soc* **2006**, *128*, 14485-97.
305. Berman, H. M., Westbrook, J., Feng, Z.*, et al.*, The Protein Data Bank. *Nucleic Acids Res* **2000**, *28*, 235-242.
306. Srivastava, A., Mukherjee, P., Desai, P. V.*, et al.*, Structural analysis of farnesyl pyrophosphate synthase from parasitic protozoa, a potential chemotherapeutic target. *Infect Disord Drug Targets* **2008**, *8*, 16-30.
307. Song, L., Poulter, C. D., Yeast farnesyl-diphosphate synthase: site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II. *Proc Natl Acad Sci U S A* **1994**, *91*, 3044-8.

308. Oldfield, E., Song, Y., Y., Z. Bisphosphonates compounds and methods. 2007.
309. Lindert, S., Zhu, W., Liu, Y. L., *et al.*, Farnesyl diphosphate synthase inhibitors from in silico screening. *Chem Biol Drug Des* **2013**, *81*, 742-8.
310. Jahnke, W., Rondeau, J. M., Cotesta, S., *et al.*, Allosteric non-bisphosphonate FPPS inhibitors identified by fragment-based discovery. *Nat Chem Biol* **2010**, *6*, 660-6.
311. Jahnke, W., NMR reporter screening for the detection of high-affinity ligands. *Angew. Chem. Int. Edn Engl.* **2002**, *41*, 3420-23.
312. Roddy, T. P., Horvath, C. R., Stout, S. J., *et al.*, Mass spectrometric techniques for label-free high-throughput screening in drug discovery. *Anal Chem* **2007**, *79*, 8207-13.
313. Glickman, J. F., Schmid, A., Farnesyl pyrophosphate synthase: real-time kinetics and inhibition by nitrogen-containing bisphosphonates in a scintillation assay. *Assay Drug Dev Technol* **2007**, *5*, 205-14.
314. Simonen, M., Ibig-Rehm, Y., Hofmann, G., *et al.*, High-Content Assay to Study Protein Prenylation. *J Biomol Screen* **2008**, *13*, 456-467.
315. Jahnke, W., Henry, C., An in vitro assay to measure targeted drug delivery to bone mineral. *ChemMedChem* **2010**, *5*, 770-776.
316. Marzinzik, A. L., Amstutz, R., Bold, G., *et al.*, Discovery of Novel Allosteric Non-Bisphosphonate Inhibitors of Farnesyl Pyrophosphate Synthase by Integrated Lead Finding. *ChemMedChem* **2015**, *10*, 1884-91.
317. Liu, J., Liu, W., Ge, H., *et al.*, Syntheses and characterization of non-bisphosphonate quinoline derivatives as new FPPS inhibitors. *Biochim Biophys Acta* **2014**, *1840*, 1051-62.
318. Jahnke, W., Bold, G., Marzinzik, A. L., *et al.*, A General Strategy for Targeting Drugs to Bone. *Angew Chem Int Ed Engl* **2015**, *54*, 14575-9.
319. Gritzalis, D., Park, J., Chiu, W., *et al.*, Probing the molecular and structural elements of ligands binding to the active site versus an allosteric pocket of the human farnesyl pyrophosphate synthase. *Bioorg Med Chem Lett* **2015**, *25*, 1117-23.
320. Park, J., Zielinski, M., Magder, A., *et al.*, Human farnesyl pyrophosphate synthase is allosterically inhibited by its own product. *Nat Commun* **2017**, *8*, 14132.
321. Gisselberg, J. E., Herrera, Z., Orchard, L. M., *et al.*, Specific Inhibition of the Bifunctional Farnesyl/Geranylgeranyl Diphosphate Synthase in Malaria Parasites via a New Small-Molecule Binding Site. *Cell Chem Biol* **2018**, *25*, 185-193.
322. Murray, C. W., Rees, D. C., The rise of fragment-based drug discovery. *Nat Chem* **2009**, *1*, 187-192.
323. Erlanson, D. A., McDowell, R. S., O'Brien, T., Fragment-based drug discovery. *J Med Chem* **2004**, *47*, 3463-82.
324. Rees, D. C., Congreve, M., Murray, C. W., *et al.*, Fragment-based lead discovery. *Nat Rev Drug Discov* **2004**, *3*, 660.
325. Erlanson, D. A.; Jahnke, W., *Fragment-based Drug Discovery Lessons and Outlook*. Wiley-VCH Verlag GmbH & Co. KGaA 2016.
326. Jencks, W. P., On the attribution and additivity of binding energies. *Proceedings of the National Academy of Sciences of the United States of America* **1981**, *78*, 4046-50.
327. Goodford, P. J., A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J Med Chem* **1985**, *28*, 849-57.
328. Miranker, A., Karplus, M., Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* **1991**, *11*, 29-34.
329. Lamoree, B., Hubbard, R. E., Current perspectives in fragment-based lead discovery (FBLD). *Essays Biochem* **2017**, *61*, 453-64.
330. Lauri, G., Bartlett, P. A., CAVEAT: a program to facilitate the design of organic molecules. *J Comput Aided Mol Des* **1994**, *8*, 51-66.
331. Böhm, H.-J., The computer program LUDI: A new method for the de novo design of enzyme inhibitors. *J Comput Aid Mol Des* **1992**, *6*, 61-78.
332. Eisen, M. B., Wiley, D. C., Karplus, M., *et al.*, HOOK: a program for finding novel molecular architectures that satisfy the chemical and steric requirements of a macromolecule binding site. *Proteins* **1994**, *19*, 199-221.
333. Shuker, S. B., Hajduk, P. J., Meadows, R. P., *et al.*, Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531-4.
334. Nienaber, V. L., Richardson, P. L., Klighofer, V., *et al.*, Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nature Biotechnology* **2000**, *18*, 1105.
335. Congreve, M., Carr, R., Murray, C., *et al.*, A 'Rule of Three' for fragment-based lead discovery? *Drug Discov Today* **2003**, *8*, 876-877.

336. Hann, M. M., Leach, A. R., Harper, G., Molecular Complexity and Its Impact on the Probability of Finding Leads for Drug Discovery. *Journal of Chemical Information and Computer Sciences* **2001**, *41*, 856-864.
337. Hubbard, R., Davis, B., Chen, I.*, et al.*, The SeeDs Approach: Integrating Fragments into Drug Discovery. *Curr Top Med Chem* **2007**, *7*, 1568-1581.
338. Erlanson, D. A., Wells, J. A., Braisted, A. C., Tethering: fragment-based drug discovery. *Annu Rev Biophys Biomol Struct* **2004**, *33*, 199-223.
339. Chessari, G., Woodhead, A. J., From fragment to clinical candidate--a historical perspective. *Drug Discov Today* **2009**, *14*, 668-75.
340. Mortenson, P. N., Erlanson, D. A., de Esch, I. J. P.*, et al.*, Fragment-to-Lead Medicinal Chemistry Publications in 2017. *J Med Chem* **2018**.
341. Alex, A., Flocco, M., Fragment-Based Drug Discovery: What has it Achieved so Far? *Curr Top Med Chem* **2007**, *7*, 1544-1567.
342. Erlanson, D. A., Fragments in the clinic: 2018 edition. http://practicalfragments.blogspot.com/2018/10/fragments-in-clinic-2018-edition.html (accessed 26th January 2019).
343. Lipinski, C. A., Lombardo, F., Dominy, B. W.*, et al.*, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **1997**, *23*, 3-25.
344. Schulz, M. N., Landstrom, J., Bright, K.*, et al.*, Design of a fragment library that maximally represents available chemical space. *J Comput Aided Mol Des* **2011**, *25*, 611-20.
345. Keseru, G. M., Erlanson, D. A., Ferenczy, G. G.*, et al.*, Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia. *J Med Chem* **2016**, *59*, 8189-206.
346. Chen, I. J., Hubbard, R. E., Lessons for fragment library design: analysis of output from multiple screening campaigns. *J Comput Aided Mol Des* **2009**, *23*, 603-20.
347. Murray, C. W., Rees, D. C., Opportunity Knocks: Organic Chemistry for Fragment-Based Drug Discovery (FBDD). *Angew Chem Int Ed Engl* **2016**, *55*, 488-92.
348. Davis, B. J., Roughley, S. D., Chapter Eleven - Fragment-Based Lead Discovery. In *Annual Reports in Medicinal Chemistry*, Goodnow, R. A., Ed. Academic Press: 2017. Vol. 50,
349. Leach, A. R., Hann, M. M., Molecular complexity and fragment-based drug discovery: ten years on. *Current Opinion in Chemical Biology* **2011**, *15*, 489-496.
350. Barelier, S., Krimm, I., Ligand specificity, privileged substructures and protein druggability from fragment-based screening. *Curr Opin Chem Biol* **2011**, *15*, 469-474.
351. Nilar, S. H., Ma, N. L., Keller, T. H., The importance of molecular complexity in the design of screening libraries. *J Comput Aid Mol Des* **2013**, *27*, 783-792.
352. Leach, A. R., Hann, M. M., Molecular complexity and fragment-based drug discovery: ten years on. *Curr Opin Chem Biol* **2011**, *15*, 489-96.
353. Erlanson, D. A., Fesik, S. W., Hubbard, R. E.*, et al.*, Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* **2016**.
354. Erlanson, D. A., Personal Essay: Fragments in the Blogosphere. In *Fragment-based Drug Discovery Vol 47*, Howard, S.; Abell, C., Eds. The Royal Society of Chemistry 2015.
355. Price, A. J., Howard, S., Cons, B. D., Fragment-based drug discovery and its application to challenging drug targets. *Essays Biochem* **2017**, *61*, 475-484.
356. Keserű, G. M., Erlanson, D. A., Ferenczy, G. G.*, et al.*, Design Principles for Fragment Libraries: Maximizing the Value of Learnings from Pharma Fragment-Based Drug Discovery (FBDD) Programs for Use in Academia. *J Med Chem* **2016**, *59*, 8189-8206.
357. Hall, R. J., Mortenson, P. N., Murray, C. W., Efficient exploration of chemical space by fragment-based screening. *Prog Biophys Mol Biol* **2014**, *116*, 82-91.
358. Morley, A. D., Pugliese, A., Birchall, K.*, et al.*, Fragment-based hit identification: thinking in 3D. *Drug Discov Today* **2013**, *18*, 1221-7.
359. Akritopoulou-Zanze, I., Hajduk, P. J., Kinase-targeted libraries: the design and synthesis of novel, potent, and selective kinase inhibitors. *Drug Discov Today* **2009**, *14*, 291-7.
360. Schuffenhauer, A., Ruedisser, S., Marzinzik, A. L.*, et al.*, Library design for fragment based screening. *Curr Top Med Chem* **2005**, *5*, 751-62.
361. Harner, M. J., Frank, A. O., Fesik, S. W., Fragment-based drug discovery using NMR spectroscopy. *J Biomol NMR* **2013**, *56*, 65-75.
362. Kozikowski, B. A., Burt, T. M., Tirey, D. A.*, et al.*, The Effect of Room-Temperature Storage on the Stability of Compounds in DMSO. *J Biomol Screen* **2003**, *8*, 205-209.

363. Singh, M., Tam, B., Akabayov, B., *NMR-Fragment Based Virtual Screening: A Brief Overview*. 2018.
364. Blomberg, N., Cosgrove, D. A., Kenny, P. W.*, et al.*, Design of compound libraries for fragment screening. *J Comput Aided Mol Des* **2009**, *23*, 513-25.
365. Fuller, N., Spadola, L., Cowen, S.*, et al.*, An improved model for fragment-based lead generation at AstraZeneca. *Drug Discovery Today* **2016**, *21*, 1272-1283.
366. Lau, W. F., Withka, J. M., Hepworth, D.*, et al.*, Design of a multi-purpose fragment screening library using molecular complexity and orthogonal diversity metrics. *J Comput Aid Mol Des* **2011**, *25*, 621-36.
367. Aimon, A., Karageorgis, G., Masters, J.*, et al.*, Realisation of small molecule libraries based on frameworks distantly related to natural products. *Org Biomol Chem* **2018**, *16*, 3160-3167.
368. Foley, D. J., Craven, P. G. E., Collins, P. M.*, et al.*, Synthesis and Demonstration of the Biological Relevance of sp(3) -rich Scaffolds Distantly Related to Natural Product Frameworks. *Chemistry* **2017**, *23*, 15227-32.
369. Bian, Y., Xie, X.-Q., Computational Fragment-Based Drug Design: Current Trends, Strategies, and Applications. *The AAPS Journal* **2018**, *20*, 59.
370. Barker, J., Courtney, S., Hesterkamp, T.*, et al.*, Fragment screening by biochemical assay. *Expert Opin Drug Discov* **2006**, *1*, 225-36.
371. Boettcher, A., Ruedisser, S., Erbel, P.*, et al.*, Fragment-based screening by biochemical assays: Systematic feasibility studies with trypsin and MMP12. *J Biomol Screen* **2010**, *15*, 1029-41.
372. Davies, T. G., Tickle, I. J., *Fragment Screening Using X-Ray Crystallography*. Springer, 2011.
373. Navratilova, I., Hopkins, A. L., Fragment screening by surface plasmon resonance. *ACS Med Chem Lett* **2010**, *1*, 44-48.
374. Perspicace, S., Banner, D., Benz, J.*, et al.*, Fragment-based screening using surface plasmon resonance technology. *J Biomol Screen* **2009**, *14*, 337-49.
375. Linke, P., Amaning, K., Maschberger, M.*, et al.*, An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery. *J Biomol Screen* **2016**, *21*, 414-421.
376. Kranz, J. K., Schalk-Hihi, C., Chapter eleven - Protein Thermal Shifts to Identify Low Molecular Weight Fragments. In *Methods in Enzymology*, Kuo, L. C., Ed. Academic Press: 2011. Vol. 493,
377. Erlanson, D. A., Poll results: affiliation, metrics, and fragment-finding methods http://practicalfragments.blogspot.com/2016/10/poll-results-affiliation-metrics-and.html (accessed 10th of March 2019).
378. Hubbard, R. E., Murray, J. B., Chapter twenty - Experiences in Fragment-Based Lead Discovery. In *Methods in Enzymology*, Kuo, L. C., Ed. Academic Press: 2011. Vol. 493,
379. Bowling, J. J., Shadrick, W. R., Griffith, E. C.*, et al.*, Going Small: Using Biophysical Screening to Implement Fragment Based Drug Discovery. **2016**.
380. Gossert, A. D., Jahnke, W., NMR in drug discovery: A practical guide to identification and validation of ligands interacting with biological macromolecules. *Prog Nucl Magn Reson Spectrosc* **2016**, *97*, 82-125.
381. Davis, B., Screening Protein–Small Molecule Interactions by NMR. In *Protein-Ligand Interactions. Methods in Molecular Biology (Methods and Protocols), Vol 1008*, Williams, M.; Daviter, T., Eds. Humana Press, Totowa, NJ: 2013.
382. Lepre, C. A., Library design for NMR-based screening. *Drug Discovery Today* **2001**, *6*, 133-140.
383. Arroyo, X., Goldflam, M., Feliz, M.*, et al.*, Computer-aided design of fragment mixtures for NMR-based screening. *PLoS One* **2013**, *8*, e58571.
384. Stark, J. L., Eghbalnia, H. R., Lee, W.*, et al.*, NMRmix: A Tool for the Optimization of Compound Mixtures in 1D (1)H NMR Ligand Affinity Screens. *J Proteome Res* **2016**, *15*, 1360-8.
385. Mayer, M., Meyer, B., Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angewandte Chemie International Edition* **1999**, *38*, 1784-88.
386. Dalvit, C., Fogliatto, G., Stewart, A.*, et al.*, WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *J Biomol NMR* **2001**, *21*, 349-59.
387. Dalvit, C., Pevarello, P., Tatò, M.*, et al.*, Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water*. *J Biomol NMR* **2000**, *18*, 65-68.
388. Lepre, C. A., Chapter Nine - Practical Aspects of NMR-Based Fragment Screening. In *Methods in Enzymology*, Kuo, L. C., Ed. Academic Press: 2011. Vol. 493,
389. Gossert, A. D., Henry, C., Blommers, M. J.*, et al.*, Time efficient detection of protein-ligand interactions with the polarization optimized PO-WaterLOGSY NMR experiment. *J Biomol NMR* **2009**, *43*, 211-7.

390. Viegas, A., Manso, J., Nobrega, F. L.*, et al.*, Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *J Chem Educ* **2011**, *88*, 990-994.
391. Hajduk, P. J., Olejniczak, E. T., Fesik, S. W., One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules. *J Am Chem Soc* **1997**, *119*, 12257-61.
392. Baldisseri, D., Practical Aspects of Fragment-Based Screening Experiments in TopSpin. *Bruker BioSpin*.
393. Vulpetti, A., Hommel, U., Landrum, G.*, et al.*, Design and NMR-Based Screening of LEF, a Library of Chemical Fragments with Different Local Environment of Fluorine. *J Am Chem Soc* **2009**, *131*, 12949-12959.
394. Vulpetti, A., Dalvit, C., Design and Generation of Highly Diverse Fluorinated Fragment Libraries and their Efficient Screening with Improved 19F NMR Methodology. *ChemMedChem* **2013**, *8*, 2057-2069.
395. Carr, H. Y., Purcell, E. M., Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Physical Review* **1954**, *94*, 630-38.
396. Meiboom, S., Gill, D., Modified Spin-Echo Method for Measuring Nuclear Relaxation Times. *Rev Sci Instrum* **1958**, *29*, 688-691.
397. General Aspects of Organofluorine Compounds. In *Organofluorine Compounds in Biology and Medicine*, Prakash Reddy, V., Ed. Elsevier: 2015.
398. Jordan, J. B., Poppe, L., Xia, X.*, et al.*, Fragment Based Drug Discovery: Practical Implementation Based on 19F NMR Spectroscopy. *J Med Chem* **2012**, *55*, 678-87.
399. Mayer, M., Meyer, B., Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J Am Chem Soc* **2001**, *123*, 6108-17.
400. Jahnke, W., Floersheim, P., Ostermeier, C.*, et al.*, NMR reporter screening for the detection of high-affinity ligands. *Angew Chem Int Ed Engl* **2002**, *41*, 3420-3.
401. Dalvit, C., Fagerness, P. E., Hadden, D. T. A.*, et al.*, Fluorine-NMR Experiments for High-Throughput Screening: Theoretical Aspects, Practical Considerations, and Range of Applicability. *J Am Chem Soc* **2003**, *125*, 7696-703.
402. Dalvit, C., Flocco, M., Veronesi, M.*, et al.*, Fluorine-NMR Competition Binding Experiments for High-Throughput Screening of Large Compound Mixtures. *Comb Chem High Throughput Screening* **2002**, *5*, 605-611.
403. Peng, C., Frommlet, A., Perez, M.*, et al.*, Fast and Efficient Fragment-Based Lead Generation by Fully Automated Processing and Analysis of Ligand-Observed NMR Binding Data. *J Med Chem* **2016**, *59*, 3303-10.
404. Davis, B. J., Erlanson, D. A., Learning from our mistakes: The 'unknown knowns' in fragment screening. *Bioorg Med Chem Lett* **2013**, *23*, 2844-52.
405. Shuker, S. B., Hajduk, P. J., Meadows, R. P.*, et al.*, Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science* **1996**, *274*, 1531-34.
406. Brutscher, B., Intraresidue HNCA and COHNCA experiments for protein backbone resonance assignment. *J Magn Reson* **2002**, *156*, 155-59.
407. Weisemann, R., Rterjans, H., Bermel, W., 3D Triple-resonance NMR techniques for the sequential assignment of NH and 15N resonances in 15N- and 13C-labelled proteins. *Journal of Biomolecular NMR* **1993**, *3*, 113-20.
408. Kelman, Z., *Isotope Labeling of Biomolecules – Labeling Methods, Volume 565*. Elsevier, Waltham, MA, 2015.
409. Atreya, H. S., *Isotope labeling in Biomolecular NMR*. Springer, Dordrecht, 2012.
410. Hajduk, P. J., Augeri, D. J., Mack, J.*, et al.*, NMR-Based Screening of Proteins Containing 13C-Labeled Methyl Groups. *J Am Chem Soc* **2000**, *122*, 7898-904.
411. Schanda, P., Kupce, E., Brutscher, B., SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* **2005**, *33*, 199-211.
412. Dias, D. M., Ciulli, A., NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Prog Biophys Mol Bio* **2014**, *116*, 101-12.
413. Williamson, M. P., Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Mag Res Sp* **2013**, *73*, 1-16.
414. Mattos, C., Ringe, D., Locating and characterizing binding sites on proteins. *Nat Biotechnol* **1996**, *14*, 595-9.
415. Fitzpatrick, P. A., Steinmetz, A. C., Ringe, D.*, et al.*, Enzyme crystal structure in a neat organic solvent. *Proc Natl Acad Sci U S A* **1993**, *90*, 8653-8657.

416. Verlinde, C. L. M. J., Kim, H., Bernstein, B. E.*, et al.*, Anti-trypanosomiasis drug development based on structures of glycolytic enzymes. In *Structure-based drug design*, Marcel Dekker; Inc: 1997.
417. Blundell, T. L., Jhoti, H., Abell, C., High-throughput crystallography for lead discovery in drug design. *Nat Rev Drug Discov* **2002**, *1*, 45-54.
418. Hartshorn, M. J., Murray, C. W., Cleasby, A.*, et al.*, Fragment-Based Lead Discovery Using X-ray Crystallography. *J Med Chem* **2005**, *48*, 403-413.
419. Schiebel, J., Radeva, N., Krimmer, S. G.*, et al.*, Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study. *ACS Chem Biol* **2016**, *11*, 1693-701.
420. O'Reilly, M., Cleasby, A., Davies, T. G.*, et al.*, Crystallographic screening using ultra-low-molecular-weight ligands to guide drug design. *Drug Discov Today* **2019**.
421. Erlanson, D. A., Poll results: what structural information is needed to optimize fragments. http://practicalfragments.blogspot.com/2017/06/poll-results-what-structural.html.
422. Davis, A. M., St-Gallay, S. A., Kleywegt, G. J., Limitations and lessons in the use of X-ray structural information in drug design. *Drug Discov Today* **2008**, *13*, 831-41.
423. Chilingaryan, Z., Yin, Z., Oakley, A. J., Fragment-based screening by protein crystallography: successes and pitfalls. *Int J Mol Sci* **2012**, *13*, 12857-79.
424. Hünefeld, F. L., *Der Chemismus in der thierischen Organisation. Physiologisch-chemische Untersuchungen der materiellen Veränderungen, oder des Bildungslebens im thierischen Organismus; insbesondere des Blutbildungsprocesses, der Natur der Blut Körperchen und ihrer Kernchen. Ein Beitrag zur Physiologie und Heilmittellehre.* Brockhaus, 1840.
425. Kendrew, J. C., Bodo, G., Dintzis, H. M.*, et al.*, A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958**, *181*, 662-66.
426. McPherson, A., Gavira, J. A., Introduction to protein crystallization. *Acta Crystallographica Section F* **2014**, *70*, 2-20.
427. Giege, R., A historical perspective on protein crystallization from 1840 to the present day. *FEBS J* **2013**, *280*, 6456-97.
428. Jancarik, J., Kim, S. H., Sparse-Matrix Sampling - a Screening Method for Crystallization of Proteins. *Journal of Applied Crystallography* **1991**, *24*, 409-411.
429. Tickle, I., Sharff, A., Vinković, M.*, et al.*, High-throughput protein crystallography and drug discovery. *Chemical Society Reviews* **2004**, *33*, 558-565.
430. Bergfors, T., Seeds to crystals. *J Struct Biol* **2003**, *142*, 66-76.
431. D'Arcy, A., Bergfors, T., Cowan-Jacob, S. W.*, et al.*, Microseed matrix screening for optimization in protein crystallization: what have we learned? *Acta Crystallogr F, Struct Biol Comm* **2014**, *70*, 1117-26.
432. Sauter, C., Ng, J. D., Lorber, B.*, et al.*, Additives for the crystallization of proteins and nucleic acids. *Journal of Crystal Growth* **1999**, *196*, 365-376.
433. McPherson, A., Nguyen, C., Cudney, R.*, et al.*, The Role of Small Molecule Additives and Chemical Modification in Protein Crystallization. *Cryst Growth Des* **2011**, *11*, 1469-1474.
434. Saraswathi, N. T., Sankaranarayanan, R., Vijayan, M., Effect of stabilizing additives on the structure and hydration of proteins: a study involving monoclinic lysozyme. *Acta Crystallogr D Biol Crystallogr* **2002**, *58*, 1162-7.
435. Wlodawer, A., Dauter, Z., Jaskolski, M., *Protein Crystallography*. Springer, 2017.
436. Luft, J. R., Wolfley, J. R., Said, M. I.*, et al.*, Efficient optimization of crystallization conditions by manipulation of drop volume ratio and temperature. *Protein Sci* **2007**, *16*, 715-722.
437. Bergfors, T. M., *Protein Crystallization - Techniques, Strategies, and Tips*. TBS The Book Service Ltd, **1999**.
438. Parker, M. W., Protein structure from x-ray diffraction. *J Biolog Physics* **2003**, *29*, 341-362.
439. Smyth, M. S., Martin, J. H., x ray crystallography. *Mol Pathology* **2000**, *53*, 8-14.
440. Blundell, T. L., Johnson, L. N., *Protein Crystallography*. Academic Press Inc. , 1976.
441. Messerschmidt, A., *X-Ray Crystallography of Biomacromolecules*. Wiley-VCH, 2007.
442. Bragg, W. H., Bragg, W. L., The Reflexion of X-rays by Crystals. *Proc R Soc Lond. Series A.* **1913**, *88*, 428-38.
443. Bragg, W. L., The Diffraction of Short Electromagnetic Waves by a Crystal. *Proc. Cambridge Phil. Soc.* **1913**, *17*, 43.
444. Taylor, G. L., Introduction to phasing. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 325-38.
445. Perutz, M. F., Isomorphous Replacement and Phase Determination in Non-Centrosymmetric Space Groups. *Acta Crystallographica* **1956**, *9*, 867-73.

446. Green, D. W., Ingram, V. M., Perutz, M. F., *et al.*, The structure of haemoglobin - IV. Sign determination by the isomorphous replacement method. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1954**, *225*, 287-307.
447. Hendrickson, W. A., Analysis of Protein Structure from Diffraction Measurement at Multiple Wavelengths. *Transactions ACA* **1985**, *21*, 11-21.
448. Karle, J., Some developments in anomalous dispersion for the structural investigation of macromolecular systems in biology. *International Journal of Quantum Chemistry* **1980**, *18*, 357-67.
449. Hendrickson, W. A., Ogata, C. M., Phase determination from multiwavelength anomalous diffraction measurements. In *Methods in Enzymology*, Academic Press: 1997. Vol. 276,
450. Rossmann, M. G., Blow, D. M., The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica* **1962**, *15*, 24-31.
451. Rossmann, M. G., *The Molecular Replacement Method* New York: Gordon & Breach., 1972.
452. Cowtan, K., Phase Problem in X-ray Crystallography, and Its Solution. In *eLS*, 2003.
453. Taylor, G., The phase problem. *Acta Crystallogr D Biol Crystallogr* **2003**, *59*, 1881-90.
454. Broennimann, C., Eikenberry, E. F., Henrich, B., *et al.*, The PILATUS 1M detector. *J Synchrotron Radiat* **2006**, *13*, 120-30.
455. Mueller, M., Wang, M. T., Schulze-Briese, C., Optimal fine phi-slicing for single-photon-counting pixel detectors. *Acta Crystallographica Section D-Biological Crystallography* **2012**, *68*, 42-56.
456. Grimes, J. M., Hall, D. R., Ashton, A. W., *et al.*, Where is crystallography going? *Acta crystallographica. Section D, Structural biology* **2018**, *74*, 152-166.
457. Industrial Liasion Office, BART sample changer - Now available on all MX beamlines. https://www.diamond.ac.uk/industry/Industry-News/Latest-News/Synchrotron-Industry-News---MXnews5/BART-update.html (accessed 26th of April 2019).
458. Bowler, M. W., Nurizzo, D., Barrett, R., *et al.*, MASSIF-1: a beamline dedicated to the fully automatic characterization and data collection from crystals of biological macromolecules. *J Synchrotron Radiat* **2015**, *22*, 1540-7.
459. Marquez, J. A., Cipriani, F., CrystalDirect: a novel approach for automated crystal harvesting based on photoablation of thin films. *Methods Mol Biol* **2014**, *1091*, 197-203.
460. Viola, R., Carman, P., Walsh, J., *et al.*, Operator-assisted harvesting of protein crystals using a universal micromanipulation robot. *J Appl Crystallogr* **2007**, *40*, 539-45.
461. Ng, J. T., Dekker, C., Kroemer, M., *et al.*, Using textons to rank crystallization droplets by the likely presence of crystals. *Acta crystallographica. Section D, Biological crystallography* **2014**, *70*, 2702-18.
462. Krojer, T., Talon, R., Pearce, N., *et al.*, The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination. *Acta Crystallogr D Struct Biol* **2017**, *73*, 267-78.
463. Pearce, N. M., Krojer, T., Bradley, A. R., *et al.*, A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat Commun* **2017**, *8*, 15123.
464. Pearce, N. M., Bradley, A. R., Krojer, T., *et al.*, Partial-occupancy binders identified by the Pan-Dataset Density Analysis method offer new chemical opportunities and reveal cryptic binding sites. *Structural Dynamics* **2017**, *4*, 032104.
465. Davis, B. J., Giannetti, A. M., The Synthesis of Biophysical Methods In Support of Robust Fragment-Based Lead Discovery. In *Fragment-based Drug Discovery Lessons and Outlook*, Erlanson, D. A.; Jahnke, W., Eds. 2016.
466. Erlanson, D. A., Davis, B. J., Jahnke, W., Fragment-Based Drug Discovery: Advancing Fragments in the Absence of Crystal Structures. *Cell Chemical Biology* **2018**.
467. Schultes, S., de Graaf, C., Haaksma, E. E. J., *et al.*, Ligand efficiency as a guide in fragment hit selection and optimization. *Drug Discov Today Technol* **2010**, *7*, e147-202.
468. Hopkins, A. L., Groom, C. R., Alex, A., Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* **2004**, *9*, 430-1.
469. Hopkins, A. L., Keseru, G. M., Leeson, P. D., *et al.*, The role of ligand efficiency metrics in drug discovery. *Nat Rev Drug Discov* **2014**, *13*, 105-21.
470. Kenny, P. W., The nature of ligand efficiency. *J Cheminform* **2019**, *11*, 8.
471. Johnson, C. N., Erlanson, D. A., Jahnke, W., *et al.*, Fragment-to-Lead Medicinal Chemistry Publications in 2016. *Journal of Medicinal Chemistry* **2018**, *61*, 1774-84.
472. Carr, R. A., Congreve, M., Murray, C. W., *et al.*, Fragment-based lead discovery: leads by design. *Drug Discov Today* **2005**, *10*, 987-92.
473. Kenny, P. W., Leitao, A., Montanari, C. A., Ligand efficiency metrics considered harmful. *J Comput Aided Mol Des* **2014**, *28*, 699-710.

474. Shultz, M. D., Improving the plausibility of success with inefficient metrics. *ACS Med Chem Lett* **2014**, *5*, 2-5.
475. Leeson, P. D., Springthorpe, B., The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat Rev Drug Discov* **2007**, *6*, 881-90.
476. Verdonk, M. L., Rees, D. C., Group efficiency: a guideline for hits-to-leads chemistry. *ChemMedChem* **2008**, *3*, 1179-80.
477. Reynolds, C. H., Bembenek, S. D., Tounge, B. A., The role of molecular size in ligand efficiency. *Bioorg Med Chem Lett* **2007**, *17*, 4258-61.
478. Ferenczy, G., Keseru, G., Ligand Efficiency Metrics and their Use in Fragment Optimizations. In *Fragment-based Drug Discovery Lessons and Outlook*, Erlanson, D. A. J., Wolfgang, Ed. Wiley-VCH Verlag GmbH & Co: 2016.
479. Matsui, Y., Yasumatsu, I., Asahi, T.*, et al.*, Discovery and structure-guided fragment-linking of 4-(2,3-dichlorobenzoyl)-1-methyl-pyrrole-2-carboxamide as a pyruvate kinase M2 activator. *Bioorg Med Chem* **2017**, *25*, 3540-3546.
480. Howard, N., Abell, C., Blakemore, W.*, et al.*, Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. *J Med Chem* **2006**, *49*, 1346-55.
481. Mondal, M., Radeva, N., Fanlo-Virgos, H.*, et al.*, Fragment Linking and Optimization of Inhibitors of the Aspartic Protease Endothiapepsin: Fragment-Based Drug Design Facilitated by Dynamic Combinatorial Chemistry. *Angew Chem Int Ed Engl* **2016**, *55*, 9422-6.
482. De Fusco, C., Brear, P., Iegre, J.*, et al.*, A fragment-based approach leading to the discovery of a novel binding site and the selective CK2 inhibitor CAM4066. *Bioorg Med Chem* **2017**, *25*, 3471-82.
483. Nazare, M., Matter, H., Will, D. W.*, et al.*, Fragment deconstruction of small, potent factor Xa inhibitors: exploring the superadditivity energetics of fragment linking in protein-ligand complexes. *Angew Chem Int Ed Engl* **2012**, *51*, 905-11.
484. Wilfong, E. M., Du, Y., Toone, E. J., An enthalpic basis of additivity in biphenyl hydroxamic acid ligands for stromelysin-1. *Bioorg Med Chem Lett* **2012**, *22*, 6521-4.
485. Ichihara, O., Barker, J., Law, R. J.*, et al.*, Compound Design by Fragment-Linking. *Mol Inform* **2011**, *30*, 298-306.
486. Nikiforov, P. O., Surade, S., Blaszczyk, M.*, et al.*, A fragment merging approach towards the development of small molecule inhibitors of Mycobacterium tuberculosis EthR for use as ethionamide boosters. *Org Biomol Chem* **2016**, *14*, 2318-26.
487. Hudson, S. A., McLean, K. J., Surade, S.*, et al.*, Application of fragment screening and merging to the discovery of inhibitors of the Mycobacterium tuberculosis cytochrome P450 CYP121. *Angew Chem Int Ed Engl* **2012**, *51*, 9311-6.
488. Taylor, S. J., Padyana, A. K., Abeywardane, A.*, et al.*, Discovery of Potent, Selective Chymase Inhibitors via Fragment Linking Strategies. *Journal of Medicinal Chemistry* **2013**, *56*, 4465-81.
489. de Vicente, J., Tivitmahaisoon, P., Berry, P.*, et al.*, Fragment-Based Drug Design of Novel Pyranopyridones as Cell Active and Orally Bioavailable Tankyrase Inhibitors. *ACS Med Chem Lett* **2015**, *6*, 1019-24.
490. Hoffer, L., Muller, C., Roche, P.*, et al.*, Chemistry-driven Hit-to-lead Optimization Guided by Structure-based Approaches. *Mol Inform* **2018**, *37*, e1800059.
491. Enamine, Enamine Golden Fragment Library. https://enamine.net/download/FL/Enamine_Golden_Fragment_Library.pdf (accessed 16th January 2019).
492. Gasteiger, E., Hoogland, C., Gattiker, A.*, et al.*, Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, Walker, J. M., Ed. Humana Press: 2005.
493. Laemmli, U. K., Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680-5.
494. Westermeier, R., *Electrophoresis in Practice: A Guide to Methods and Applications of DNA and Protein Separations, 4th, Revised and Updated Edition*. WILEY-VCH, 2004.
495. Alvarez-Garcia, D., Barril, X., Molecular simulations with solvent competition quantify water displaceability and provide accurate interaction maps of protein binding sites. *J Med Chem* **2014**, *57*, 8530-9.
496. Hwang, T. L., Shaka, A. J., Water Suppression That Works - Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients. *J Magn Reson Ser A* **1995**, *112*, 275-79.
497. Kabsch, W., Xds. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 125-32.
498. Evans, P. R., Murshudov, G. N., How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **2013**, *69*, 1204-14.
499. Sharff A, K. P., Vonrhein C, Smart O, Womack T, Flensburg C, Paciorek C and Bricogne G, Pipedream, version 1.2.1. *Global Phasing Ltd* **2011**, *Cambridge*, United Kingdom.

500. Vonrhein, C., Flensburg, C., Keller, P.*, et al.*, Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D Biol Crystallogr* **2011**, *67*, 293-302.
501. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D.*, et al.*, Phaser crystallographic software. *J Appl Crystallogr* **2007**, *40*, 658-74.
502. Smart, O. S., Womack, T. O., Flensburg, C.*, et al.*, Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* **2012**, *68*, 368-80.
503. Emsley, P., Lohkamp, B., Scott, W. G.*, et al.*, Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486-501.
504. Emsley, P., Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2126-32.
505. Adams, P. D., Afonine, P. V., Bunkoczi, G.*, et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 213-21.
506. Kleywegt, G. J., Brünger, A. T., Checking your imagination: applications of the free R value. *Structure* **1996**, *4*, 897-904.
507. Painter, J., Merritt, E. A., Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D* **2006**, *62*, 439-50.
508. Yang, H., Guranovic, V., Dutta, S.*, et al.*, Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 1833-9.
509. Collins, P. M., Ng, J. T., Talon, R.*, et al.*, Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta Crystallogr D Struct Biol* **2017**, *73*, 246-55.
510. Winter, G., xia2: an expert system for macromolecular crystallography data reduction. *Journal of Applied Crystallography* **2010**, *43*, 186-90.
511. Waterman, D. G., Winter, G., Gildea, R. J.*, et al.*, Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr D* **2016**, *72*, 558-575.
512. Evans, P., Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **2006**, *62*, 72-82.
513. Murshudov, G. N., Skubak, P., Lebedev, A. A.*, et al.*, REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* **2011**, *67*, 355-67.
514. Collaborative Computational Project, N., The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **1994**, *50*, 760-3.
515. Winn, M. D., Ballard, C. C., Cowtan, K. D.*, et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Biological Crystallography* **2011**, *67*, 235-42.
516. Delageniere, S., Brenchereau, P., Launer, L.*, et al.*, ISPyB: an information management system for synchrotron macromolecular crystallography. *Bioinformatics* **2011**, *27*, 3186-92.
517. Long, F., Nicholls, R. A., Emsley, P.*, et al.*, AceDRG: a stereochemical description generator for ligands. *Acta Crystallogr D Struct Biol* **2017**, *73*, 112-122.
518. Steiner, R. A., Tucker, J. A., Keep it together: restraints in crystallographic refinement of macromolecule-ligand complexes. *Acta Crystallogr D Struct Biol* **2017**, *73*, 93-102.
519. Cipriani, F., Rower, M., Landret, C.*, et al.*, CrystalDirect: a new method for automated crystal harvesting based on laser-induced photoablation of thin films. *Acta Crystallogr D Biol Crystallogr* **2012**, *68*, 1393-9.
520. Zander, U., Hoffmann, G., Cornaciu, I.*, et al.*, Automated harvesting and processing of protein crystals through laser photoablation. *Acta Crystallogr D* **2016**, *72*, 454-66.
521. Svensson, O., Malbet-Monaco, S., Popov, A.*, et al.*, Fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr* **2015**, *71*, 1757-67.
522. McCarthy, A. A., Barrett, R., Beteva, A.*, et al.*, ID30B - a versatile beamline for macromolecular crystallography experiments at the ESRF. *J Synchrotron Radiat* **2018**, *25*, 1249-60.
523. Incardona, M. F., Bourenkov, G. P., Levik, K.*, et al.*, EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. *J. Synchrot. Radiat.* **2009**, *16*, 872-79.
524. Monaco, S., Gordon, E., Bowler, M. W.*, et al.*, Automatic processing of macromolecular crystallography X-ray diffraction data at the ESRF. *J Appl Crystallogr* **2013**, *46*, 804-10.
525. Sparta, K. M., Krug, M., Heinemann, U.*, et al.*, Xdsapp2.0. *J Appl Cryst* **2016**, *49*, 1085-92.
526. Studier, F. W., Protein production by auto-induction in high-density shaking cultures. *Prot Expr Puri* **2005**, *41*, 207-34.
527. Rosano, G. L., Ceccarelli, E. A., Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol* **2014**, *5*, 172.

528. Obmolova, G., Malia, T. J., Teplyakov, A.*, et al.*, Promoting crystallization of antibody-antigen complexes via microseed matrix screening. *Acta Crystallogr Sect D, Biol Crystallogr* **2010**, *66*, 927-33.
529. Newman, J., Pearce, L., Lesburg, C. A.*, et al.*, Crystallization of an apo form of human arginase: using all the tools in the toolbox simultaneously. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **2011**, *67*, 90-3.
530. Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A.,, Vonrhein, C., Bricogne, G. , STARANISO. *Cambridge, United Kingdom: Global Phasing Ltd* **2018**.
531. Pflugrath, J. W., The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr* **1999**, *55*, 1718-25.
532. Kabsch, W., Evaluation of Single-Crystal X-Ray-Diffraction Data from a Position-Sensitive Detector. *J Appl Crystallogr* **1988**, *21*, 916-924.
533. Galaka, T., Ferrer Casal, M., Storey, M.*, et al.*, Antiparasitic Activity of Sulfur- and Fluorine-Containing Bisphosphonates against Trypanosomatids and Apicomplexan Parasites. *Molecules* **2017**, *22*, 82.
534. Ashkenazy, H., Abadi, S., Martz, E.*, et al.*, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **2016**, *44*, W344-50.
535. Harner, M. J., Frank, A. O., Fesik, S. W., Fragment-based drug discovery using NMR spectroscopy. *Journal of biomolecular NMR* **2013**, *56*, 65-75.
536. Pearce, N. M., Krojer, T., von Delft, F., Proper modelling of ligand binding requires an ensemble of bound and unbound states. *Acta Crystallogr D Struct Biol* **2017**, *73*, 256-66.
537. Kortagere, S., Ekins, S., Welsh, W. J., Halogenated ligands and their interactions with amino acids: Implications for structure–activity and structure–toxicity relationships. *J Mol Graph Modell* **2008**, *27*, 170-77.
538. Wilcken, R., Zimmermann, M. O., Lange, A.*, et al.*, Principles and applications of halogen bonding in medicinal chemistry and chemical biology. *J Med Chem* **2013**, *56*, 1363-88.
539. Hudson, L., Mui, J., Vázquez, S.*, et al.*, Novel Quinazolinone Inhibitors of ALK2 Flip between Alternate Binding Modes: Structure–Activity Relationship, Structural Characterization, Kinase Profiling, and Cellular Proof of Concept. *J Med Chem* **2018**, *61*, 7261-72.
540. Malhotra, S., Karanicolas, J., When Does Chemical Elaboration Induce a Ligand To Change Its Binding Mode? *J Med Chem* **2017**, *60*, 128-145.
541. Hardegger, L. A., Kuhn, B., Spinnler, B.*, et al.*, Halogen Bonding at the Active Sites of Human Cathepsin L and MEK1 Kinase: Efficient Interactions in Different Environments. *ChemMedChem* **2011**, *6*, 2048-54.
542. Panesar, H. K., Solano, J., Minehan, T. G., Synthesis and DNA binding profile of N-mono- and N,N'-disubstituted indolo[3,2-b]carbazoles. *Org Biomol Chem* **2015**, *13*, 2879-83.
543. Nie, Z. S., Jeffrey Alan; Veal, James Marvin; Wallace, Michael Brennan Preparation of pyrazolylpyridine derivatives as histone demethylase inhibitors. *WO 2014089364* **2014**.
544. Souers, A. J., Collins, C. A., Gao, J., Judd, A. S., Kym, P. R., Mulhern, M. M., Sham, H. L., Wodka, D. Antagonists of melanin concentrating hormone effects on the melanin concentrating hormone receptor. 2004.
545. Jeankumar, V. U., Reshma, R. S., Vats, R.*, et al.*, Engineering another class of anti-tubercular lead: Hit to lead optimization of an intriguing class of gyrase ATPase inhibitors. *Eur J Med Chem* **2016**, *122*, 216-31.
546. Sato, S., Sakamoto, T., Miyazawa, E.*, et al.*, One-pot reductive amination of aldehydes and ketones with α-picoline-borane in methanol, in water, and in neat conditions. *Tetrahedron* **2004**, *60*, 7899-06.
547. Dunsmore, C. J., Carr, R., Fleming, T.*, et al.*, A chemo-enzymatic route to enantiomerically pure cyclic tertiary amines. *J Am Chem Soc* **2006**, *128*, 2224-5.
548. Ombrato, R., Garofalo, B., Mangano, G.*, et al.* Preparation of heterocyclic compounds as antibacterial agents. WO2016096631A1, 2016.
549. Mayrose, I., Graur, D., Ben-Tal, N.*, et al.*, Comparison of Site-Specific Rate-Inference Methods for Protein Sequences: Empirical Bayesian Methods Are Superior. *Mol Biol Ev* **2004**, *21*, 1781-91.

10. Appendix

10.1 Plasmid maps

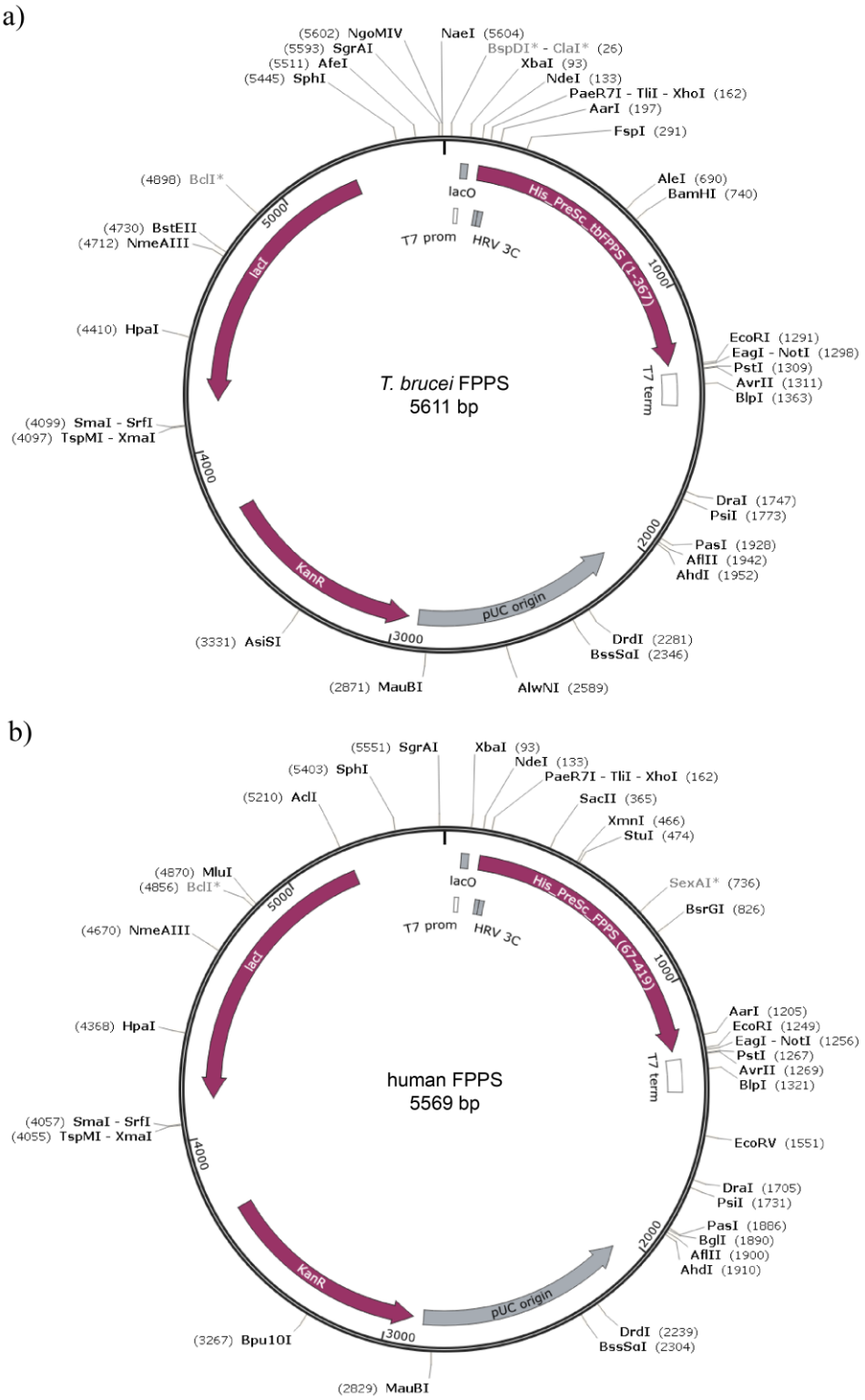


Figure 73: *T. brucei* and human FPPS plasmid maps. Plasmid maps of **(a)** *T. brucei* FPPS and **(b)** human FPPS are shown. Plasmid maps were generated with SnapGene.

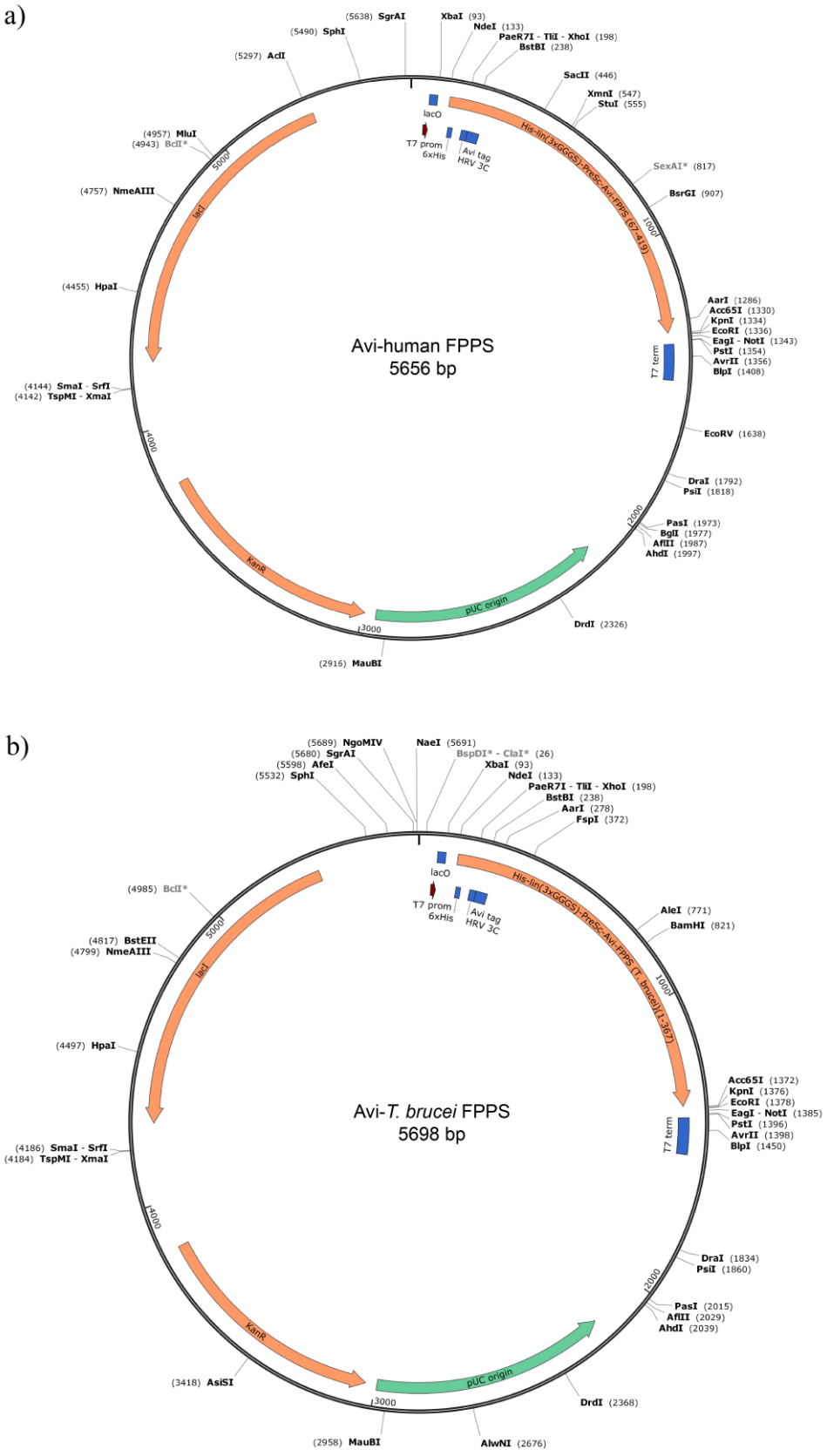


Figure 74: Avi-*T. brucei*, avi-human, and avi-*T. cruzi* FPPS plasmid maps. Plasmid maps of **(a)** avi-*T. brucei* FPPS **(b)** avi-human FPPS and (c) avi-*T. cruzi* FPPS are shown. Plasmid maps were generated with SnapGene.

c)

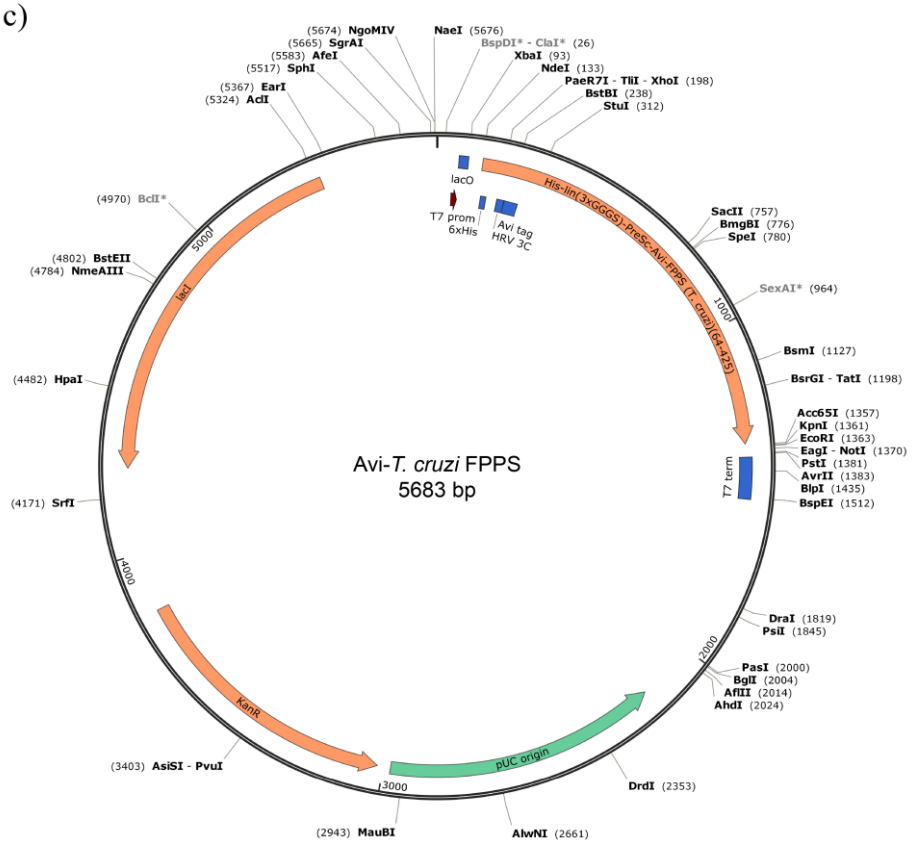


Figure 74: continued.

10.2 Compounds

10.2.1 Human allosteric binders and analogues thereof

Table 34: Compound list of human allosteric binders and analogues thereof.

Cmpd Nr.	Ligand in PDB	PDB ID	SMILES code	MW [Da]	clogP[1]	A: CAS, B: PubChem, C: Reaxys	prot.NMR[2]	Binding site
113	JQE	6R38	ClC1=CC=C2C(C(CC(O)=O)=C(Cl)S2)=C1	261.1	3.92	B: 612075	n/a	allosteric
114	-	-	ClC1=CC2=C(C=C1)SC(C(C)C(O)=O)=C2	240.7	3.50	B:131453858	medium	-
115	-	-	CC(C(C=C1Cl)=C(C=C1)S2)=C2CC(O)=O	240.7	3.64	A: 51527-19-6	medium	-
116	3N2	6R37	ClC1=CC=C2C(C(CC(O)=O)=CS2)=C1	226.7	3.19	A: 17266-30-7	medium	allosteric
117	-	-	OC(C1=NC(C(C2=CC=CC3=CC=CC=C23)=CC=C4)=C4C=C1)=O	299.3	5.45	A : 1185407-78-6	aggregation	-
118	-	-	ClC1=CC(Cl)=CC2=C1NC(C(O)=O)=C2CC(O)=O	288.1	2.64	C: 24889800	strong	-
119	-	-	O=C(O)CC1=CSC2=CC=C3C(C=CC=C3)=C21	242.3	3.62	A: 108900-25-0	strong	-
120	-	-	O=C(O)CC1=COC2=C1C=CC(OC)=C2	206.2	1.90	A: 69716-05-8	weak	-
121	BFH	6R39	OC(C1=CC2=CC=C3C=CC=CC3=C2N1CC(O)=O)=O	269.3	3.07	C: 24135224	strong	allosteric
122	-	-	OC(C1=CC2=CC=C3C=CC=CC3=C2N1CC4=CC(C(O)=O)=NO4)=O	336.3	3.41	C: 24889801	weak	-

[1] clogP values obtained from ChemDraw.

[2] Classification of compounds in weak, medium and strong shifter on *T. brucei* FPPS according to [^{13}C, ^{1}H] -SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K. Compound **113** was not tested (n/a).

10.2.2 *T. brucei* FPPS hits of Novartis core fragment library

Table 35: Compound list of *T. brucei* FPPS hits of Novartis core fragment library.

Cmpd Nr.	Ligand in PDB	PDB ID	SMILES code	MW	clogP[1]	A:CAS,B:PubChem, C: Reaxys	TbrFPPS[2]	TcrFPPS[2]	hFPPS[2]	Prot. NMR[3]
123	PJD	5QTA	C1(CS(C)(=O)=O)C(=CC=CC=1Cl)Cl	239.1	1.70	A: 20018-02-4	1	0	1	M
124			C2(OC1C=CC(O)=CC=1)C(=CC(Cl)=CN=2)Cl	256.1	3.39	B: 620880	1	1	1	M/Prec
125			C12C(=C(C=C(C=1)OC)N)N=CC=C2	174.2	1.80	A: 90-52-8	1	1	1	W
126			C1=CC=C2C(=C1)CC(N2C(=O)OCC3=CC=CC=C3)CO	283.3	3.15	A: 135829-04-8	1	1	1	M/Prec
127			C2(C1N=C(N)SC=1)=C(C=CC=C2)O	192.2	1.57	A: 60135-72-0	1	1	1	W
128			C23C1=C(CCCC1)SC=2N=CN=C3N(C)C	233.3	3.38	B: 4112641	1	1	1	M/Prec
129	P9Y	5QT4	C1(=CC=C(Cl)C=C1)OCCCN2C=CN=C2	236.7	3.00	B: 3599333	1	1	1	W
130	PJG	5QTB	C12(CC3CC(C1)CC(C2)C3)NCCN	194.3	1.26	A: 37818-93-2	1	1	1	S
131	JNE	5QT6	C1(N=C(C=CC=1)C)NCC2C=CC=CC=2	198.3	3.06	A: 70644-47-2	1	1	1	S
132	PBV	5QT5	C1(C=CC=C(N=1)N)CC	122.2	1.35	A: 21717-29-3	1	1	1	M
133			C23N(CC1C=CC(Cl)=CC=1)C=NC=2N=CN=C3NC	273.7	2.66	B: 110171704	1	1	0	M
134			C1(C3C(N(C)C2C(N=1)=CC=CC=2)=CC=CC=3)N4CCNCC4	292.4	2.65	B: 69041412	1	1	0	M
135			C23(N(C(C1=CC=CC=C1C2)=O)C)CCCCC3	229.3	2.85	B: 21818044	1	1	1	M/Prec
136	PJ4	5QT9	C23C(N1CCN(C)CC1)=NC(Cl)=CC=2C=CC=C3	261.8	3.18	B: 23155989	1	1	1	M
137			C23C(CN1C=CN=C1N2)=CC(OC)=C(C=3OC)OC	261.3	1.12	B: 20509913	1	1	1	W
138	PBY	5QT7	N(C1=CC(=CC=C1)Cl)C2C=C(N=CN=2)N	220.7	3.07	B: 66669982	1	1	1	M

[1] clogP values obtained from ChemDraw.

[2] Validated hits by ligand-observed NMR on *T. brucei*, *T.cruzi* and human FPPS. For a validated hit on the respective protein the number 1 is given and for no binding a 0 is noted.

[3] Classification of a compound in weak (W), medium (M), medium and also showing precipitation (M/Prec.) and strong shifter (S) on *T. brucei* FPPS according to [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K.

10.2.3 *T. brucei* FPPS hits of Novartis fluorine fragment library

Table 36: Compound list of *T. brucei* FPPS hits of Novartis fluorine fragment library.

Cmpd Nr.	SMILES code	MW	clogP[1]	A: CAS, B: PubChem, C: Reaxys	TbrFPPS[2]	TcrFPPS[2]	hFPPS[2]	Prot. NMR[3]
139	NC1=NC2=CC=C(C(F)(F)F)C=C2S1	218.2	4.85	A: 777-12-8	1	1	1	strong
140	CC1=CC([N+]([O-])=O)=C(C(F)(F)F)C=C1N	220.2	2.76	B: 129319121	1	1	1	strong
141	FC(F)(F)C1=CC(/C=N/NC(N)=N)=CC(C(F)(F)F)=C1	298.2	3.60	A: 23557-66-6	1	1	0	weak/prec.
142	ClC1=C(C(F)(F)F)C=C([N+]([O-])=O)C(N)=C1	240.6	3.40	A: 35375-74-7	1	1	1	medium
143	FC(F)(F)/C(C1=CSC=C1)=N/O	195.2	2.52	A: 138395-47-8	1	1	1	medium
144	NC1=C(C(NC2=CC(C(F)(F)F)=CC=C2)=O)C=CC=C1	280.3	3.68	A: 20878-52-8	1	1	1	weak
145	FC(F)(F)C1=NN2C(C(C3=CC=CS3)=C1)=NN=C2	270.2	2.07	B: 760142	1	1	1	strong
146	NC1=CC(C(F)(F)F)=CC=C1C(OC2)=NC2(C)C	258.2	3.62	A: 1361005-81-3	1	1	1	medium
147	FC(F)(F)C(C=C1)=CC(N)=C1[N+]([O-])=O	206.1	2.85	A: 402-14-2	1	1	1	medium
148	FC(F)(F)C1=NC(NC2=CC=C(Cl)C=C2)=NC(N)=N1	289.6	3.23	A: 53387-70-5	1	1	1	weak
149	OC1=C(C(F)(F)F)C=C([N+]([O-])=O)C=C1C(F)(F)F	275.1	3.71	B: 13784430	1	1	0	strong

[1] clogP values obtained from ChemDraw.

[2] Validated hits by ligand-observed NMR on T. brucei, T.cruzi and human FPPS. For a validated hit on the protein the number 1 is given and for no binding a 0 is given

[3] Classification of a compound in weak (W), medium (M), medium and also showing precipitation (M/Prec.) and strong shifter (S) on *T. brucei* FPPS according to [^{13}C, ^{1}H]-SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K.

10.2.4 Fragment analogues of active-site binder 129 – SAR by archive and catalogue

Table 37: Compound list of fragment analogues of active-site binder 129.

Cmpd Nr.	SMILES code	MW	clogP[2]	A: CAS, B: PubChem, C: Reaxys	Prot. NMR[3]
150	N1(C=CN=C1)CCCCCC(N)=O	181.2	-0.23	A: 88940-82-3	very weak
151	N1(CCCC(O)=O)C=CN=C1	154.2	-0.05	A: 72338-58-0	very weak
152	C2(C(CCN1C=CN=C1)=O)C=CC(Cl)=CC=2	234.7	2.16	A: 13193-72-9	weak
153[1]	C2(N(CC(CN1C=CN=C1)O)C)C=CC(Cl)=CC=2	265.7	2.15	A: 34105-16-3	medium
154	C1(=CC(=C(Cl)C(=C1)C)C)OCCCN2C=CN=C2	264.8	4.00	B: 2852681	aggregation
155	C1(C(=CC(Cl)=CC=1)Cl)OCCCN2C=CN=C2	271.1	3.52	C: 28655004	strong
156	N2=CN(CCCOC1=CC=C(C=C1)Br)C=C2	281.2	3.14	A: 415936-60-6	medium
157	ClC1C=CC=C(C=1)OCCCN2C=NC=C2	236.7	2.99	B: 2843569	medium
158	COC2=CC=C(OCCCN1C=CN=C1)C=C2C	246.3	2.73	A: 1796945-26-0	very weak
159	N1(C=NC=C1)C(CC(NC2=CC=C(C=C2)Cl)=O)C	263.7	2.33	A: 930547-67-4	medium
160	ClC2=CC=C(NC(=O)CN1C=CN=C1)C=C2	235.7	1.73	A: 72502-24-0	very weak
161	CC2N=CN(CC(=O)NC1=CC=C(C)C=C1C)C=2C	257.3	1.60	A: 1376365-08-0	very weak
162	ClC1=CC=C(C=C1)C(=O)NCCN2C=CN=C2	249.7	1.74	A: 1009049-54-0	very weak
163	O=C(CCN1C=CN=C1)NC3=CC=C2C=CC=CC2=C3	265.3	2.23	A: 793679-33-1	weak

[1] Ligand in PDB: PJ1, PDB ID 5QT8

[2] clogP values obtained from ChemDraw.

[3] Classification of compounds in very weak, weak, medium and strong shifter on *T. brucei* FPPS according to [^{13}C, ^{1}H] -SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K.

10.2.5 Compounds derived from DSPL and Edelris Keymical Fragments library

Table 38: Compound list of Oxford fragments.

Cmpd Nr.	SMILES code	MW	clogP[1]	A: CAS, B: PubChem, C: Reaxys	TbruFPPS/TcruFPPS[2]
164	CC1(C)[C@]2([H])C(CN3[C@@]([H])(CO)CCC3)=CC[C@]1(C2)[H]	235.37	4.24	B: 40514543	TbruFPPS
165	COC1=CC=C2SC(N)=NC2=C1	180.23	2.13	A: 54346-87-1	TbruFPPS
166	CC1=C(C(NCC(C)C)=O)N(C=CC=C2)C2=N1	231.30	2.59	A: 352562-29-9	TbruFPPS
167	CN(C(CN1C2)C3=CC=CC=C3S1(=O)=O)C2=O	252.29	1.08	B:132461954	TbruFPPS
168	O[C@H]1[C@H](NC(=O)CC2CCCC2)[C@H]3O[C@@H]1c4ccccc34	287.36	2.48	n/a	TbruFPPS
169	OC1=CC=C(CNC2CCCCC2)C=C1	205.30	2.87	A: 940366-58-5	TbruFPPS
170	CC(NC(C1=CC=C(C=CC=C2)C2=N1)=O)C	214.27	2.68	A: 586982-18-5	TbruFPPS
171	FC1=CC=CC(F)=C1C(N2CCCCCC2)=O	239.27	2.75	A: 385380-70-1	TbruFPPS
172	CC([C@@](S1)(NO)[H])(C)N(C)C1=S	192.30	1.14	B: 818441	TbruFPPS
173	COC(NC1=CC=C(CC2=CC=NC=C2)C=C1)=O	242.28	2.18	A: 329921-01-9	TbruFPPS
174	CN1[C@@]([H])([C@@](O)(C2=CC=CC=C2)[H])[C@](O)(CC3)CC[C@@]13[H]	247.34	2.08	n/a	TbruFPPS
175	ClC1=CC=C(NCCC(N(C2CC2)C3)=O)C3=C1	250.73	2.35	A: 1307894-22-9	TbruFPPS
176	CC1(CNCC(NC1)=O)C	142.2	0.15	A: 67744-44-9	TbruFPPS
177	c1cc(cc(c1)Cl)NC(=O)C(F)(F)F	223.6	3.25	A : 40410-54-6	TbruFPPS
178	C1(NCCC2=CC=CC=C2)=NC3=CC=CC=C3N1	237.30	4.13	C : 11220754	TcruFPPS
179	N1(CCCCCC1)CC2=CNC3=C2C=CC=C3	228.34	3.72	C : 401260	TcruFPPS
180	CCC1CCC(N2CCOCC2)CC1	197.32	3.16	B: 756963	TcruFPPS
181	O=S1(CCN(CC(C=C2)=CC=C2C)CC1)=O	239.33	1.42	A: 477858-35-8	TcruFPPS

Table 38: continued.

Cmpd Nr.	SMILES code	MW	clogP[1]	A: CAS, B: PubChem, C: Reaxys	TbruFPPS/TcruFPPS[2]
182	CC(CC1)CCN1CC2=CNC3=C2C=CC=C3	228.3	3.68	A: 21000-95-3	TcruFPPS
183	CN1CCN(CC1)C2=NC(C=CC=C3)=C3S2	233.3	2.69	A: 34173-40-5	TcruFPPS
184	O[C@@H]1[C@H](CCN(CC#C)C1)OC2=CC=C(C)C(C)=C2	259.4	2.74	C: 13315827	TcruFPPS

[1] clogP values obtained from ChemDraw.
[2] TbruFPPS/TcruFPPS stands for compounds derived from the respective fragment screen on the two proteins.

10.2.6 Compounds derived from Enamine fragment library

Table 39: Compound list of iNEXT fragments.

Cmpd Nr.	Ligand in PDB	PDB ID	SMILES code	MW	clogP[1]	A: CAS, B: PubChem, C: Reaxys	Prot. NMR[2]
185	PJP	5QTK	O=S(C1=CC=C(C)C=C1)(N(CC)CC)=O	227.3	2.81	A: 649-15-0	n/a
186	PJS	5QTJ	O=C(C1=NC(C)=NC=C1)N(CC2=C(F)C=CC=C2Cl)C	293.7	2.09	B: 47256035	medium
187	PJV	5QTI	BrC1=CC=CC=C1OC2=NN=CC=C2	251.1	2.15	A: 40580-53-8	medium
188	PJY	5QTH	O=C(C1=CC=CN=C1SC)N(C)C	196.3	0.66	B: 2799882	weak
189	PKO	5QTE	O=C(C1N(C(C2=CC=C(Cl)C=C2)=O)CSC1)O	271.7	2.20	A: 1042437-09-1	-
190	PK7	5QTF	NCC1=CC=C(F)C(Cl)=C1	159.6	1.95	A: 72235-56-4	-
191	PK4	5QTG	CNS(=O)(=O)C=1C=CC=C(C)C1F	203.2	1.87	B: 126782062	-

[1] clogP values obtained from ChemDraw.
[2] Classification of compounds in weak, medium and strong shifter on *T. brucei* FPPS according to [^{13}C, ^{1}H] -SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K. Compound **185** was not tested (n/a).

10.2.7 Synthesised fragments as part of this thesis

Table 40: Compound list of synthesised fragments.

Cmpd Nr.	SMILES code	Yield [%]	MW	clogP[2]	Prot. NMR[3]
LM1	C1=C(C=CC(=C1)OCCCN2C=CN=C2CNC3=CC=CC(=C3)Cl)Cl	3.4	376.3	4.95	medium
LM2	C1C(=CC3C(C=1)=NN(CCN2C=CN=C2)C=3)Cl	15.8	246.7	2.13	medium
LM3	C1C(=CC3=C(C=1)N(CCN2C=CN=C2)N=C3)Cl	7.6	246.7	2.13	medium
LM4	C1(=CNC2=C1C=CC=C2)CN4CCN(CC3C=C(Cl)C=CC=3)CC4	46.3	339.9	4.88	medium
LM5	C(N2CCN(CC1=CC=CC(=C1)Cl)CC2)C3=CNC4=C3C=NC=C4	22.9	340.9	3.92	weak
LM6	C(N2CCN(CC1=CC=CC(=C1)Cl)CC2)C3=CNC4=C3N=CC=C4	45.3	340.9	4.13	weak
LM7	C(N2CCN(CC1=CC=CC(=C1)Cl)CC2)C3=CNC4=C3C=C(C=C4)O	6.9	355.9	4.21	medium
LM8	C(N2CCN(CC1=CC=CC(=C1)Cl)CC2)C3=CNC4=C3C=CC=C4OC	46.1	369.9	4.90	medium
LM9	C(N2CCN(CC1=CC=CC(=C1)Cl)CC2)C3=CNC4=C3C=CC=C4Cl	27.1	374.3	5.76	precipitation
LM10[1]	C1=CC=C2C(=C1)NC=C2CN3CCC(CC3)NCC4=CC=CC(=C4)Cl	8.8	353.9	4.06	strong

[1] Ligand in PDB: LEZ, PDB ID 6SII

[2] clogP values obtained from ChemDraw.

[3] Classification of compounds in very weak, weak, medium and strong shifter on *T. brucei* FPPS according to [^{13}C, ^{1}H] -SOFAST-HMQC (32 scans) experiment with 30 µm *T. brucei* FPPS protein measured at 310 K.

10.3 Surface Plasmone Resonance

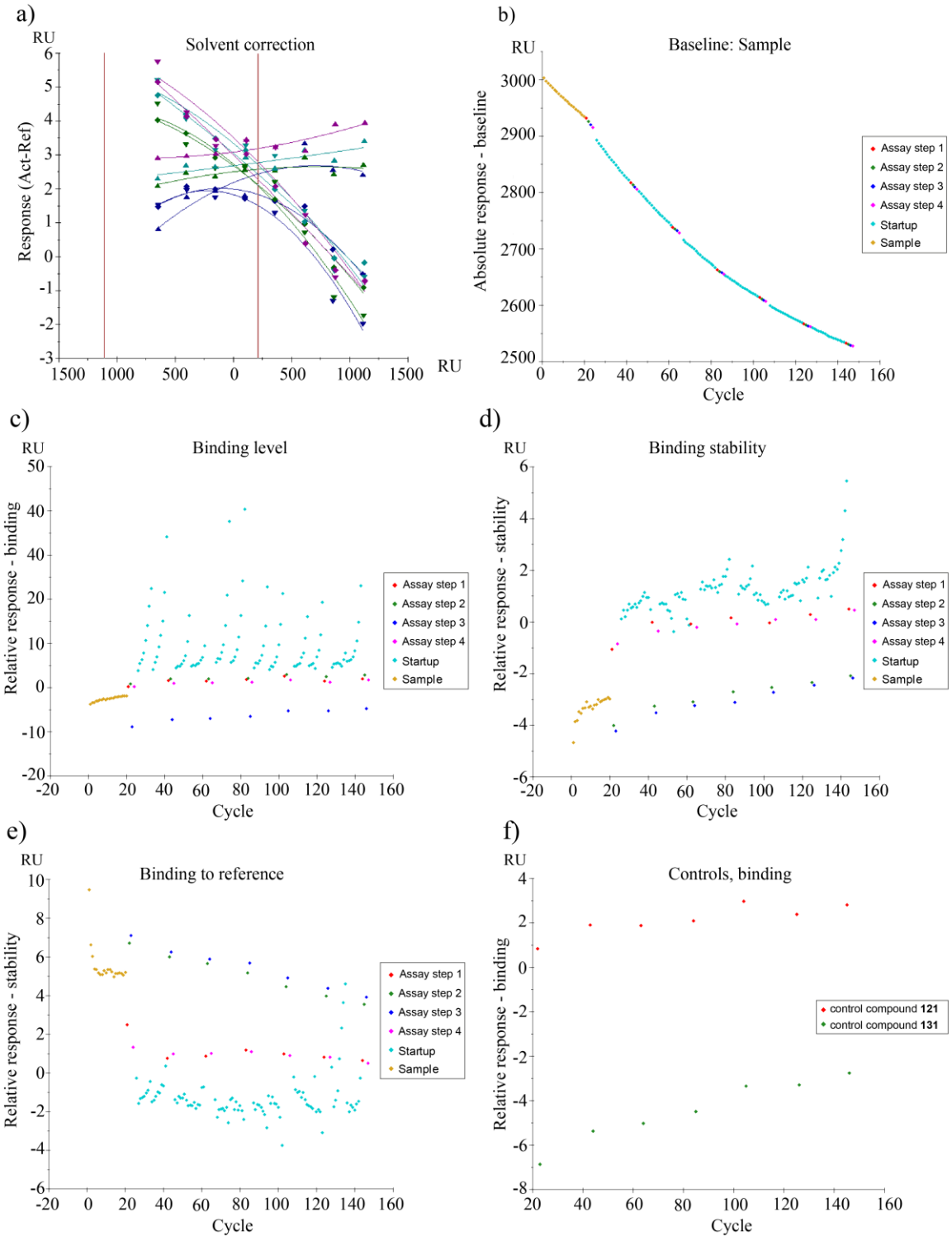


Figure 75: SPR plots of K_d determination of human allosteric binder. The following plots are shown (a) Solvent correction. **(b)** Baseline of immobilised T. brucei FPPS over measurement time**. (c)** Binding levels of compounds. **(d)** Binding stability of compounds **(e)** Binding to reference. **(f)** Binding of control samples to reference channel.

10.4 X-ray crystallography

10.4.1 X-ray data plots of Novartis data sets

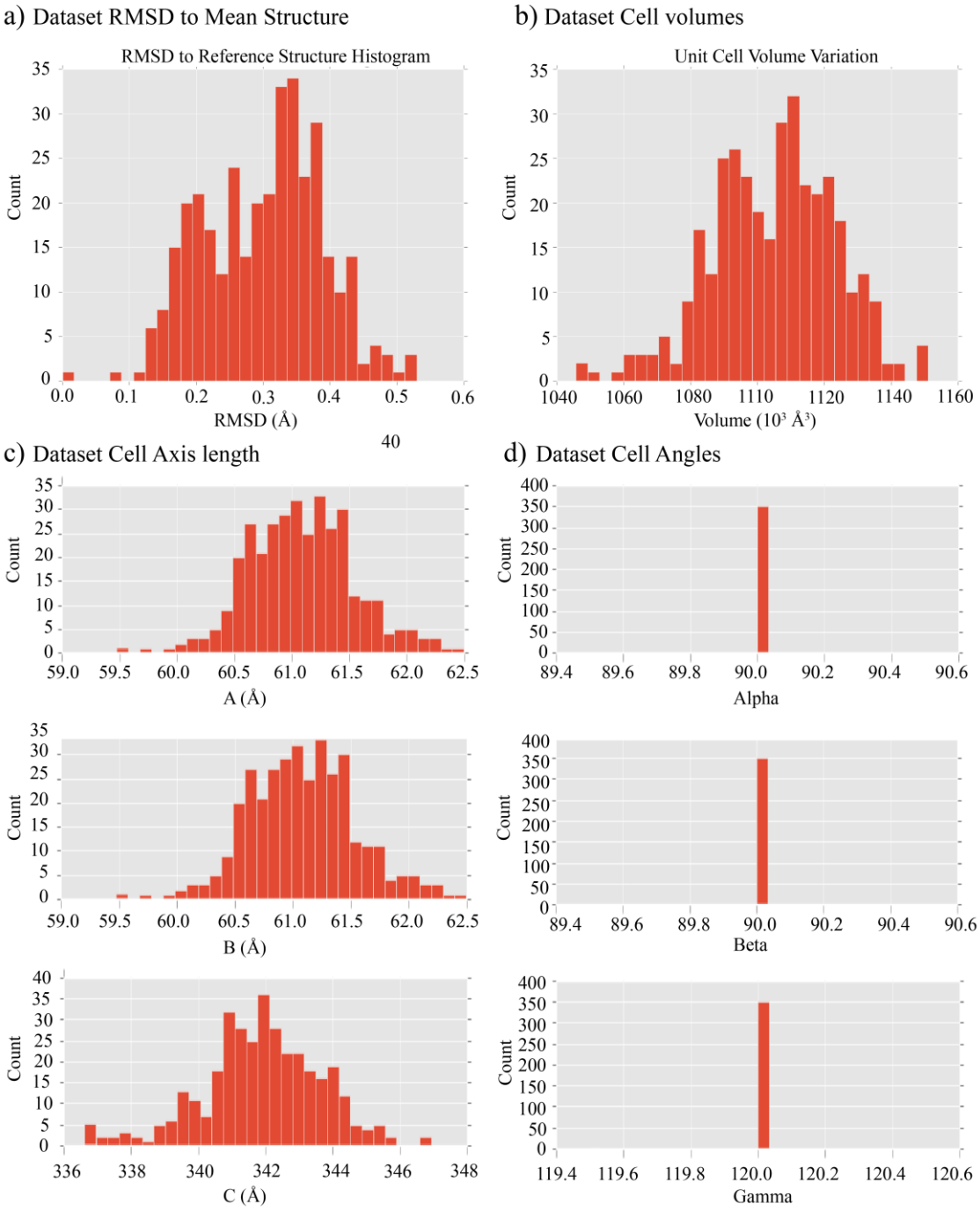


Figure 76: Summary graphs of Novartis X-ray data sets. The figure shows the following plots: **(a)** Dataset RMSD to Mean Structure with the RMSD (Å) on the x-axis and the number or structures on the y-axis. **(b)** Dataset Cell Volumes with the Volume (10^3 A^3) on the x-axis and the number of structures on the y-axis. **(c)** Dataset Cell Axis length with the A, B and C axis length on the respective x-axis and the number of structures on the y-axis. **(d)** Dataset Cell Angles with the α, β, γ angle on the respective x-axis and the number of structures on the y-axis.

10.4.2 X-ray data plots of iNEXT data

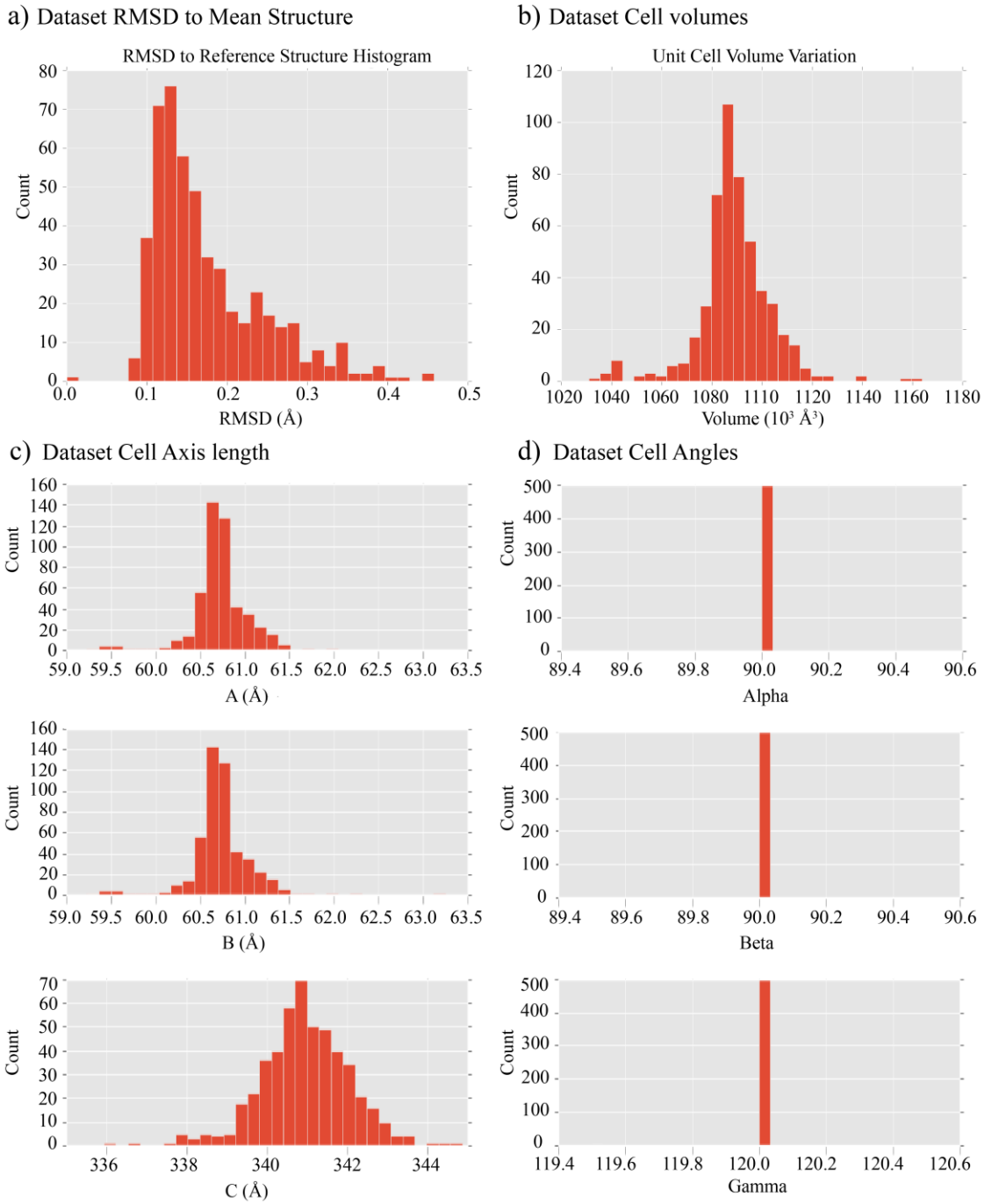


Figure 77: Summary graphs of iNEXT X-ray data sets. The figure shows the following plots. **(a)** Dataset RMSD to Mean Structure with the RMSD (Å) on the x-axis and the number or structures on the y-axis. **(b)** Dataset Cell Volumes with the Volume (10^3 A^3) on the x-axis and the number of structures on the y-axis. **(c)** Dataset Cell Axis length with the A, B and C axis length on the respective x-axis and the number of structures on the y-axis. **(d)** Dataset Cell Angles with the α, β, γ angle on the respective x-axis and the number of structures on the y-axis.

10.4.3 Data collection and refinement statistics of complexes crystallized at Novartis

Table 41: Crystallographic tables on data collection and refinement of apo protein and *T. brucei* FPPS in complex with human allosteric binders.

PDB code	6R37	6R36	6R38	6R39
Ligand in PDB	3N2	-	JQE	BFH
Compound Nr.	116	-	113	121
Data collection				
Diffraction source	SLS X10SA	SLS X10SA	SLS X10SA	SLS X10SA
Wavelength	0.99999 A	0.99974 A	0.99996 A	1.00001 A
Temperature	100	100	100	100
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M	Pilatus 6M
Indexing and scaling				
Unit cell parameters				
a,b,c (Å)	60.8, 60.8, 341.2	60.5, 60.5, 341.0	60.8, 60.8, 340.0	61.3, 61.3, 340.0
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Resolution range (Å)[1]	52.63 - 2.10 (2.14 - 2.10)	52.38 - 1.67 (1.88 - 1.67)	56.67 - 2.33 (2.37 - 2.33)	53.12 - 2.55 (2.60 - 2.55)
Total No. of reflections	383717	521843	318597	244264
No. of unique reflections	21086	28487	17098	13351
Multiplicity/data redundancy	18.2	18.3	18.6	18.3
Mean(I)/σ(I)	19.4	21.7	9.5	14.4
R_{meas}	0.094	0.083	0.337	0.152
Completeness spherical	91.3	63.8	100	100
Completeness ellipsoidal	-	92.7	-	-
Anomal. completeness (spherical)	90.7	60.8	100	99.9
Anomal. completeness (ellipsoidal)	-	92	-	-
Wilson B-factor	29.55	56.15	50.67	74.43
$CC_{1/2}$	1.000	1.000	0.998	1.000
Structure refinement				
R_{work} final	0.238	0.219	0.224	0.234
R_{free} final	0.256	0.243	0.249	0.257
No of non-H atoms				
Protein atoms	2628	2589	2643	2612
Heterogen atoms	18	15	19	20
Solvent atoms	0	81	1	2
RMS deviations				
Bond lengths (Å)	0.01	0.010	0.01	0.01
Bond angles (°)	0.93	1.00	1.00	1.03
Mean B value overall $Å^2$	92.83	48.95	84.44	127.85
Ramachandran plot				
Most favoured (%)	98.75	99.05	97.83	98.43
Outliers (%)	0.00	0.00	0.31	0.00
Allowed (%)	1.25	0.95	1.39	1.57
Clashscore	2.12	2.1	2.28	1.54
Matthews Coefficient	2.14	2.16	2.16	2.19
Solvent content	42.6	43.17	43.1	43.83

[1] Values in brackets are for outer resolution shell.

Table 42: Crystallographic tables on data collection and refinement of fragments idenitified in the Novartis fragment screen, one fragment analogue and a synthesised merged fragment.

PDB code	5QT5	5QT6	5QT7	5QT4
Ligand in PDB	PBV	JNE	PBY	P9Y
Compound Nr.	132	131	138	129
Data collection				
Diffraction source	SLS X10SA	SLS X10SA	SLS X10SA	SLS X10SA
Wavelength	0.99870 A	0.99979 A	0.99993 A	1.00000 A
Temperature	100	100	100	100
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M	Pilatus 6M
Indexing and scaling				
Unit cell parameters				
a,b,c (Å)	61.0, 61.0, 343.0	60.7, 60.7, 341.9	61.1, 61.1, 344.2	61.1, 61.1, 344.2
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Resolution range[1] (Å)	57.16 – 2.14	52.55 – 2.19	52.89 – 2.29	52.89 - 2.32
	(2.34 – 2.14)	(2.19 – 2.05)	(2.50 - 2.29)	(2.60 - 2.32)
Total No. of reflections	221565	217664	230096	157952
No. of unique reflections	12690	12355	12949	9284
Multiplicity/data redundancy	17.5	17.6	17.8	17.0
Mean(I)/σ(I)	14.7	11.0	15.8	26.4
R_{meas}	0.178	0.492	0.137	0.075
Completeness spherical	57.4	60.3	70.8	52.8
Completeness ellipsoidal	80.5	89.9	84.4	82.4
Anomal. completeness (spherical)	53.6	55.9	68.0	47.9
Anomal. completeness (ellipsoidal)	78.9	88.6	83	80.4
Wilson B-factor	31.3	32	51.7	57.8
$CC_{1/2}$	0.999	0.998	0.999	0.999
Structure refinement				
R_{work} final	0.237	0.250	0.235	0.283
R_{free} final	0.270	0.280	0.284	0.325
No of non-H atoms				
Protein atoms	2595	2595	2595	2595
Heterogen atoms	20	26	26	27
Solvent atoms	52	66	7	16
RMS deviations				
Bond lengths (Å)	0.002	0.002	0.002	0.001
Bond angles (°)	0.382	0.399	0.383	0.347
Mean B value overall $Å^2$	53.83	56.27	67.36	79.53
Ramachandran plot				
Most favoured (%)	98.11	98.11	99.05	98.11
Outliers (%)	0.00	0.31	0.00	0.00
Allowed (%)	1.89	1.89	0.95	1.89
Clashscore	2.89	2.69	2.89	3.85

[1] Values in brackets are for outer resolution shell.

Table 42: continued.

PDB code	5QTA	5QTB	5QT9	5QT8
Ligand in PDB	PJD	PJG	PJ4	PJ1
Compound Nr.	123	130	136	153
Data collection				
Diffraction source	SLS X10SA	SLS X10SA	SLS X10SA	SLS X10SA
Wavelength	1.00000 A	0.99993 A	0.99997 A	0.99997 A
Temperature	100	100	100	100
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M	Pilatus 6M
Indexing and scaling				
Unit cell parameters				
a,b,c (Å)	60.7, 60.7, 341.3	60.9, 60.9, 340.7	60.9, 60.9, 343.3	61.0, 61.0, 343.5
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Resolution range[1] (Å)	56.88 – 2.11 (2.37 – 2.11)	52.79 – 1.97 (2.11 – 1.97)	52.77 – 2.20 (2.48 – 2.20)	57.24 - 2.35 (2.64 - 2.35)
Total No. of reflections	252649	326470	216968	188854
No. of unique reflections	14308	17981	12144	10512
Multiplicity/data redundancy	17.7	18.2	17.9	18.0
Mean(I)/σ(I)	18.4	17.8	16.3	16.2
R_{meas}	0.114	0.143	0.129	0.156
Completeness spherical	63.0	63.9	59.4	62.1
Completeness ellipsoidal	90.6	83.3	87.8	89.9
Anomal. completeness (spherical)	59.1	61.1	55.1	57.6
Anomal. completeness (ellipsoidal)	89.6	82.5	86.4	88.5
Wilson B-factor	35.6	27.3	47.4	46.8
$CC_{1/2}$	1.000	0.999	1.000	0.999
Structure refinement				
R_{work} final	0.2537	0.2310	0.2295	0.2363
R_{free} final	0.2943	0.2718	0.2742	0.2724
No. of non-H atoms				
Protein atoms	2595	2595	2595	2595
Heterogen atoms	24	25	29	29
Solvent atoms	54	84	33	17
RMS deviations				
Bond lengths (Å)	0.001	0.004	0.002	0.002
Bond angles (°)	0.339	0.608	0.466	0.383
Mean B value overall $Å^2$	58.30	48.52	66.08	60.48
Ramachandran plot				
Most favoured (%)	98.42	98.74	97.79	97.48
Outliers (%)	0.31	0.00	0.00	0.00
Allowed (%)	1.58	1.26	2.21	1.42
Clashscore	2.89	2.50	2.88	3.08

[1] Values in brackets are for outer resolution shell.

Table 42: continued.

PDB code	5QTC	6SII
Ligand in PDB	-	LEZ
Compound Nr.	-	LM10
Data collection		
Diffraction source	SLS X10SA	SLS X10SA
Wavelength	1.00000 A	1.00002 A
Temperature	100	100
Detector	Pilatus 6M	Pilatus 6M
Indexing and scaling		
Unit cell parameters		
a,b,c (Å)	60.0, 60.0, 340.6	61.6, 61.6, 342.2
α, β, γ (°)	90, 90, 120	90, 90, 120
Space group	$P6_122$	$P6_122$
Resolution range[1] (Å)	55.77 – 1.65 (1.88 – 1.65)	57.04 – 2.33 (2.37 – 2.33)
Total No. of reflections	445510	326285
No. of unique reflections	24792	17624
Multiplicity/data redundancy	18.0	18.5
Mean(I)/σ(I)	20.3	15.1
R_{meas}	0.095	0.116
Completeness spherical	54.1	99.4
Completeness ellipsoidal	89.4	-
Anomal. completeness (spherical)	50.6	99.3
Anomal. completeness (ellipsoidal)	88.6	-
Wilson B-factor	27.1	79.3
$CC_{1/2}$	1.000	0.999
Structure refinement		
R_{work} final	0.211	0.248
R_{free} final	0.251	0.277
No. of non-H atoms		
Protein atoms	2585	2619
Heterogen atoms	11	31
Solvent atoms	145	20
RMS deviations		
Bond lengths (Å)	0.006	0.010
Bond angles (°)	0.831	1.04
Mean B value overall $Å^2$	44.31	121.9
Ramachandran plot		
Most favoured (%)	99.05	97.81
Outliers (%)	0.00	0.31
Allowed (%)	0.95	2.11
Clashscore	5.60	2.1

[1] Values in brackets are for outer resolution shell.

Table 43: Crystallographic tables on data collection and refinement of fragments idenitified in X-ray fragment screen at EMBL Grenoble.

PDB code	5QTK	5QTJ	5QTI	5QTH
Ligand in PDB	PJP	PJS	PJV	PJY
Compound Nr.	185	186	187	188
Data collection				
Diffraction source	ID30B	MASSIF-1	MASSIF-3	MASSIF-1
Wavelength	0.97625 A	0.96600 A	0.96770 A	0.96770 A
Temperature	100	100	100	100
Detector	Pilatus 6M	Pilatus 2M	EIGER X4M	Pilatus 2M
Indexing and scaling				
Unit cell parameters				
a,b,c (Å)	60.5, 60.5, 340.8	60.5, 60.5, 342.2	60.5, 60.5, 340.5	60.4, 60.4, 344.9
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Resolution range[1] (Å)	56.81 – 1.86	57.03 – 2.10	56.74 – 2.10	57.48 – 2.09
	(2.06 – 1.86)	(2.28 – 2.10)	(2.32 – 2.10)	(2.32 - 2.09)
Total No. of reflections	776652	629377	597677	604251
No. of unique reflections	22050	17890	16042	15971
Multiplicity/data redundancy	35.2	35.2	37.3	37.8
Mean(I)/σ(I)	14.3	21.7	15.6	18.9
R_{meas}	0.300	0.136	0.201	0.16
Completeness spherical	67.4	77.9	70.2	68.6
Completeness ellipsoidal	91.6	91.7	91.5	91.3
Anomal. completeness (spherical)	64.5	75.5	67.1	65.3
Anomal. completeness (ellipsoidal)	91	91	90.8	90.6
Wilson B-factor	20.4	34.1	35.9	23.4
$CC_{1/2}$	0.992	0.999	1.000	0.996
Structure refinement				
R_{work} final	0.253	0.242	0.243	0.243
R_{free} final	0.290	0.278	0.279	0.283
No. of non-H atoms	2786	2753	2713	2769
Protein atoms	2629	2629	2629	2623
Heterogen atoms	26	37	31	29
Solvent atoms	131	87	53	117
RMS deviations				
Bond lengths (Å)	0.003	0.002	0.001	0.003
Bond angles (°)	0.479	0.492	0.385	0.524
Mean B value overall $Å^2$	38.13	57.31	57.02	23.4
Ramachandran plot				
Most favoured (%)	98.45	98.14	98.45	97.51
Outliers (%)	0.00	0.00	0.00	0.00
Allowed (%)	1.55	0.86	1.55	2.49
Clashscore	3.24	3.42	4.05	3.05

[1] Values in brackets are for outer resolution shell.

Table 43: continued.

PDB code	5QTE	5QTF	5QTG	5QTD
Ligand in PDB	PKD	PK7	PK4	-
Compound Nr.	189	190	191	-
Data collection				
Diffraction source	MASSIF-1	MASSIF-1	ID30B	MASSIF-1
Wavelength	0.9677 A	0.9660 A	0.9677 A	0.9660 A
Temperature	100	100	100	100
Detector	Pilatus 2M	Pilatus 2M	Pilatus 6M	Pilatus 2M
Indexing and scaling				
Unit cell parameters				
a,b,c (Å)	60.6, 60.6, 341.2	60.6, 60.6, 341.3	60.4, 60.4, 344.9	60.5, 60.5, 340.1
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120	90, 90, 120
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$
Resolution range[1] (Å)	56.87 - 2.07 (2.26 - 2.07)	56.89 – 1.78 (2.01 – 1.78)	57.48 – 2.09 (2.32 – 2.09)	56.68 – 1.64 (1.86 – 1.64)
Total No. of reflections	629732	878882	604251	1001423
No. of unique reflections	17265	24054	15971	27755
Multiplicity/data redundancy	36.5	36.5	37.8	36.1
Mean(I)/σ(I)	14.6	20.3	18.9	24.0
R_{meas}	0.348	0.127	0.160	0.132
Completeness spherical	72.3	65.0	68.6	59.6
Completeness ellipsoidal	91.8	92.2	91.3	89.7
Anomal. completeness (spherical)	69.4	62.1	65.3	56.5
Anomal. completeness (ellipsoidal)	91.1	91.7	90.6	89.0
Wilson B-factor	23.1	30.7	38.4	23.7
$CC_{1/2}$	0.998	0.998	0.999	0.999
Structure refinement				
R_{work} final	0.255	0.253	0.274	0.253
R_{free} final	0.283	0.293	0.307	0.298
No. of non-H atoms				
Protein atoms	2754	2754	2720	2764
Heterogen atoms	34	27	30	17
Solvent atoms	101	98	61	118
RMS deviations				
Bond lengths (Å)	0.002	0.002	0.002	0.007
Bond angles (°)	0.383	0.481	0.391	0.851
Mean B value overall $Å^2$	44.73	53.35	58.10	43.71
Ramachandran plot				
Most favoured (%)	98.44	97.83	99.07	98.14
Outliers (%)	0.00	0.00	0.00	0.00
Allowed (%)	1.56	2.17	0.93	1.86
Clashscore	2.29	3.24	3.05	7.06

[1] Values in brackets are for outer resolution shell.

10.5 Different density maps

10.5.1 Human allosteric binders on *T. brucei* FPPS

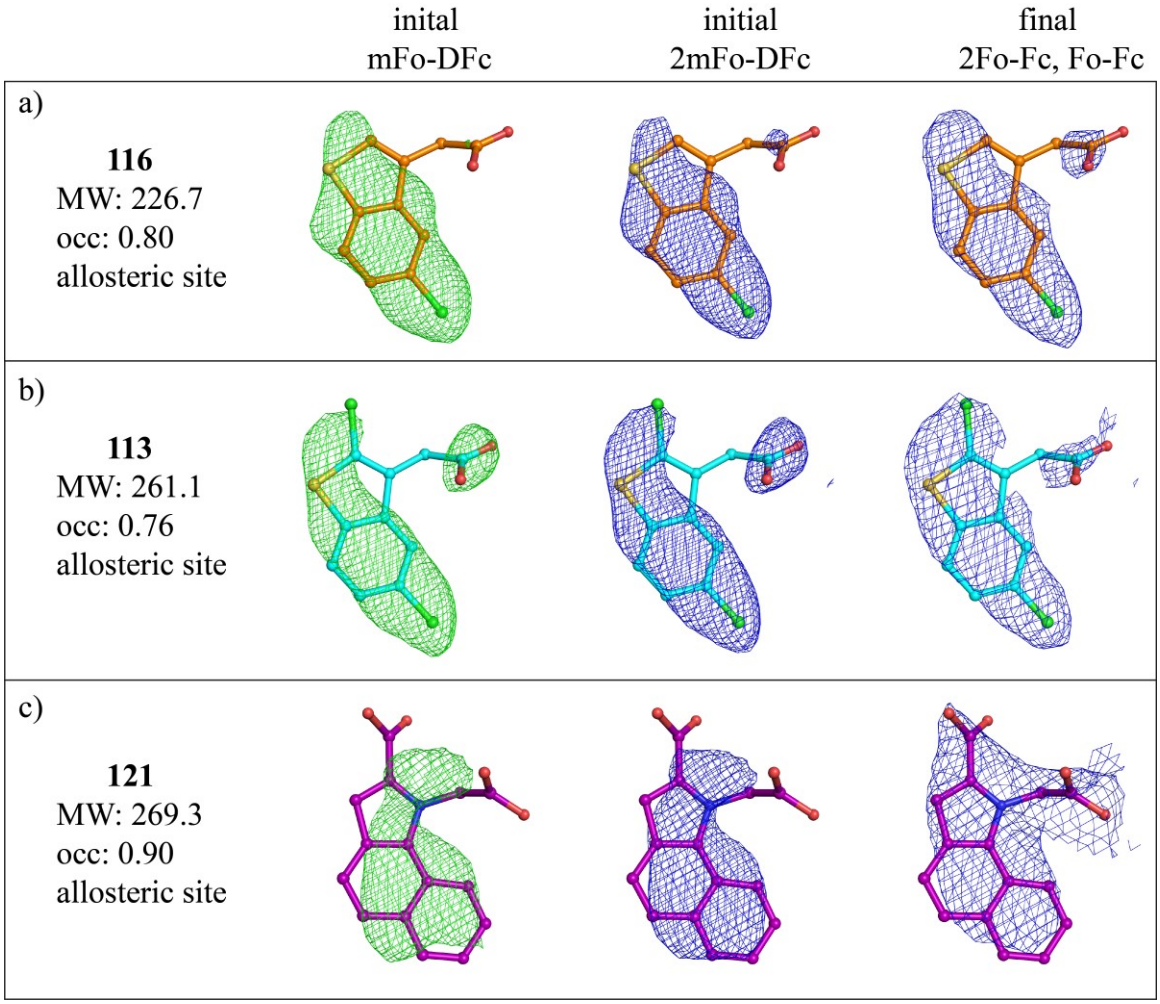


Figure 78: Different density maps of human allosteric binders on *T. brucei* FPPS. In **(a)**, **(b)**, and **(c)** different density maps of compounds **116**, **113**, and **121** are shown. The molecular weight (MW), occupancy (occ.) and binding site is given. From left to right: initial mFo-DFc difference electron density map contoured at 3 σ, initial 2mFo-DFc electron density map contoured at 1 σ, final refined 2Fo-Fc electron density map contoured at 1 σ overlaid with final refined Fo-Fc map contoured at 3 σ.

10.5.2 *T. brucei* FPPS hits of Novartis core library

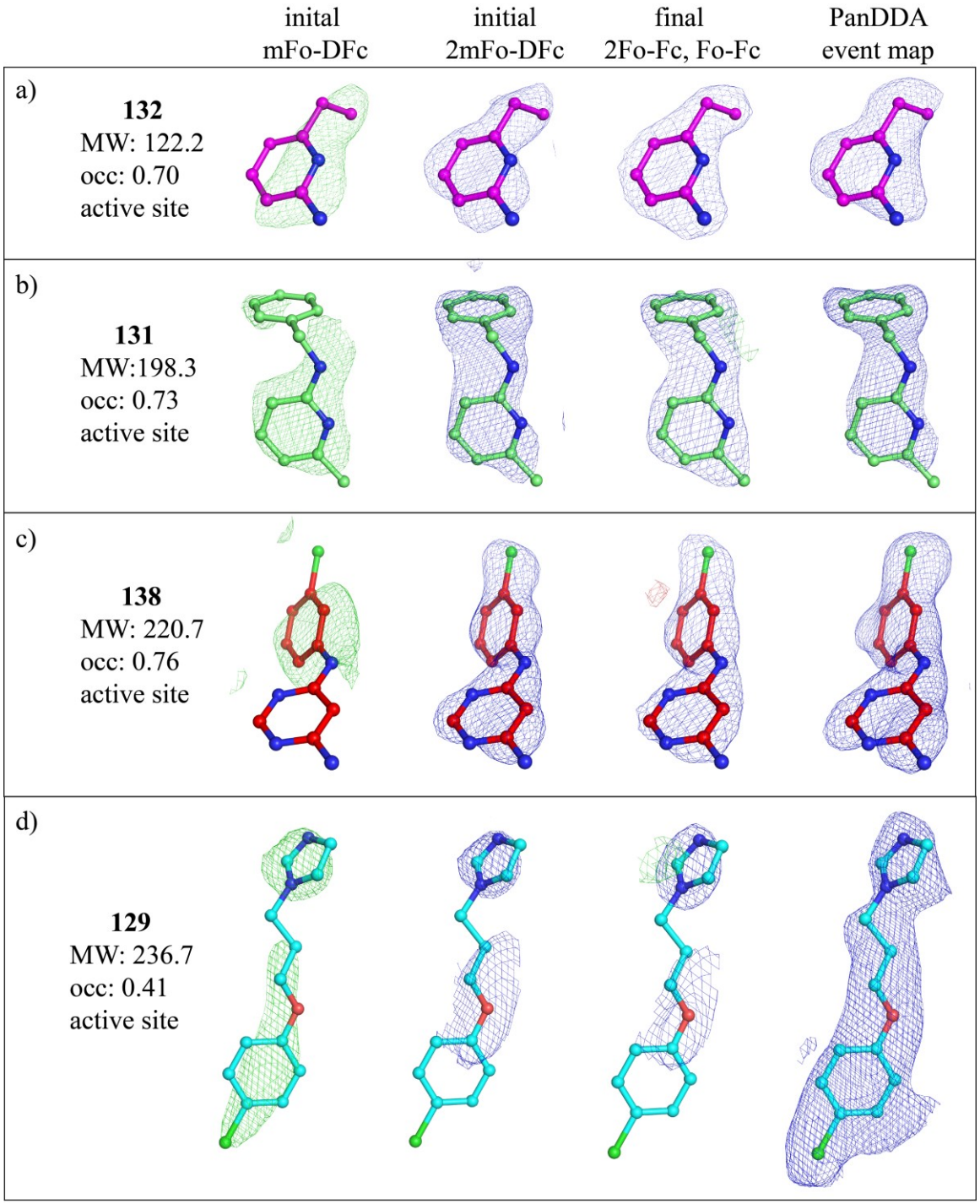


Figure 79: Different density maps of fragment hits identified in the NMR fragment screen. In **(a)**, **(b)**, **(c)**, **(d)**, **(e)**, **(f)**, and **(g)** different density maps of compounds **132**, **131**, **138**, **129**, **123**, **130**, and **136** are shown. The molecular weight (MW), occupancy (occ.) and binding site is given. From left to right: initial mFo-DFc difference electron density map contoured at 3 σ, initial 2mFo-DFc electron density map contoured at 1 σ, final refined 2Fo-Fc electron density map contoured at 1 σ overlaid with final refined Fo-Fc map contoured at 3 σ and on the right: PanDDA difference density map contoured at 2 σ.

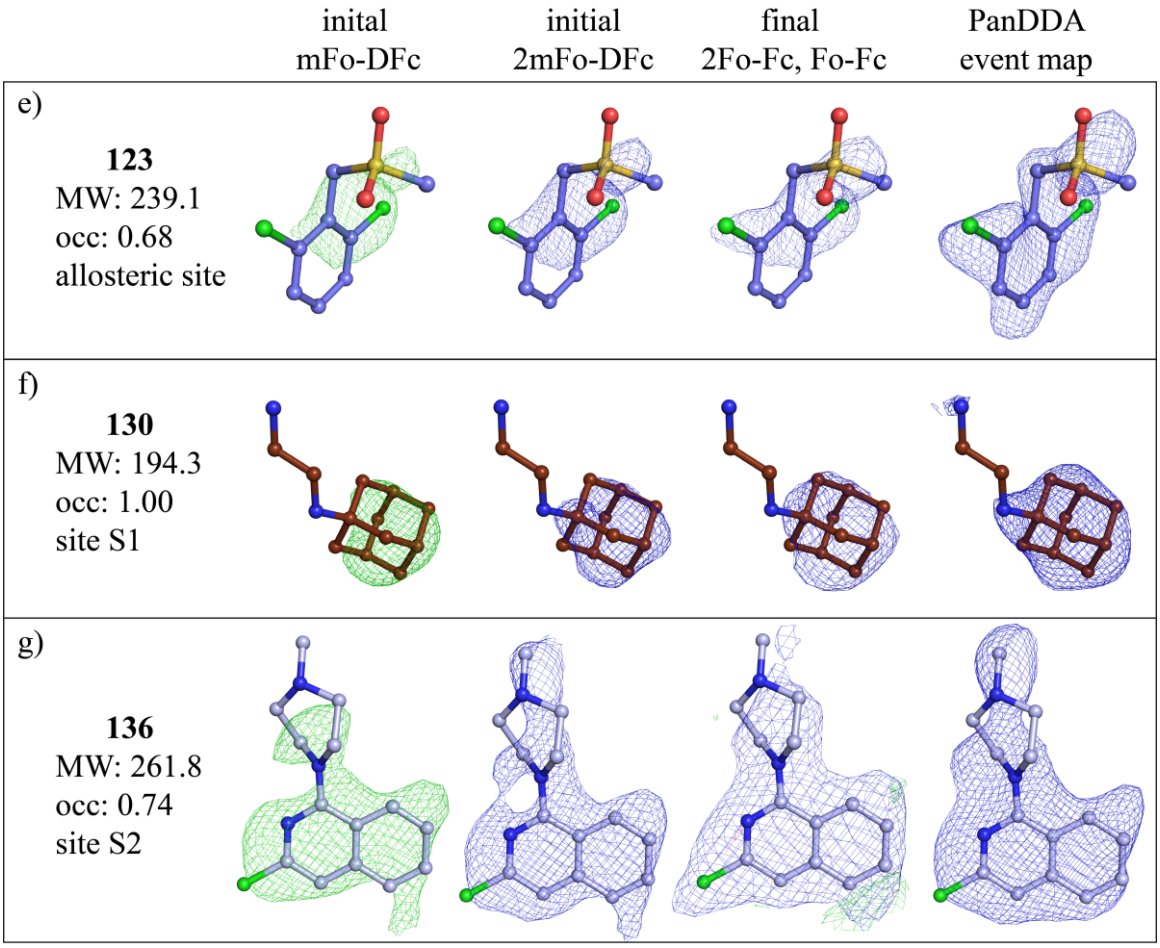


Figure 79: continued.

10.5.3 Fragment analogue and synthesised fragment merge

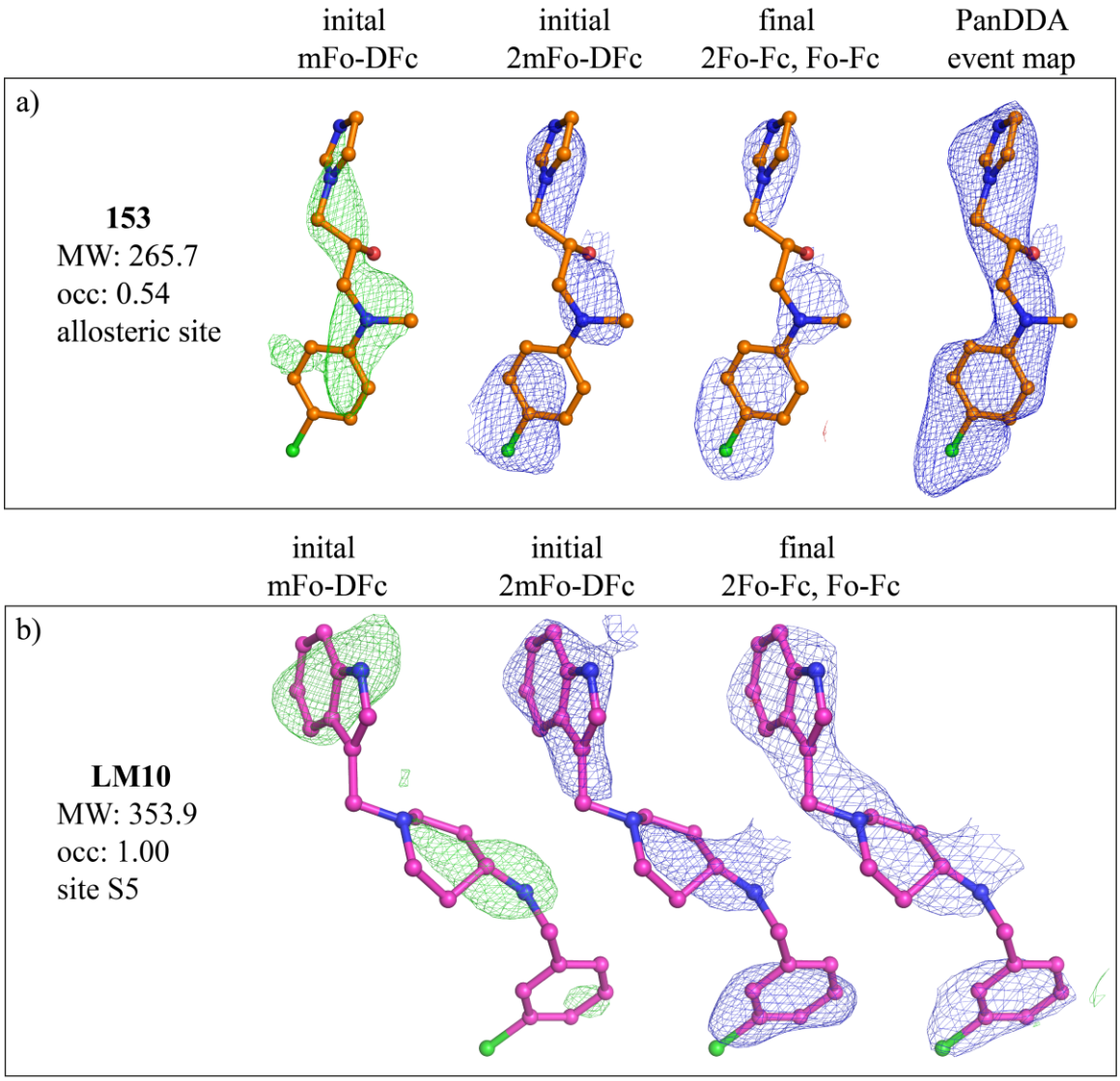


Figure 80: Different density maps of fragment analogue and synthesised merged compound . In **(a)** and **(b)** different density maps of fragment analogue **153**, and synthesised merged compound LM10 are shown. The molecular weight (MW), occupancy (occ.) and binding site is given. From left to right: initial mFo-DFc difference electron density map contoured at 3 σ, initial 2mFo-DFc electron density map contoured at 1 σ, final refined 2Fo-Fc electron density map contoured at 1 σ overlaid with final refined Fo-Fc map contoured at 3 σ and on the right: PanDDA difference density map contoured at 2 σ. The PanDDA event map is only shown for compound **153**.

10.5.4 Fragments identified in X-ray fragment screen at EMBL Grenoble

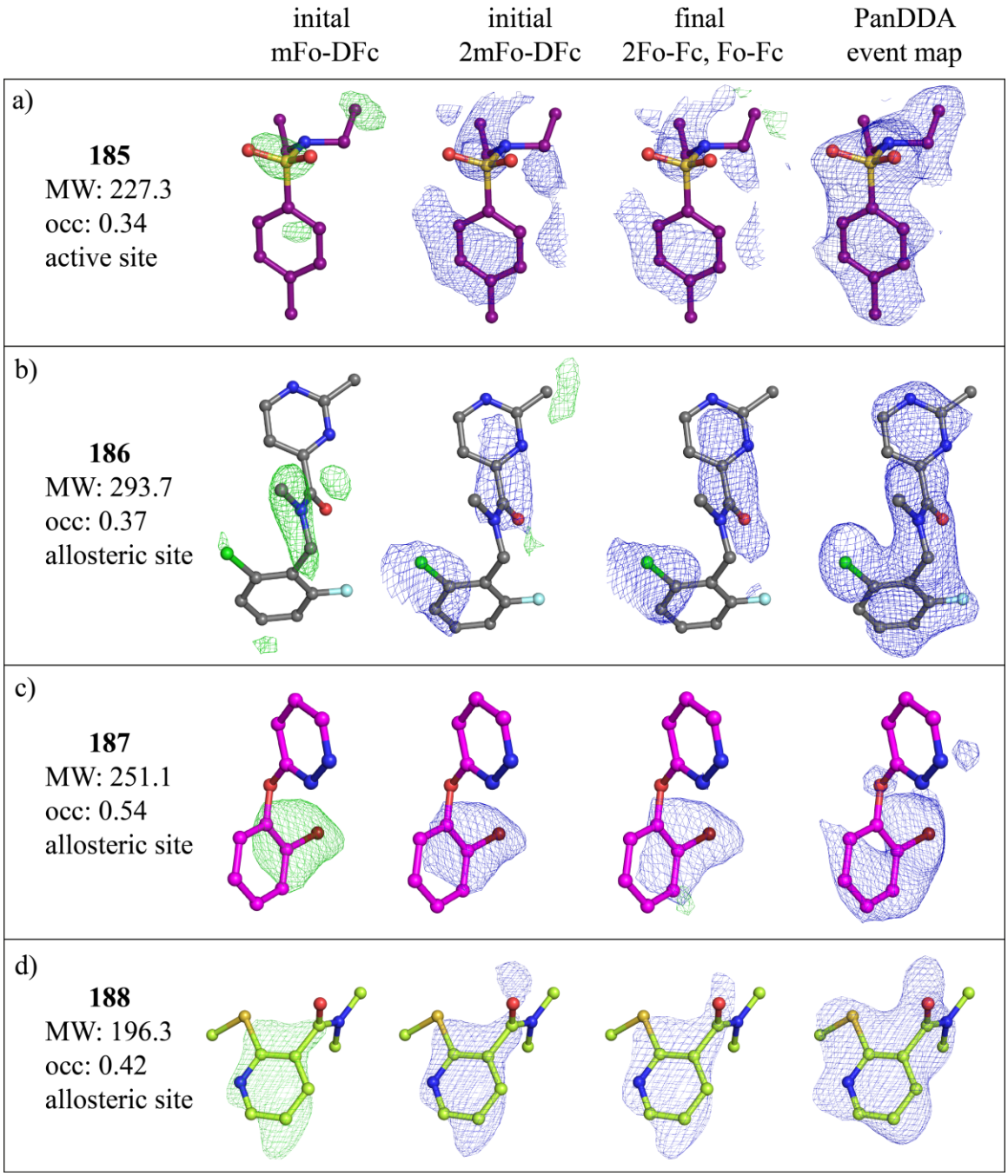


Figure 81: Different density maps of fragment hits identified in the X-ray fragment screen at EMBL Grenoble. In **(a)**, **(b)**, **(c)**, **(d)**, **(e)**, **(f)**, and **(g)** different density maps of compounds **185**, **186**, **187**, **188**, **189**, **190**, and **191** are shown. The molecular weight (MW), occupancy (occ.) and binding site is given. From left to right: initial mFo-DFc difference electron density map contoured at 3 σ, initial 2mFo-DFc electron density map contoured at 1 σ, final refined 2Fo-Fc electron density map contoured at 1 σ overlaid with final refined Fo-Fc map contoured at 3 σ and on the right: PanDDA difference density map contoured at 2 σ.

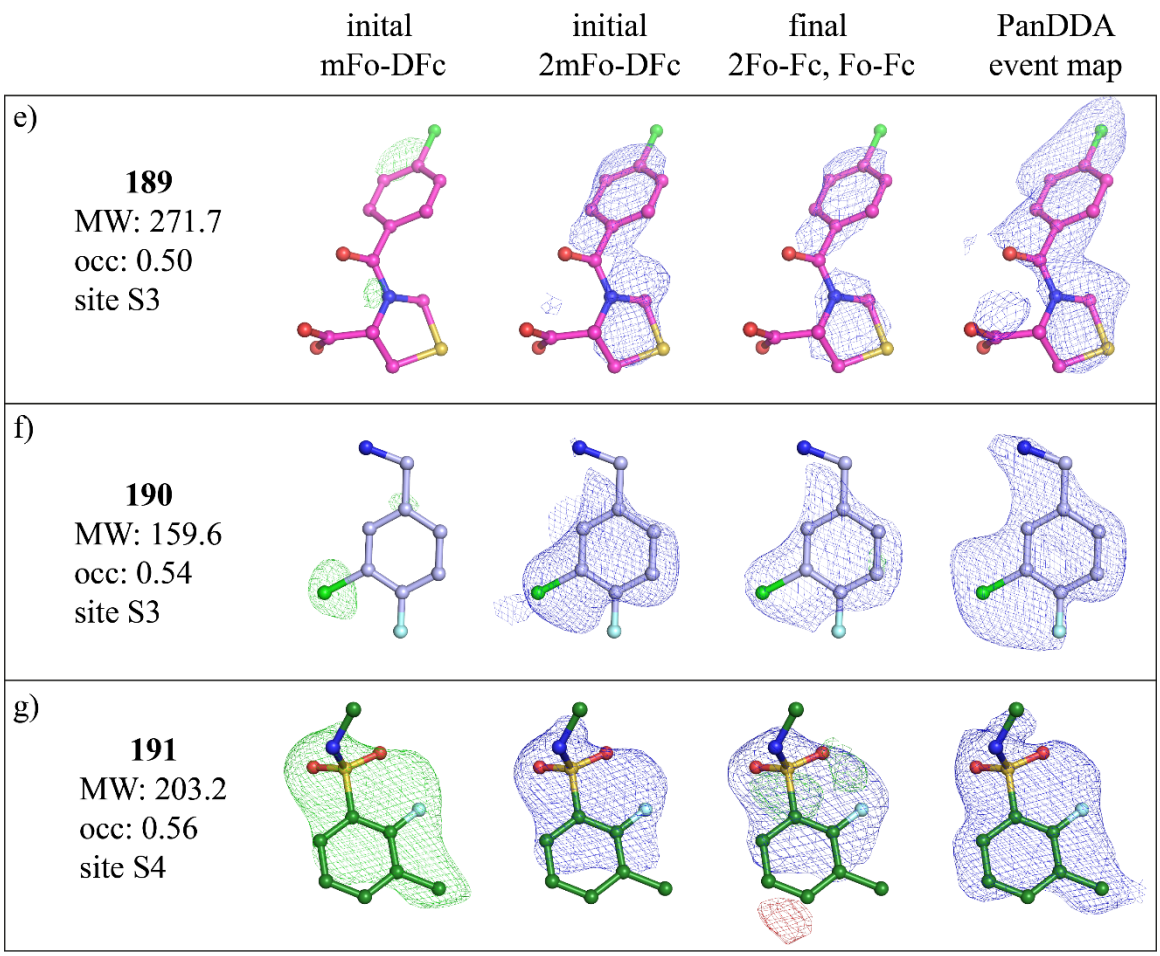


Figure 81: continued.

10.6 ConSurf analysis

For the identification of conserved regions in the *T. brucei* FPPS protein a ConSurf run was performed (consurf.tau.ac.il) on the *T. brucei* FPPS (PDB: 2I19). No multiple sequence alignment was provided and the homology search algorithm HMMER was applied with the default parameter settings performing one iteration and using a E-value cut-off of 0.001. The search was performed on the "clean UniProt" protein database with a manual selection option of sequences. CLUSTALW was chosen as an alignment method and the run was submitted. In total, 4750 entries were found and 497 UniProt entries with the following titles were retrieved: farnesyl pyrophosphate synthase, bifunctional GPPS/FPPS, Erg20p. The entries were manually selected for the alignment (Figure **82**). Conservation scores were calculated using the empirical Bayesian paradigm[549] with the default settings for the evolutionary substitution model.

001 A0A0C6AJW1
002 A0A0C5ZRH0
003 A0A0D4T1R4
004 A0A0C6B0S8
005 A0A0C6B0S1
006 A0A0C6APL4
007 A0A0D4T9K5
008 A0A0C6A2A2
009 A0A0C5ZE82
010 A0A0C5ZIR3
011 A0A0D4T1V4
012 A0A0D4SZ21
013 A0A0C6A289
014 A0A0C5ZG15
015 A0A0C6AXZ4
016 A0A0C6AIK9
017 A0A0C5ZIF8
018 W7RIX2
019 A0A0D4STG5
020 A0A0C5ZNX3
021 A0A0C5ZS50
022 A0A0C6AD69
023 A0A0D4TC48
024 A0A0D4TAN5
025 A0A0D4T3M8
026 A0A0C5Z977
027 B3LPV6
028 A0A0C6AI01
029 A0A0C6A8Q9
030 A0A0C6AEQ1
031 A0A0C6AMJ1
032 P08524
033 A0A0D4T5P2
034 A0A0D4T0Z7
035 A0A0C5YXK2
036 A0A0D4T070
037 A0A0C6AAQ9
038 A0A0C5ZDB1
039 A0A0C6B5R6
040 A0A0C5ZA99
041 A0A0C6AK49
042 A0A0D4T8Q2
043 A0A0C5ZW54
044 A0A0D4T9J8
045 A0A0D4T202
046 A0A0C6AWR8
047 A0A0C6AAY1
048 A0A0D4T364
049 A0A0C5ZB93
050 A0A0D4SUD8
051 A0A0C6AYV7
052 A0A0C5ZHE7
053 A0A0D4SWM4
054 A0A024XLG4
055 A0A0C6AAD8
056 A0A0C6A7U0
057 A0A0D4T8H1
058 A0A0D4SQJ9
059 A0A0C5ZUW7
060 A0A0C6AVW0
061 C8ZB52
062 A0A0C6AX30
063 A0A0D4T5W4
064 A0A0C6AX22
065 A0A0C5ZXS0
066 A0A0C5ZFD4
067 A0A0C6AUV5
068 A0A0C6A5W6
069 A0A0C6A7U4
070 A0A0C6AD58
071 A0A0C5ZI00
072 A0A0C5ZM65
073 A0A0C6A221
074 A0A0C6A0E4
075 A0A0C6AMN5
076 A0A0C5ZT97
077 A0A0C6B4Q8
078 A0A0C6ABW1
079 A0A0C5ZK62
080 A0A0C6AI31
081 A0A0D4SVS4
082 A0A0C6AHB7
083 A0A0D4SUK0
084 A0A023ZLR8
085 A0A0D4SSY8
086 A0A0C5Z7T5
087 A0A0C6AM07
088 A0A0C5Z6W2
089 A0A0C5ZT69
090 A0A0C6A2Y9
091 A0A0C6AF79
092 A0A0C5ZN12
093 W7Q0N2
094 A0A0C6AKU1
095 A0A0C5ZWM7
096 A0A0C6B034
097 A0A0C6AAU2
098 A0A0C5IJX5
099 A0A0C5ZY10
100 N1P1X6
101 W0T5C0
102 A0A090BIU1
103 P49349
104 C4YIE1
105 A5DX72
106 C4QY32
107 J9VSY8
108 A0A0D2LBX7
109 D8Q297
110 A0A0D7A366
111 R7SLM5
112 R7SSD6
113 R7SLS1
114 B4YA15
115 S7RL38
116 A0A0H2SDP8
117 Q75ND9
118 M7WUP1
119 A0A109FE05
120 U5II3V1
121 A0A0C4MWF7
122 X5DUM0
123 W9KQW3
124 W9MXS4
125 X0MTQ2
126 X0DCQ3
127 X0K6U8
128 X0HYK3
129 W9IYA0
130 X0GP72
131 N1REE7
132 A0A0D2XDQ9
133 W9PWY9
134 X0AC92
135 N4TNY5
136 Q92235
137 W7M8T5
138 V6RG22
139 C7YU73
140 G3J9F4
141 J4KQ38
142 A0A0D9P3P3
143 A0A0D9Q605
144 E9E7Y0
145 A0A0A1T3M9
146 L2G676
147 N4W485
148 G2XEE5
149 K7XUM4
150 A0A024S6R9
151 L7JEL9
152 L7HNR0
153 G4MLE9
154 R8BH59
155 J3P125
156 A0A0C4DPT6
157 S3BVU7
158 U7PSZ8
159 F0XRU4
160 G0S7R8
161 G2QUM8
162 G2Q9I5
163 G4UER6
164 Q92250
165 F8MF51
166 B5MEL7
167 M7TRE8
168 M7UUX8
169 Q0ZS32
170 A7EQQ6
171 K1WQV7
172 A0A132B2N0
173 A0A0C3H8N0
174 A0A061IIII8
175 W3X3S4
176 R1EUF1
177 A0A074WNM1
178 F9XIB9
179 A0A0D2B757
180 L8FNQ0
181 A0A0J5PUK7
182 B0Y2C6
183 A1D067
184 A1C591
185 Q0CAS4
186 G7XBA4
187 A0A0U5GM00
188 F0UP55
189 C6HFF5
190 C0NTA4
191 A6QSQ7
192 F2TLI5
193 T5BTG5
194 C5GIE6
195 C5JFT0
196 V9DD28
197 A0A0D2E370
198 A0A0D1ZPY7
199 A0A0D1YTY8
200 A0A0D2BYX2
201 A0A0D2BGH7
202 A0A0D2C7W1
203 A0A0D1WRX1
204 W2S9B3
205 U1HZZ1
206 B6QF61
207 H9U991
208 B2WM73
209 O14230
210 Q8NJ20
211 S2JNS8
212 A0A0C9MTL9
213 H9CDX3
214 A0A0B7NLL9
215 A0A0C7B597
216 A0A0A1N7Q3
217 I1CNE1
218 A0A077WI13
219 A0A0D2X444
220 A0A0L0G965
221 Q4JHN6
222 G0T3G2
223 S5W638
224 A0A0H4AYG9
225 D9IXA9
226 Q4TTY7
227 H2EV94
228 Q5S267
229 A0A075B8K1
230 Q94IE8
231 M9TF24
232 A0A0F6PMF9
233 S4S9J9
234 D1M868
235 A0A0K0PTV0
236 Q9AR37
237 Q94G65
238 W9RMI8
239 H2ER25
240 E3W209
241 A0A088CNE3
242 A0A0B0PHA5
243 A0A0K1II354
244 A0A0B2S247
245 A0A0B2PYM7
246 D7NM49
247 A0A075EAM3
248 P49351
249 D3K2X4
250 P49352
251 Q8I0D8
252 A0A0A6ZDM8
253 A0A0B0PT58
254 W5ZRY0
255 O04838
256 A0A061GHH9
257 A0A061GGK8
258 D7RII1
259 X2D2M7
260 A0A140GWW1
261 A0A140GWW2
262 A0A0S3J2X4
263 Q8L7F4
264 A6N2H2
265 A9ZN19
266 A0A0II3WME2
267 B9T7W9
268 A0A140GWW4
269 A9PFL3
270 C0LSI7
271 A0A0F6PN04
272 A0A097RN62
273 Q1XIT3
274 F8RWH1
275 E6Y2M9
276 Q94F74

Figure 82: UniProt codes of sequences used for ConSurf analysis.

277 A0A0U3B401
278 A0A088B2N2
279 W5S157
280 G0TA25
281 J9XLH1
282 J9XH87
283 A0A0A1E5Q0
284 A0A0G3BFY6
285 K7NBV9
286 A0A076JDS1
287 Q8RVK7
288 F8R7Z1
289 W5VY26
290 A0A0K1H3L0
291 D7UAM1
292 Q56CY4
293 F6KUJ4
294 Q94EW0
295 A0A0V0HWG6
296 W8SI43
297 A0A0H5AZB7
298 C8CJE5
299 O65004
300 Q43315
301 F4JNF1
302 V4LYG7
303 D7MPY7
304 Q09152
305 V4LC14
306 A0A087GN96
307 X2L3Q6
308 A7L691
309 K7W9P4
310 L7RFK0
311 K7WCJ2
312 K7X481
313 M4I1V5
314 Q9SYX3
315 Q9ZPJ3
316 E2D028
317 Q7XYS9
318 P49350
319 K7W5B2
320 O24241
321 K7WCJ4
322 O24242
323 O64905
324 A0A109QM03
325 A0A024BUG6
326 K7W5B4
327 K7WQ53
328 K7WCJ7
329 K7W5B9
330 K7WQ55
331 Q7XYT0
332 K7W9P6
333 K7X489
334 K7W9Q1
335 K7X485
336 A0A140GWW0
337 A0A140GWW3
338 B9S9Y3
339 A0A061F3M2
340 Q8GTG0
341 Q9FXR9
342 A0A0K9R9T1
343 Q8RVQ7
344 V5LFZ7
345 F2YI32
346 E7EJI5
347 E7EDU5
348 K7WTR3
349 V9QM18
350 I7FHQ2
351 C6KH59
352 B8XPY7
353 B1A9K8
354 C0PQT1
355 A0A077HCY1
356 Q5YJK5
357 Q2VY68
358 J9Q908
359 P49353
360 Q548K3
361 K7USV3
362 M7YJA3
363 L7RXA0
364 L7RYL1
365 L7S310
366 F2D4M6
367 M8CZ52
368 Q9XEN0
369 A0A0E0L589
370 B6TAB8
371 B6SM90
372 K7V8H3
373 C5XHL1
374 A0A0D9YDR4
375 A0A0E0FRU0
376 O04882
377 M7ZE00
378 L7RYK4
379 L7PDB0
380 L7PE91
381 R7W5G1
382 B8A815
383 I7C7X2
384 I7CL89
385 I6TIIC7
386 D8T6V9
387 B8CD36
388 A0A061RPF1
389 A0A061R8F8
390 E1Z2N1
391 A8IX41
392 C1E3K6
393 A0A0L0D3B5
394 A0A0P5TMM9
395 A0A0P5KKW5
396 A0A0P5D1K2
397 A0A0P5X624
398 A0A0N8DBT0
399 Q9NH03
400 F0Z8Y9
401 F2DL53
402 B7GA81
403 A0A0G2T690
404 D8LGP5
405 D0NFW9
406 H3HB70
407 A0A067C844
408 F0WZM9
409 F0WUD3
410 Q1X875
411 Q1X876
412 B1PI49
413 A9Z1D3
414 C7EY43
415 C7EY44
416 A9Z1D4
417 A0A067RTI2
418 W4W3D3
419 L9KIW6
420 S7PKH9
421 Q8WMY2
422 Q0VCD7
423 Q2NL07
424 L8IDK5
425 K9K3N0
426 S9XHJ7
427 D0G6X4
428 G7NV64
429 F7FI27
430 P14324
431 H2Q079
432 K7AMG1
433 F7GUQ3
434 B0CM97
435 U6CPE7
436 M1EQP2
437 Q920E5
438 Q5M8R9
439 Q4FJN9
440 A0A0P6DDG6
441 F1LND7
442 A0A0G2JXT3
443 P05369
444 G5ARB7
445 A0A091CYM3
446 A0A0F7ZE20
447 T1E6J7
448 A0A0B8RUM7
449 P08836
450 A0A091FUY1
451 M7ATH9
452 Q4V916
453 Q5XJ03
454 A0A0F8ACB7
455 W8B5K5
456 A0A0N1IHS1
457 A5A7A5
458 Input_pdb_SEQRES_A
459 Q86C09
460 C9ZSP7
461 Q57WF1
462 I3QF12
463 I3QF21
464 I3QF05
465 I3QF18
466 I3QEZ6
467 I3QF06
468 I3QF09
469 G4WJU4
470 I3QEZ8
471 I3QF15
472 I3QF10
473 I3QF24
474 Q8WS26
475 Q95WL3
476 I3QF22
477 I3QF02
478 I3QF19
479 Q8WS25
480 I3QF11
481 G4WJU5
482 I3QF17
483 I3QEZ9
484 I3QF07
485 I3QF08
486 I3QF25
487 I3QEZ7
488 I3QF23
489 I3QF00
490 G0TYE1
491 E9BFY6
492 Q0GKD7
493 E9AII04
494 Q4QBL1
495 E9AVW8
496 A4IICI8
497 A0A088RT42

Figure 82: continued

List of scientific contributions

Poster:

- Muenzker L., Petrick J., Klebe G., Marzinzik A., Jahnke W., Targeting *Trypanosoma brucei* FPPS by fragment-based drug discovery, 9th International Conference on Structural Biology, 18th-20th September 2017, Zurich, Switzerland.
- Muenzker L., Petrick J., Klebe G., Marzinzik A., Jahnke W., Fragment-based discovery of novel active and allosteric site binders of *T. brucei* farnesyl pyrophosphate synthase, 3rd Integrative Structural Biology School, 16th – 21st July 2018, Institut Pasteur, Paris, France.
- Muenzker L., Petrick J., Klebe G., Marzinzik A., Jahnke W., Fragment-based discovery of novel active and allosteric site binders of *T. brucei* farnesyl pyrophosphate synthase, 7th RSC-BMCS Fragment-based Drug Discovery meeting, 24th – 26th March 2019, Cambridge, United Kingdom.

Presentations:

- Muenzker L., Fragment-based discovery of novel active and allosteric site binders of *T. brucei* farnesyl pyrophosphate synthase, Fragment-based Lead Discovery Conference, 7th – 10th October 2018, San Diego, United States of America.
- Muenzker L., Fragment-based discovery of novel active and allosteric site binders of *T. brucei* farnesyl pyrophosphate synthase, 7th RSC-BMCS Fragment-based Drug Discovery meeting, 24th – 26th March 2019, Cambridge, United Kingdom, Flash talk.

Poster presentation:

- Petrick, J., Muenzker, L., Klebe, G., Marzinzik, K., Jahnke, W., Targeting *T. cruzi* FPPS by fragment based drug discovery, Revolutions in Structural Biology: Celebrating the 100th Anniversary of Sir John Kendrew, 16th – 17th Nov 2017, EMBL Heidelberg, Germany.

Acknowledgements

There is a long list of direct and indirect contributors to this work, and I wish to acknowledge them all. First of all, I would like to thank my supervisor Dr. Wolfgang Jahnke for his continuous support and encouragement during my PhD studies and for regular meetings with stimulating discussions and guidance which helped me to push the project further. Thanks for sharing your knowledge to help me grow in the last three years. Thanks a lot also to my second supervisor at Novartis, Dr. Andreas Marzinzik, for his positive energy and motivation and especially his input on the synthesis. Furthermore, I wish to thank my academic supervisor, Prof. Gerhard Klebe, for his constant interest and for his insight and valuable feedback to improve this thesis. I would also like to thank my committee members Prof. Carsten Culmsee and Prof. Moritz Bünemann for taking their time to evaluate my thesis.

This thesis would not have been possible without the support, patience and guidance of many colleagues at the Chemical Biology and Therapeutics (CBT) and the Global Discovery Chemistry (GDC) group at NIBR. I am very grateful to Felix Freuler, Simon Hänni, Julia Klopp, Cecile Delmas, Dr. Alvar Gossert and Valentina Tundo for their help and support on the plasmid preparation, protein expression and purification. In addition to this, I would like to thank Daniela Scherrer-Becker, Magdalena Maschlej and Rene Hemmig for their support whenever needed.

Special thanks to Dr. Lukasz Skora and Dr. Alvar Gossart, who helped to immerse in the NMR field and who contributed significantly to setup my first experiments. Furthermore, thanks to Chrystèlle Henry, Andrea Wiget, Anke Blechschmidt and Nicola Aubin for answering any question related to NMR that came up during this work.

In addition I would like to thank Dr. Christian Bergsdorf for his valuable guidance on performing SPR experiments. Thanks for answering all the questions on the analysis and for helping to understand the data.

I would also like to express my sincere gratitude to Gabriele Rummel, Sylvie Lehmann, Frederic Villard and Aude Izaac for introducing me into the practice of crystallography and Gabriele Rummel especially for helping me to set up my very special crystallization plates. Additionally, I would like to thank Dr. Jean-Michel Rondeau, Dr. Markus Krömer and Dr. Carien Dekker for their help on the X-ray data processing and for providing an introduction on the background of the method. I am particularly grateful to Dr. Christian Schleberger for his patient support and for answering any question that I had concerning PanDDA analysis and data submission. I have greatly benefited from your knowledge and you were always willing to help me. I wish to thank Rainer Wilcken for performing docking experiments and for his modelling input and Sébastien Ripoche for his support on the synthesis of the compounds in this thesis.

I also wish to thank David Papale and Adriano Marra for all the computer related support. Thanks to Sarah Barbet and Lydia Hartleben for their bureaucracy assistance throughout this thesis.

My grateful thanks are also extended to my collaborators. Thanks to Elliot Nelson and Dr. Frank von Delft for giving me the opportunity to perform an XChem fragment screen at Diamond. I would also like to express my sincere gratitude to the HTX lab at EMBL Grenoble, especially Dr. Irina Cornaciu, Dr. Damien Clavel and Dr. Jose Márquez and for the iNEXT framework to enable an X-ray fragment screen at the EMBL/ESRF in Grenoble. Thanks for your patience, for not giving up on my target to perform a successful fragment screen and for offering me a visit at the HTX lab.

I also wish to thank the beamline staff at the Swiss Light Source, European Synchrotron Radiation Facility and the Diamond light source for their outstanding work and support.

Thanks to Prof. Xavier Barril, who provided me the opportunity to join his team for two months and who gave me an introduction in the work of his lab. Thanks also to his lab members Moira Rachman, Maciej Mjewski, Serena Piticchio for their help during my visit. I also wish to acknowledge the collaboration with Moira Rachman on the fragment evolution. Thanks also to Francesca Magari and Giulia Opassi on the collaboration in this project.

I would like to thank my funding agency for the Marie Curie H2020 fellowship and Dr. Iwan de Esch and Dr. Jacqueline Muijlwijk-Koezen for coordinating the FragNET ITN. Thanks for all their work to make this program so valuable, fun and memorable for us, students. A special thanks also to all my ESR colleagues for the great network and good times we had together. I hope we all stay in touch. Special thanks to Angelo for being an amazing PhD buddy and for supporting each other throughout this journey.

I also wish to thank Matthias, Lukasz and Ole for sharing their experiences, for their words of encouragement and for all the little chats that made this journey much more enjoyable. Furthermore, I would like to thank my colleagues Anna, Katharina, Claudia, Philipp, and Sara for a fun time together and great conversations during their time here at Novartis and beyond.

Special thanks to Joy for the great team work, for the unique friendship and for all the memories through the years. Thanks for supporting each other on this journey. It would not have been the same incredible experience without you.

Thanks to my friends here in Basel but also all over the world for providing happy distraction and for unforgettable memories.

Thanks to my boyfriend Nicolas for his constant support, and especially I cannot thank you enough for encouraging me throughout this experience. I am thankful for your love and patience. And last but not least, I am extremely thankful to my family, my parents and my brother, for their unlimited support, and their belief in me. Without them I would not be here today. Thank you!

Curriculum Vitae

Aus Gründen des Persönlichkeitsschutzes wird von der elektronischen Veröffentlichung des Lebenslaufes abgesehen.