



Kent Academic Repository

Huleani, Sergiu (2018) *The identification of novel fragment leads by Biophysical methods for the inhibition of TbrPDEB1 as an anti-parasitic approach to Human African Trypanosomiasis.* Master of Science by Research (MScRes) thesis, University of Kent,.

Downloaded from

<https://kar.kent.ac.uk/71666/> The University of Kent's Academic Repository KAR

The version of record is available from

This document version

Publisher pdf

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).



**The University of Kent
School of Biosciences**

**The identification of novel fragment leads by Biophysical
methods for the inhibition of TbrPDEB1 as an anti-
parasitic approach to Human African Trypanosomiasis**

By:

Sergiu Huleani

MSc Degree in: Drug Design

Year of Submission: 2018

Total word count: 23,419

Signed:

A handwritten signature in black ink, appearing to be the name "Sergiu Huleani".

Declaration

No part of this thesis has been submitted in support of an application for any degree or other qualification of the University of Kent, or any other University or Institution of learning.

Acknowledgements

I would first like to thank my Master's by research supervisor Professor David Brown of the University of Kent School of Biosciences. His door was always open when I needed help, and his immense knowledge, understanding and friendly character always kept me motivated and hungry to learn more. His teaching and encouragement allowed my work to blossom into something I am proud to call my own. Alongside Professor Brown, I would also like to thank Dr. Paul England, Dr. David Bailey and Dr. Fredrik Svenson from IOTA Pharmaceuticals for giving me a chance to participate in this fantastic project and always being there to give an expert's advice on whatever it is I had trouble with.

I would also like to thank the experts and fellow colleagues that I am proud to call my friends: Dr. Abhimanyu Singh, Dr. Susanne Schroeder, Maria Dodina and Lorena Zara for their continued support, patience and company throughout my research. A special thank you to Dr. Gary Thompson from the Wellcome Trust Biomolecular NMR Facility at the University of Kent for his continued support and endless assistance.

Finally, I must express my profound gratitude to my parents, my partner Samantha and my close friends for providing me with endless love, support and encouragement throughout my studies. This accomplishment would not have been possible without them, and I hope to make them proud.

Table of Contents

1	Introduction.....	16
1.1	Neglected Parasitic Diseases	16
1.2	African Trypanosomiasis	17
1.3	Parasitic Phosphodiesterases (PDEs)	21
1.4	Fragment-based Screening (FBS)	27
1.5	Aims	29
2	Experimental.....	31
2.1	Chemicals and Materials	31
2.2	TbrPDEB1 Catalytic Domain Construct	33
2.3	Preparation and Transformation of Competent cells.....	34
2.4	Cell Growth and Expression of Target Protein	36
2.5	Purification of Target Protein.....	38
2.5.1	Immobilized Metal Affinity Chromatography (IMAC).....	38
2.5.2	Secondary IMAC	41
2.5.3	Ion Exchange Chromatography (IEX)	41
2.6	Biochemical Assays of Target Protein	44
2.6.1	Verifying TbrPDEB1 activity and determining working concentrations	45
2.7	X-Ray Crystallography.....	46
2.7.1	Crystallisation by XCHEM.....	47
2.7.2	Manual Crystallisation	48
2.8	Nuclear Magnetic Resonance (NMR)	49
2.8.1	Compound Quality Control (QC).....	49
2.8.2	Protein NMR	49
2.8.3	Protein-Ligand Binding Affinity NMR.....	50
3	Results and Discussion.....	52
3.1	Mass Spectrometry	52

3.2	Purification Analysis	53
3.2.1	IMAC AKTA Graph	53
3.2.2	IMAC SDS-PAGE Gel	55
3.2.3	IEX AKTA Graph	56
3.2.4	IEX SDS-PAGE Gel	57
3.3	Quality Control (QC).....	59
3.3.1	Ligand 6 QC	59
3.3.2	Ligand 29 QC	61
3.3.3	Ligand 27 QC	62
3.4	X-Ray Crystallography.....	66
3.4.1	Target Protein Apo Structure	67
3.4.2	Inhibitor Confirmation by X-Ray Crystallography.....	73
3.4.3	XCHEM Data Analysis.....	85
3.5	Biochemical Assays	87
3.5.1	Determining a suitable TbrPDEB1 concentration.....	87
3.5.2	Single concentration assessment of identified X-Ray Fragment Hits against TbrPDEB1	89
3.5.3	Determining a suitable hPDE4D concentration	92
3.5.4	Single concentration assessment of identified X-Ray Fragment Hits against hPDE4D.....	93
3.5.5	Km determination of TbrPDEB1	96
3.5.6	Dose response curve Biochemical assays	97
3.6	Ligand-Protein NMR.....	102
3.6.1	TbrPDEB1 1D Proton NMR Spectrum.....	103
3.6.2	Protein-Ligand Interaction NMR	105
3.6.3	Ligand 27 competitive STD NMR.....	114
3.7	Substructure Similarity Search	117
4	Conclusion.....	119

List of Figures and Tables

Table 1 – List of known PDEs

Table 2 – Brief overview of biophysical screening techniques

Table 3 – Transformed competent cell growth

Table 4 – Ligand 6 Proton NMR annotated table

Table 5 – Ligand 29 Proton NMR annotated table

Table 6 – Ligand 27 Proton NMR annotated table

Table 7 – Ligand hits by X-Ray Crystallography overview

Table 8 – X-Ray data collection and refinement statistics for TbrPDEB1-Lig29 complex

Table 9 – X-Ray data collection and refinement statistics for TbrPDEB1-Lig27 complex

Table 10 – X-Ray data collection and refinement statistics for TbrPDEB1-Lig6 complex

Table 11 – Biochemical Assay determined inhibition of X-Ray hits and controls to TbrPDEB1

Table 12 – Biochemical Assay determined inhibition of X-Ray hits and controls to hPDE4D

Table 13 – ChemBL Reported IC₅₀ and K_i values of fragments against hPDE4D (μM)

Table 14 – NPD-008 IC₅₀ table

Table 15 – Overview of X-Ray Crystallography hits screened by various protein-ligand NMR techniques

Table 16 – Brief overview comparison of screening techniques used

Figure 1 – Mapped distribution of *T brucei* subspecies in 2010

Figure 2 – Diagram showing the accumulation of Trypomastigotes in red blood cells

Figure 3 – Charted number of reported cases of sleeping sickness (grey columns) and population screened for human African trypanosomiasis (black circles) between 1939-2004

Figure 4 – Reaction scheme showing PDE catalysis of cAMP substrate to AMP, drawn on Chems sketch

Figure 5 – Reaction scheme showing AC catalysis of ATP substrate to cAMP, drawn on Chems sketch

Figure 6 – Trypanosoma brucei parasitic PDEs

Figure 7 – Effect of NPD-008 inhibitor on intracellular cAMP levels

Figure 8 – Fluorescence microscopy of *T. brucei* incubated with 10 μ M NPD-008 inhibitor for: 0hr (a), 6hr (b), 12hr (c) and 24 hr (d)

Figure 9 – Brief overview of drug discovery process

Figure 10 – pET28a-TbrPDEB1 plasmid map, generated by SnapGene

Figure 11 – Graph depicting BL21(DE3*) Codon-Plus competent cell growth

Figure 12 – TbrPDEB1 catalytic domain used for transformation

Figure 13 – Lonza PDELIGHT assay kit reaction scheme

Figure 14 – XCHEM platform workflow A.K. Singh 2017

Figure 15 – TbrPDEB1 deconvoluted mass spectrum

Figure 16 – TbrPDEB1 IMAC sample Load and Wash graph

Figure 17 – TbrPDEB1 IMAC Elution and Fractionation graph

Figure 18 – TbrPDEB1 IMAC SDS-PAGE Gel

Figure 19 – Protein ladder

Figure 20 – TbrPDEB1 IEX Elution and Fractionation graph

Figure 21 – TbrPDEB1 IEX SDS-PAGE Gel

Figure 22 – Ligand 6 Proton NMR QC Spectrum

Figure 23 – Ligand 6 assigned peaks

Figure 24 – Ligand 29 Proton NMR QC Spectrum

Figure 25 – Ligand 29 assigned peaks

Figure 26 – Ligand 27 Proton NMR QC Spectrum

Figure 27 – Ligand 27 Proton NMR QC spectrum provided by Enamine

Figure 28 – Ligand 27 assigned peaks

Figure 29 – DMSO₆ Solvent Proton NMR Spectrum

Figure 30 – Crystallised TbrPDEB1 protein – 20µm in length

Figure 31 – TbrPDEB1 PDB Crystal structure overview – 1.65Å – Made with PyMOL

Figure 32 – 1: TbrPDEB1 PDB Crystal structure – Divalent metal ions and water network
– Made with PyMOL

Figure 33 – 2: TbrPDEB1 PDB Crystal structure – Labelled M-loop – Made with PyMOL

Figure 34 – 3 & 4: TbrPDEB1 Crystal Structure – Substrate binding pocket, Hydrophobic clamp, Glutamine + water network and Parasite specific ‘P-pocket’ – Made with PyMOL

Figure 35 – P-pocket region differences in TbrPDEB1 (A), LmjPDEB1 (B), TcrPDEC (C) and hPDE4B (D)

Figure 36 – NPD-008 bound to TbrPDEB1 protein structure, made with PyMOL using the 5G2B PDB Structure code

Figure 37 – TbrPDEB1 crystal soaked with Ligand 29 snapshot

Figure 38 – TbrPDEB1 crystal soaked with Ligand 29 1.52Å diffraction

Figure 39 – TbrPDEB1-Lig29 successful ligand binding, indicated by green electron density

Figure 40 – TbrPDEB1-Lig29 3D structure and binding from 3 different angles, made with PyMOL

Figure 41 – TbrPDEB1-Lig27 3D structure and binding made with PyMOL

Figure 42 – TbrPDEB1-Lig6 3D structure and different binding modes (1 and 2) and binding mode overlap (3), made with PyMOL

Figure 43 – Molecular structure of IBMX – a known PDE inhibitor, drawn on Chems sketch

Figure 44 – Molecular structures of cAMP (left) and cGMP (right), drawn on Chems sketch

Figure 45 – Determination of a suitable working TbrPDEB1 concentration by use of a Biochemical Assay

Figure 46 – Assessment screen of X-Ray Fragment Hits against determined optimal TbrPDEB1 concentration by use of a Biochemical Assay

Figure 47 – Remaining TbrPDEB1 activity at 200 μ M X-Ray Fragment Hit Concentration in graph form

Figure 48 – Determination of a suitable working hPDE4D concentration by use of a Biochemical Assay

Figure 49 – Assessment screen of X-Ray fragment hits against determined optimal hPDE4D concentration by use of a Biochemical Assay

Figure 50 – Remaining hPDE4D activity at 200 μ M X-Ray Fragment Hit Concentration in graph form

Figure 51 – TbrPDEB1 enzyme K_m determination Biochemical assay

Figure 52 – Ligand 29 dose response curve against TbrPDEB1

Figure 53 – Ligand 27 dose response curve against TbrPDEB1

Figure 54 – NPD-008 dose response curve against TbrPDEB1

Figure 55 – TbrPDEB1 protein Proton 1D NMR Spectrum

Figure 56 – TbrPDEB1 Proton 1D NMR spectrum impurity (left) and glycerol Proton 1D NMR spectrum (right)

Figure 57 – Ligand 29-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 29 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)

Figure 58 - Ligand 29-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 29 on-resonance control spectrum (Middle), difference spectrum (Bottom)

Figure 59 - Ligand 29 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 29 on-resonance control spectrum (Middle), difference spectrum (Bottom)

Figure 60 - Ligand 29-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 29 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)

Figure 61 - Ligand 27-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 27 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)

Figure 62 - Ligand 27-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 27 on-resonance control spectrum (Middle), difference spectrum (Bottom)

Figure 63 - Ligand 27-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 27 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)

Figure 64 - Ligand 1-TbrPDEB1 STD Competitive NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 1 on-resonance control spectrum (Middle), difference spectrum (Bottom)

Figure 65 - Ligand 1+27-TbrPDEB1 Competitive STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 1+27 on-resonance control spectrum (Middle), difference spectrum (Bottom)

Figure 66 - Ligand 27 in DMSO Proton 1D

Figure 67 – Ligand 27 in phosphate buffer Proton 1D

Figure 68 - Ligand 29 (1) ROCS similarity search, VM_STL367219 (2), CB-5346663 (3), EN_Z2157984556 (4)

Figure 69 – Project overview

Abbreviations

PDE = Phosphodiesterase

Tbr = *Trypanosoma brucei*

Tcr = *Trypanosoma cruzi*

HAT = Human African Trypanosomiasis

PCR = Polymerase Chain Reaction

IPTG = isopropyl β -D-1-thiogalactopyranoside

NPD = Neglected Parasitic Diseases

cAMP = cyclic adenosine monophosphate

cGMP = cyclic guanosine monophosphate

AMP = adenosine monophosphate

GMP = guanosine monophosphate

AC = adenylyl cyclase

Tris = tris(hydroxymethyl)aminomethane

BME = 2-Mercaptoethanol

MES = 4-Morpholineethanesulfonic acid

PEG 3350 = Polyethylene glycol 3350

RPM = Rotations per minute

FPLC = Fast Performance Liquid Chromatography

IMAC = Immobilized Metal Affinity Chromatography

KPSI = Kilo pounds per square inch

C.V. = Column Volume

SDS-PAGE = Sodium dodecyl sulfate-polyacrylamide gel electrophoresis

DTT = Dithiothreitol

kDa = Kilodalton

IEX = Ion Exchange Chromatography

pI = Isoelectric point

ATP = Adenosine triphosphate

HEPES = 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

DMSO = Dimethyl sulfoxide

FBS = Fragment based screening

NMR = Nuclear Magnetic Resonance

EDTA = Ethylenediaminetetraacetic acid

E. coli = Escherichia coli

QC = Quality Control

TMS = Tetramethyl silane

DDT = Dichlorodiphenyltrichloroethane

PDB = Protein Data Bank

hPDE = Human phosphodiesterase

Lmj = *Leishmania major*

MRC-5 = Medical Research Council cell strain 5

IBMX = (3-isobutyl-1-methylxanthine)

RLU = Relative Light Unit

STD NMR = Saturation-Transfer Difference NMR

CPMG NMR = Carr-Purcell-Meiboom-Gill NMR

WaterLOGSY NMR = Water-Ligand Observed via Gradient Spectroscopy NMR

NOE = Nuclear Overhauser Effect

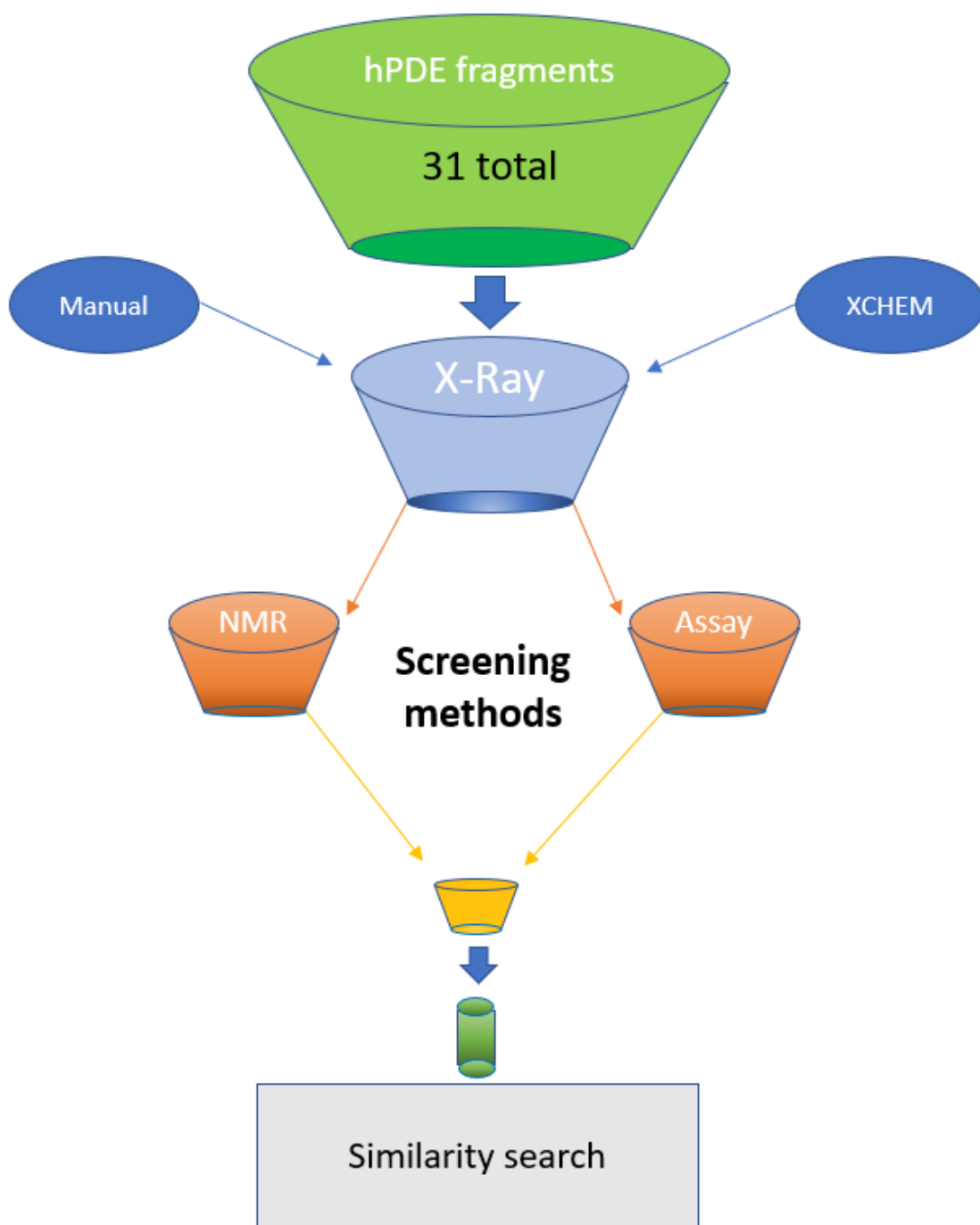
IC₅₀ = Half Maximal Inhibitory Concentration

SAR = Structure-activity relationship

Abstract

This drug design research project is based on identifying novel compound ‘hits’ against an established drug target by various screening methods such as NMR, X-Ray Crystallography and Biochemical Assays. *Trypanosoma brucei* cyclic nucleotide phosphodiesterase B1 (TbrPDEB1) is a validated drug target reported to be essential in the life cycle of the *Trypanosoma brucei* parasite – the causative agent of Human African Trypanosomiasis, one of several Neglected Parasitic Diseases which affects underdeveloped countries, and in some cases even First World countries. Otherwise known as Sleeping Sickness, the insect-borne parasitic disease occurs in 36 sub-Saharan African countries. The estimated mortality rate is up to 500,000 deaths per year and up to 80 million people are at risk of infection. With current methods of treatment being limited, ineffective and in some cases unsafe, there is an urgent need for new, more effective and safer medication.

In this research project, the TbrPDEB1 catalytic domain construct plasmid (provided by the host laboratory at the University of Kent) was used to transform an established *E. coli* bacteria strain for the purpose of protein expression. A small fragment library put together by IOTA Pharmaceuticals consisting of 31 active human PDE inhibitors was initially validated by Proton NMR and used to soak TbrPDEB1 protein crystals. Ligand uptake was determined by X-Ray diffraction at the Diamond Light Source. Through the use of the three previously mentioned screening approaches, fragment binders to the TbrPDEB1 target enzyme were identified. Additionally, inhibition of the fragment binders to the target was also measured and assessed in comparison to NPD-008, a known TbrPDEB1 inhibitor of moderate potency. Out of the 31 fragments, one fragment was identified as an optimal binder and low potency inhibitor to TbrPDEB1, which underwent a similarity search follow-up thus to suggest similar chemotypes of interest for future work.



1 Introduction

Since the dawn of time, mankind has continuously shared a common instinct throughout the decades – survival. In modern times, the development and understanding of medicinal practice in combination with advanced technology has greatly supported our most basic instinct, with key scientific discoveries such as Penicillin, Germ Theory and DNA alongside major research projects such as the Human Genome Project gifting mankind an extended average lifespan of 30 years¹. Despite this, illness and disease remain a great issue in many regions of the world, particularly in underdeveloped countries lacking proper sanitation and healthcare services. Hence, new research and further ground-breaking discoveries in drug development continue to be pursued in endless topics of medicinal treatment, thus continuing to aid and support mankind’s survival. The topic of interest in the case of this research project is based on neglected parasitic diseases (NPDs).

1.1 Neglected Parasitic Diseases

A parasitic disease is tied with the term ‘neglected’ when contributing investments for controlling said disease are extremely low considering the health impact it has on human life. This impact causes most damage in underdeveloped countries lacking proper sanitation and access to public health care services; however, this is not to say that developed countries remain unaffected.

*Table 1:
Known NPDs²*

Known NPDs as reported by WHO			
Buruli Ulcer	Chagas disease	Dengue and Chikungunya	Dracunculiasis
Echinococcosis	Foodborne trematodiasis	Human African Trypanosomiasis	Leishmaniasis
Leprosy	Lymphatic filarasis	Mycetoma and chromoblastomycosis	Onchocerciasis
Rabies	Scabies and other ectoparasites	Schistosomiasis	Soil-transmitted helminthiasis
Snakebite envenoming	Taeniasis	Trachoma	Yaws

A table listing known NPDs can be seen in *Table 1* on the previous page, interpreted from the World Health Organisation (WHO)². NPDs of interest in this project are seen highlighted in red.

Chagas disease, otherwise known as American trypanosomiasis is a NPD caused by the parasitic kinetoplastid *Trypanosoma cruzi*, or Tcr. American trypanosomiasis affects 6 to 7 million people worldwide, most of which are located in Latin America - the location of origin of the Tcr parasite³. For a long time, the parasite was confined to the Latin America region carried in a vector being the Triatominae insect. The vector-borne parasite is transmitted to hosts following a blood meal. The lack of symptoms and non-specific manifestations in the acute phase of this parasitic disease allows it to remain unnoticed and spread in the body by infecting cells and muscle tissue, eventually leading to serious chronic health issues such as cardiomyopathy within 20 to 30 years following infection. A 2016 research paper estimates around 1 million people were infected by the parasite in 2006 in Mexico, with an additional risk of infection of up to 29.5 million people⁴. In recent years, Chagas disease has become a worldwide issue due to the urban migration and spread of the parasite from rural and underdeveloped areas⁵. Being a NPD, migration can occur in multiple ways, such as congenital transmission, organ transplantation, and especially the transfusion of contaminated blood which can quickly spread even in developed countries due to the inability to diagnose Chagas disease in patients as a result of clinical unawareness. Hence, the number of people infected by the NPD has risen greatly in the last decade, with an estimate of more than 300,000 people living in the United States being infected⁶. Current therapies for Chagas disease treatment are limited and no effective vaccines are available, emphasising the importance and urgent need for effective treatment. Chagas Disease is just one example of Neglected Parasitic Diseases – another being African Trypanosomiasis, the main focus in this research project.

1.2 African Trypanosomiasis

Like Chagas Disease, African trypanosomiasis is a NPD caused by the same parasitic kinetoplastid genus – *Trypanosoma*, a monophyletic group of unicellular parasitic flagellate protozoa⁷. The *Trypanosoma brucei* parasite species is the causative agent of African trypanosomiasis, otherwise known as ‘Sleeping sickness’ – one of the most important but equally most neglected parasitic diseases known to impact more than 70 million people¹³. Most commonly found in rural areas of sub-Saharan Africa¹⁰, Human African trypanosomiasis (HAT) exists in two different forms caused by morphologically

indistinguishable subspecies of *T. brucei*⁸. The parasite is transmitted to humans through the bite of numerous different species of the Tsetse fly vector: *T brucei gambiense* is found in West and Central Africa, whilst *T brucei rhodesiense* is more common in the Eastern and Southern regions of Africa^{8,9} – the former of these subspecies accounting for 97% of reported cases of HAT¹²; These are the main forms of sleeping sickness with varying clinical presentations that infect humans, as seen mapped out in *Figure 1* - interpreted from a 2010 research paper¹⁰.

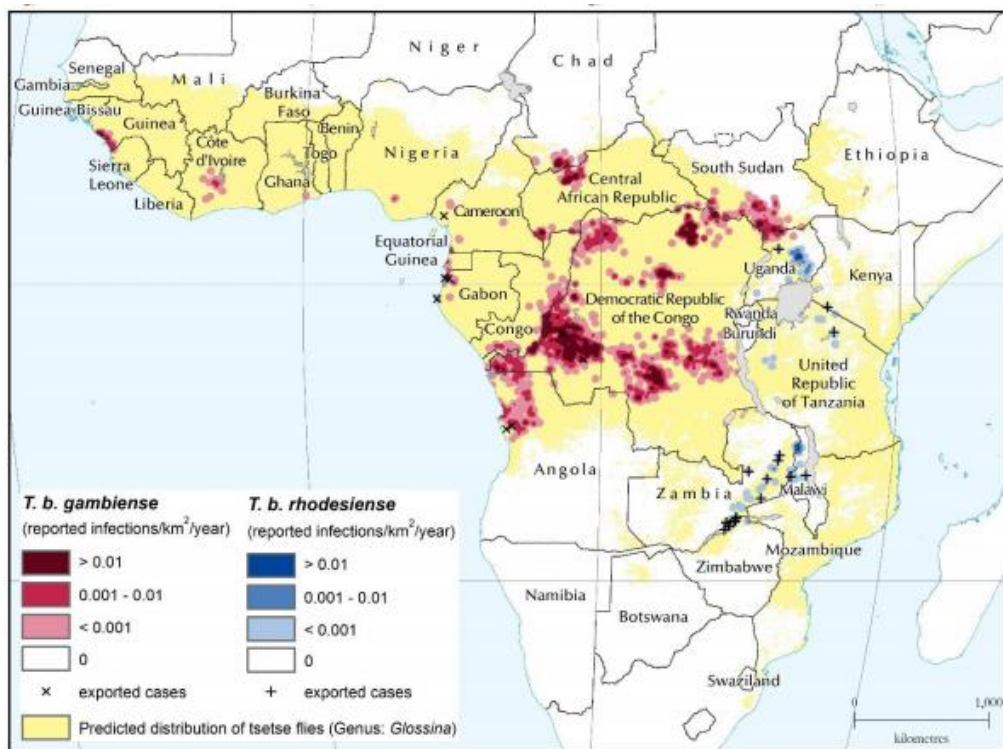


Figure 1:
Mapped distribution of T brucei subspecies in 2010¹⁰

Animals can also host these forms of parasite and can act as parasitic reservoirs for uninfected tsetse flies, allowing the disease to quickly manifest and spread in areas at risk. A third pathogenic subspecies of the *T brucei* parasite, *T brucei brucei* is known to cause Animal trypanosomiasis, otherwise known as nagana. Nagana affects wild and domestic animals alike and therefore causes further impact on economic and agricultural development in Africa¹¹ due to the loss of meat and milk production from infected and chronically ill cattle. HAT evolves through clinically distinct stages, with the initial symptoms of fever being non-specific to the disease and therefore making it difficult to diagnose – a common trait in NPDs also seen in Chagas disease. Trypanosomal chancre, a skin reaction caused by the immediate bite of an infected tsetse fly is seen in the faster progressing East African trypanosomiasis, making the disease somewhat more

distinguishable compared to the more inconspicuous, slower acting West African trypanosomiasis⁸ variation.

During the primary stages of infection, an increased severity of fever, headaches and joint pains is seen in patients as the trypanosomes enter the lymphatic system and mature into the Trypomastigotes stage. Lymphadenopathy, or ‘Winterbottom’s sign’ – an abnormal inflammation of the lymph nodes as a result of parasitic growth is a common characteristic seen in the early stages of HAT^{8,12}. Trypomastigotes accumulated in the lymph nodes multiply by binary fission and enter the bloodstream, soon passing the blood brain barrier¹⁴ and infecting the central nervous system; *Figure 2* depicts the accumulation of Trypomastigotes in red blood cells. This second, neurological encephalopathy stage is where symptoms of the disease become most apparent, and the stage at which death is imminent if no method of treatment is provided.

Changes of behaviour, lethargy and confusion are common in this stage of infection, as well as a disturbance of the sleep cycle – a feature giving the disease its notorious name.

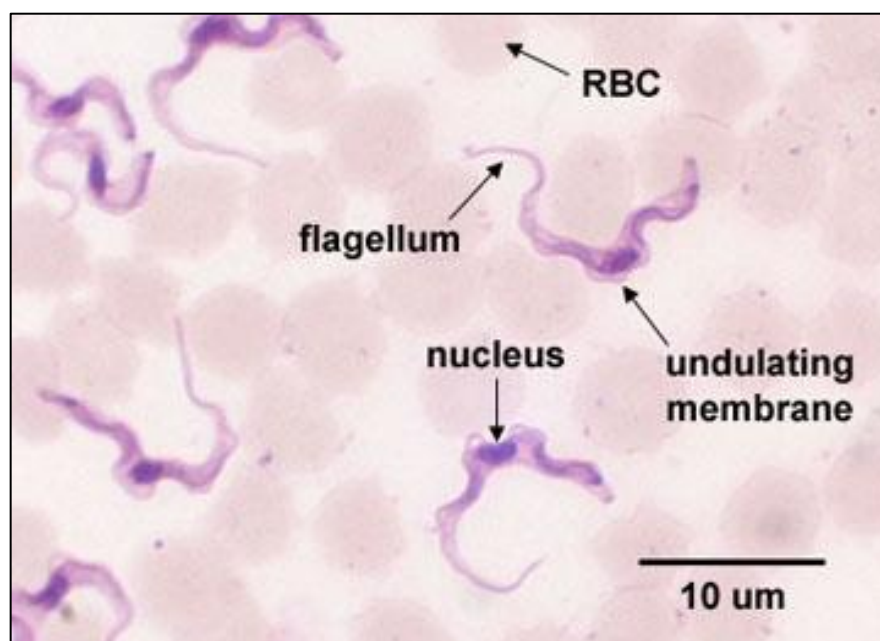


Figure 2:
Diagram showing the accumulation of Trypomastigotes in red blood cells^{8,52}

Leading up to modern times, three severe sleeping sickness epidemics occurred in Africa in the 20th century – with the first lasting 8 years until 1906 and resulting in the deaths of 800,000 people¹⁵ in Uganda and Kenya – with the colonisation of Africa creating optimal disease spreading conditions due to poor healthcare, lack of sanitation and famine. A second epidemic was seen in the 1920s, which eventually subsided and led to the decline

of sleeping sickness in Sub-Saharan Africa by 60%¹⁵ due to the development of early forms of treatment against sleeping sickness including Suramin and an organo-arsenic based trypanamide (later developed into Melarsoprol) which was used to combat the second-stage infection of both HAT¹⁵ variations. Pentamidine – an antimicrobial - was eventually developed in 1937 and was used to treat first stage west African trypanosomiasis⁸, with Suramin being used against the East African variant. Numerous methods of vector control such as bush clearing, destruction of parasitic reservoirs such as infected cattle and game animals, and the used of the DDT (dichlorodiphenyltrichloroethane) insecticide (several years after its discovery in 1939¹⁵) also contributed greatly to the drastic NPD decline, leading to an all-time low of under 10,000 reported cases in 1960¹⁵. This however was short lived, as the disease once again began to spread at an increasing rate by the mid-1970s, eventually leading to the third and most recent epidemic in the 1990s with 500,000 reported cases up until 2015^{15,17}. This can be seen in *Figure 3* below - a graph depicting the number of reported cases of sleeping sickness and population screened between 1939-2004, interpreted from a 2008 review article on the history of the relevant NPD by D. Steverding^{15,16}.

The rapid decolonisation of Africa following World War II led to political instability in many African countries, resulting in neglect of health services and in turn an ignorance to the sleeping sickness disease¹⁵. In combination with a lack of support from new governments, screening programs and parasitic control methods gradually declined as seen circled in red in *Figure 3*, leading to neglect of sleeping sickness endemic countries and a gradual increase of infection.

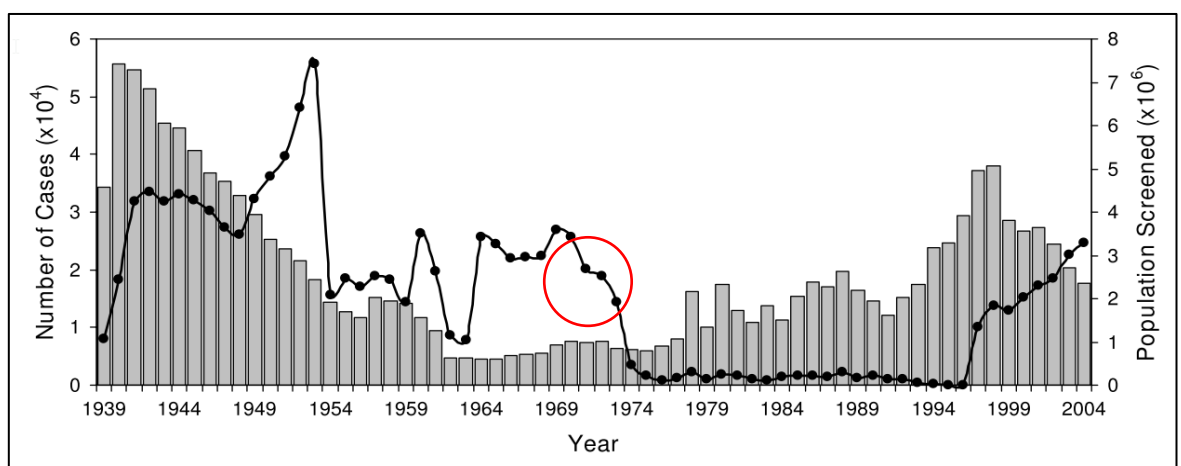


Figure 3:
Number of reported cases of sleeping sickness (grey columns) and population screened for the disease (black circles) between 1939-2004^{15,16}

In recent years, the increased awareness of the NPD as a result of a WHO campaign in 2015 which donated necessary drugs to be distributed in affected countries¹⁷ increased patient screening and disease diagnosis, and continued methods of vector control have helped increase control over the disease and reduce the overall spread and reported cases. Despite an estimate of 20,000 reported cases, the number of people affected are still high, with an even greater number of 65 million people still at risk of infection. Current forms of treatment still include those previously mentioned; Pentamidine and Suramin are relatively safe drugs with little side effects and continue to be used to treat first-stage sleeping sickness, however some cases of pathogenic resistance of the latter drug have been reported^{18,19}. Melarsoprol is still one of the only drugs used for second-stage sleeping sickness treatment despite causing many side effects and sometimes cases of reported toxicity to humans, even proving to be fatal in up to 10% of cases¹². Eflornithine²⁰ is another form of treatment less toxic than Melarsoprol, however it is much more difficult to administer to a patient, making it an unreliable treatment choice. Repeated treatment over a period of 8-10 days is also required, with second-stage cases including hospital admission and intensive patient monitoring, therefore also demanding a high degree of expertise and required training for the treatment of patients⁸.

Hence, current methods of treatment have proven to be undesirable, difficult to use and in some cases unsafe. There is an urgent need for novel, safer and more effective medication for the treatment of HAT. Recent studies established the role of cyclic nucleotide phosphodiesterase (PDE) enzymes in the life cycle of the *Trypanosoma brucei* parasite. The importance of this enzyme in parasitic functionality makes it a promising research drug target for the treatment of HAT.

1.3 Parasitic Phosphodiesterases (PDEs)

Phosphodiesterases, or PDEs comprise a group of enzymes capable of catalysing the hydrolysis of secondary messenger molecules cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) to adenosine monophosphate (AMP) and guanosine monophosphate (GMP) respectively, following the degradation of a phosphodiester bond²¹. The reaction scheme (drawn using chemsketch⁶⁸) can be seen in *Figure 4* on the following page, adapted from a study by F. Svennson, A. Bender and D. Bailey²²:

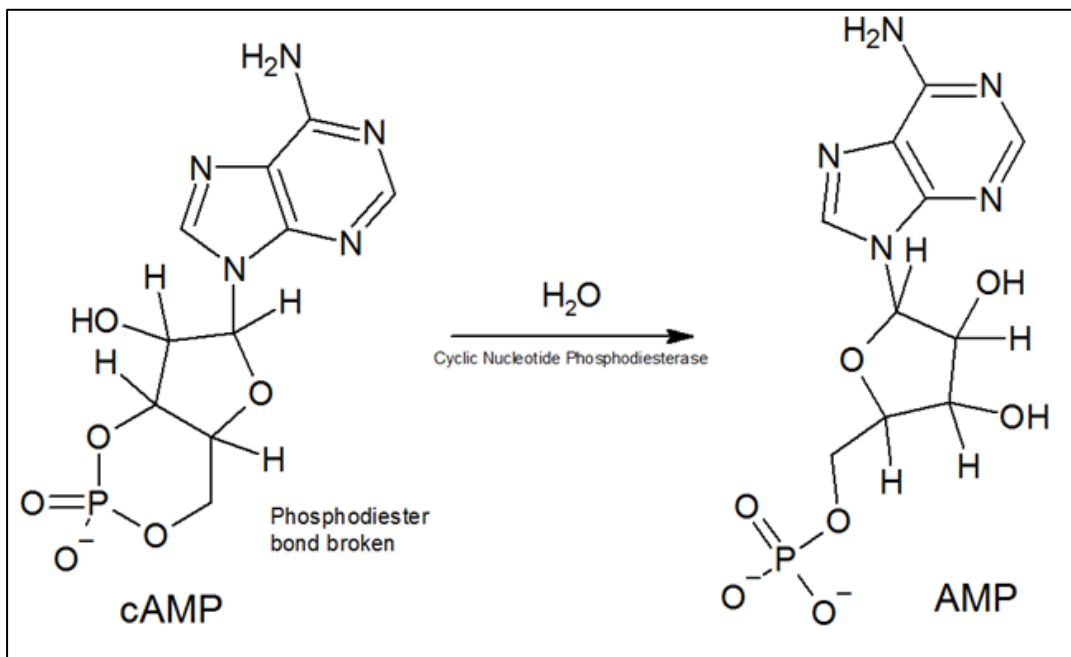


Figure 4:
Reaction scheme showing PDE catalysis of cAMP substrate to AMP²²

The secondary cyclic nucleotide messenger molecules cAMP and cGMP are important in mediation of biological responses, and are essential in hormone regulation, neurotransmitters, protein signalling and cell proliferation and differentiation²¹. At high cyclic nucleotide concentrations, protein kinases A and G (dependent on cAMP and cGMP, respectively) are activated in eukaryotes, resulting in the phosphorylation of numerous substrate protein. The phosphorylation of these substrates is vital in a number of physiological processes such as immune responses, ion channel conductance and cell apoptosis²¹.

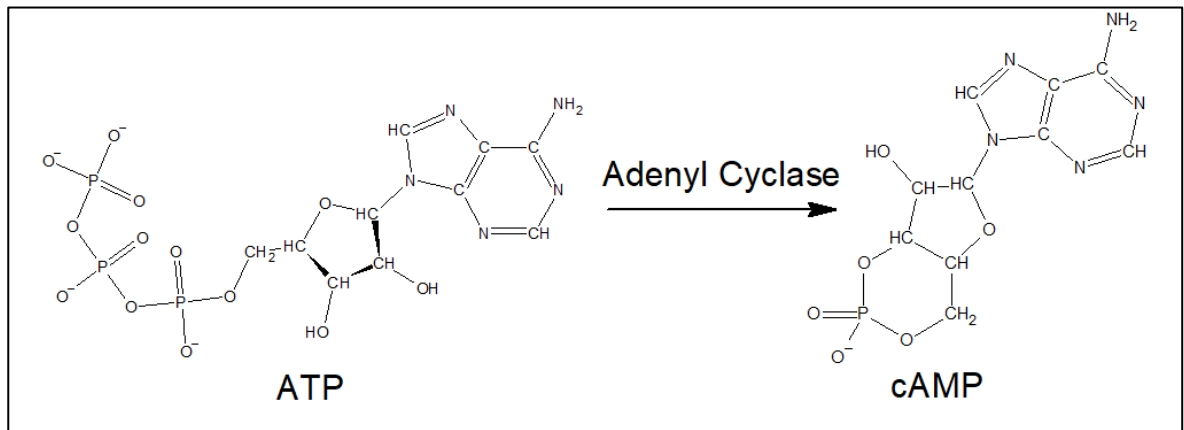


Figure 5:
Reaction scheme showing AC catalysis of ATP substrate to cAMP²³

Similarly, a second reaction scheme shown in *Figure 5*²³ shows the synthesis of cAMP from ATP, catalysed by an adenyl cyclase (AC) enzymes. Both reaction scheme described in *Figure 4* and *Figure 5* aid in regulating the level of secondary messengers and are therefore essential for most cellular processes and intracellular signalling. PDE enzyme inhibitors would block this reaction pathway, resulting in high levels of cyclic nucleotide build up due to the inability of catalytic breakdown via hydrolysis and hence a disturbance in cell functionality.

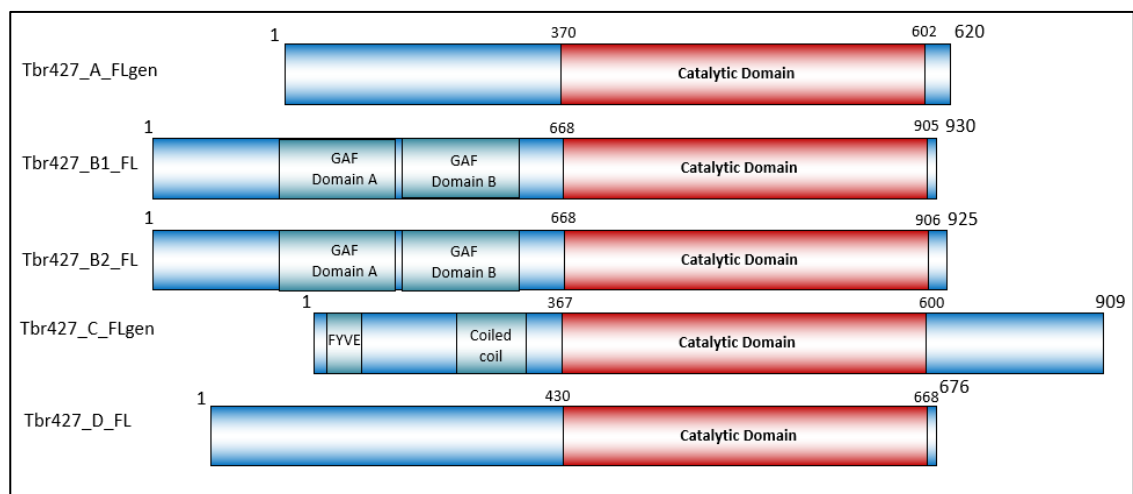


Figure 6:
Trypanosoma brucei parasitic PDEs

Parasitic PDE phenotypes can be classed into differing families of genes present in all kinetoplastid genomes; Four genetically distinct PDE families exist known as PDE A-D²⁴. These can be seen in *Figure 6* in the case of the *T. brucei* parasite project focus.

Figure 6 gives a brief overview of the sequence differentiation in the 4 presented Tbr PDE forms. As can be seen, each family varies in terms of amino acid residues in the case of the full-length sequence, and also in types of protein domains; TbrB is seen to contain two GAF (cGMP-specific phosphodiesterase, adenylyl cyclase, FhlA transcriptional activator) domains²⁵ in the N terminal region, whilst TbrC a FYVE²⁶ domain – not seen present in the other forms. The catalytic domain region, however, appears to be consistently located in the C-terminal of the polypeptide chain across all 4 PDE families. Interestingly, TbrB exists in two distinct, non-identical isoforms; TbrPDEB1 and TbrPDEB2, differing only in 5 amino acid residues. Theoretically, both cAMP-specific enzyme isoforms would react in a very similar, almost identical way – due to the variability of only one amino acid in the catalytic domain. According to a 2005 study by Laxman *et al*²⁷, it is also suggested that the two GAF A domain present in both TbrB isoforms have a high affinity for both cAMP and cGMP.

Understanding the existence of varying parasitic PDE types, their role as catalysts in cyclic nucleotide reaction pathways and in turn why that would make them drug targets of interest for the treatment of HAT is essential. However, understanding the biological function of the presented enzymes in the life cycle of the *Trypanosoma brucei* parasite is also vital for drug target validation. The chapter ‘*Therapeutic Potential of Phosphodiesterase Inhibitors in Parasitic Diseases*’ by Y. Shakur *et al.* in *Phosphodiesterases as Drug Targets* – a 2011 book – discusses the biological function of the varying parasitic PDEs²⁸:

It is reported that the characterisation and homozygous deletion of the cAMP-specific TbrPDEA enzyme shows very minor effects on cell growth due to a slight increase in cAMP concentration²⁸. A 2001 study by Gong *et al*²⁹ also states that despite being active, the deletion of TbrPDEA has no effect on cell proliferation, motility or morphology of bloodstream Trypomastigotes, thus suggesting the gene to be ‘non-essential’ in the cAMP pathway previously shown in *Figure 4*.

On the contrary, inhibition of the TbrPDEB isoforms appears to have a promising effect: A 2002 article by R. Zoraghi and T. Seebeck³⁰ reports that inhibition of the cAMP-specific TbrPDEB1 parasitic enzyme blocks the proliferation of bloodstream trypanosomes *in vitro*³⁰. RNA interference experiments demonstrated that TbrB knock-out – particularly the B1 isoform – causes a drastic increase in intracellular cAMP concentration, reported to be

toxic for bloodstream trypanosome parasites. An article by H. de Koning *et al*³¹ further supports this claim; This article discusses the effects of known human PDE inhibitors on TbrPDEB1. Human and parasitic PDE enzymes are highly conserved^{31,32} due to having structurally very similar catalytic domains, hence in this study, a library of known and already explored human PDE inhibitors were screened against TbrPDEB1. A potent inhibitor, ‘Cpd A’, resulted in successful inhibition of TbrPDEB1, causing a rapid cAMP concentration increase and eventual parasite death via cell lysis after 42 hours³¹ *in vitro* and also *in vivo* in mice models. These findings validate the pharmacological potential of TbrPDEB – more particularly TbrPDEB1 – as novel drug targets³¹. Furthermore, *Figure 7* and *Figure 8* – interpreted from a recent 2018 research paper by A. Blaazer *et al*³³ – demonstrate this phenomenon. In this study a synthesised ligand, ‘NPD-008’, was used to inhibit TbrPDEB1 in bloodstream form trypanosomes:

As can be seen in *Figure 7*, there is a direct correlation between increasing NPD-008 inhibitor concentration and intracellular cAMP levels³³.

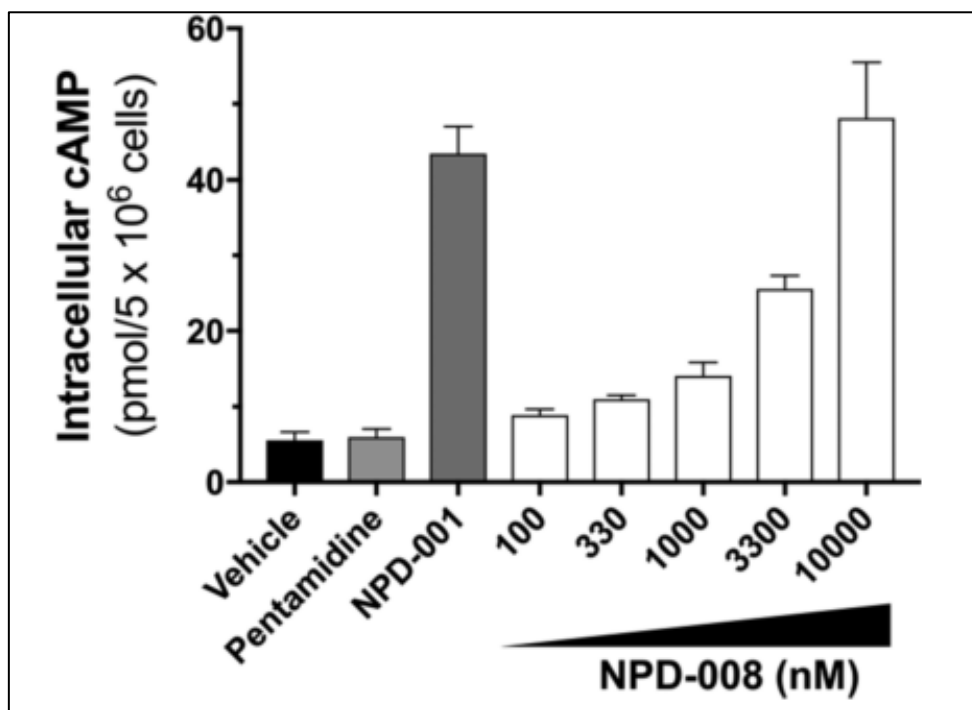


Figure 7:
*Effect of NPD-008 inhibitor on intracellular cAMP levels*³³

Hence, the effect of high intracellular cAMP is depicted in *Figure 8*³³ on the following page: Healthy Tbr trypanosomes were incubated with NPD-008 at 10 μ M concentration in a cytokinesis assay. The parasites were examined after 0, 6, 12, and 24 hours of incubation (a-d)³³.

As can be seen, the parasite begins to degrade after 6 hours *Figure 8(b)*. Large intracellular cAMP levels produced as a result of successful parasitic TbrPDEB1 inhibition in the case of *T. brucei* causes disrupted cell functionality as previously mentioned. The inability of successful cell division and cytokinetic defects³³ prevents reproduction of parasitic cells, resulting in eventual parasite death, as seen in *Figure 8(d)*. This equally confirms and proves the importance of the TbrPDEB1 enzyme in the life cycle of *T. brucei*^{34,37} parasite.

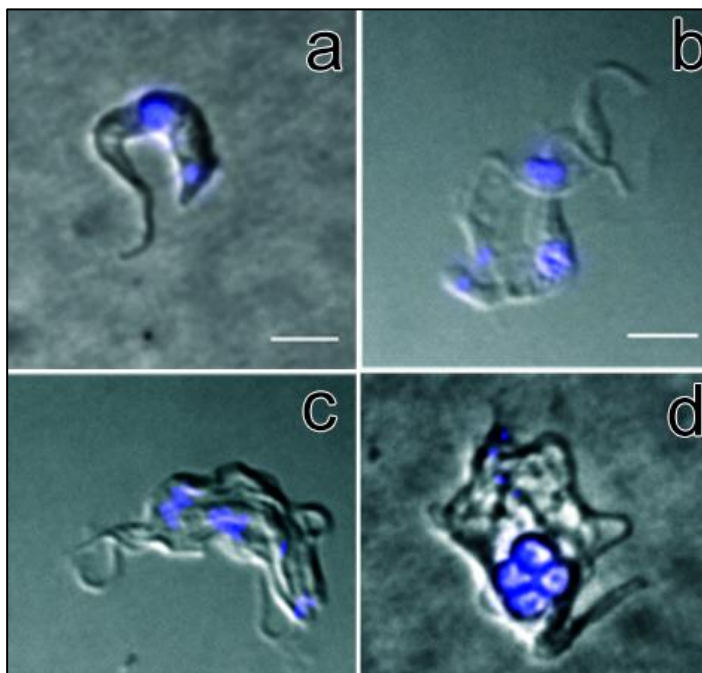


Figure 8:
*Fluorescence microscopy of *T. brucei* incubated with 10μM NPD-008 inhibitor for: 0hr(a), 6hr(b), 12hr(c) and 24hr(d)*

As previously mentioned, it is predicted that both TbrPDEB1 and TbrPDEB2 enzymes would behave in a similar way due to their almost identical catalytic domain. Due to this structural similarity, it is also suggested that both isoforms can functionally complement each other^{28,34,35}. Therefore, it is hypothesised that strong dual-inhibition of both enzymatic isoforms would likely result in drastic cAMP increase and eventual parasitic death³⁰. Note that it is apparent that enzyme inhibition does not alter DNA synthesis or nuclei division^{28,37}.

A 2012 study found that PDEC gene is inactive in some kinetoplastid species³⁶. Despite successful biological and structural characterisation of PDEC in *T. cruzi* by H. Wang *et al*³⁶, the study determines that this particular parasitic gene is not active, and hence not an

essential enzyme in the parasite life cycle. Similarly, little is known about PDED other than the DNA sequence³². However, the validation of the TbrPDEB isoforms, more particularly TbrPDEB1 makes it a clear choice as a promising drug target for the treatment of HAT. The vast crystallographic knowledge existent for TbrPDEB1 over its B2 isoform counterpart make it an increasingly more appealing choice for research; Hence, the TbrPDEB1 enzyme is the key focus of this research project. The enzyme will be isolated and characterised appropriately, before undergoing fragment screening as a method of drug discovery.

1.4 Fragment-based Screening (FBS)

Fragment-based Screening (FBS) is a rapidly developing and widely used method as a modernized drug discovery approach. The principle of this technique relies on the screening of fragments – small, low molecular weight compounds – against a particular validated drug target³⁸. Like High Throughput Screening (HTS)³⁹, FBS aims to identify active fragments, also known as ‘hits’ – compounds which have a weak binding affinity to the chosen biological target (or the desired activity in a compound screen, as a broader term¹¹). The identification of fragment hits to a drug target is a key starting point and the first step of novel drug design, as the understanding of fragment-target interactions and fragment hit optimization greatly aid in identifying promising lead compounds, as described by the ‘hit-to-lead’ principle⁴⁰. A 2011 article by JP Hughes *et al.* describes the basic principles of early drug discovery⁴¹:

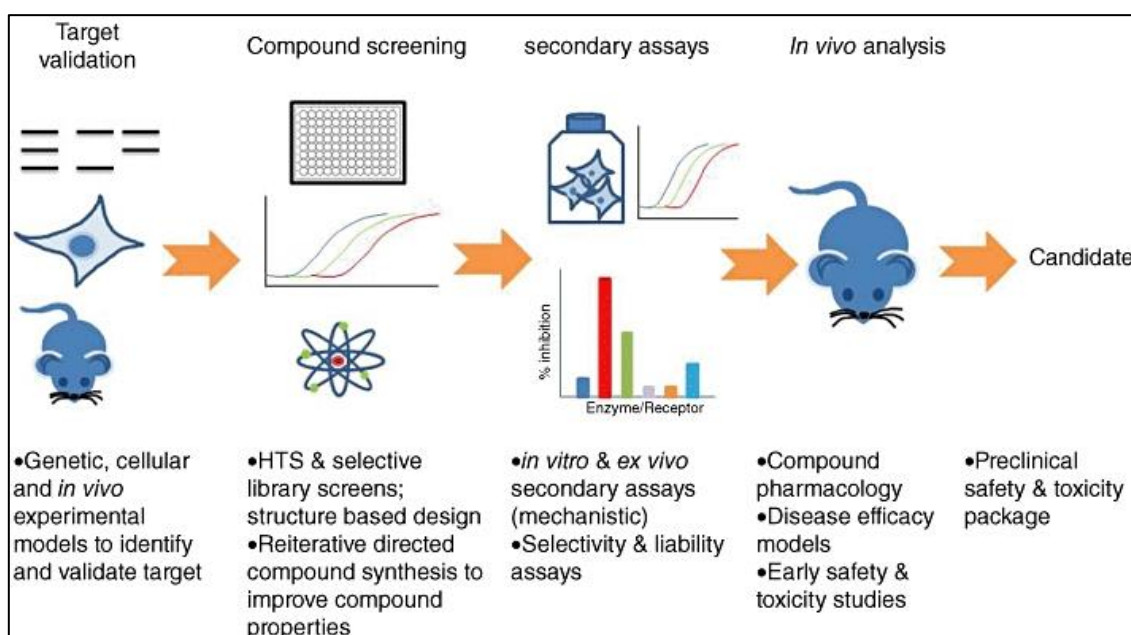


Figure 9:
Brief overview of a drug discovery process⁴¹

The article discusses the purpose of drug discovery programmes – a disease or medical condition without suitable medical products available⁴¹ initiates a drug research programme and in turn target identification and validation. The drug discovery process may be summed up in a brief flow chart interpreted from the relevant 2011 study⁴¹, as seen in *Figure 9* on the previous page.

Following target validation, the importance and efficiency of hit discovery processes such as FBS comes to light. The ability to screen large libraries of small fragments directly against the drug target quickly gives rise to lead compounds which may be further optimized and expanded to suitable drug candidates. A brief summary of biophysical screening methods can be seen described in *Table 2*, also adapted from the previously mentioned study⁴¹.

Table 2:
Brief overview of biophysical screening techniques⁴¹

Screen	Description
High throughput	Large number of compounds
Focused	Previously identified hits of similar targets
Fragment	Soaking of compounds into target crystals to determine interaction by X-Ray Crystallography
Structural-aided drug design	Use of crystal structures for molecule design
Virtual	Docking models through computational programs, hypothetical compound-target activity determination via a known crystal structure
Physiological	Tissue-based approach, drug affects at tissue level
Nuclear Magnetic Resonance (NMR)	Soaking of compounds into protein targets of known NMR structures to determine interaction

Aside from aiding in hit generation via screening techniques, NMR and X-ray crystallography also aid in validating hits by confirming binding interactions to the target⁴². These techniques are used alongside *in vitro* Biochemical assays⁴³ in order to establish and measure ligand efficiency and the activity of each compound to the target

through analytical techniques including dose-response curves⁴¹. As part of the optimization phase, modifications of lead compounds and active scaffolds by synthesizing similar chemotypical analogues aids in optimizing ligand efficiency by changing certain side chains or functional groups of said compounds and observing how the potency is affected. Docking models may also serve this purpose in the case of virtual screening. Screening and optimization techniques layered in different levels aids in bringing compound leads ‘cherry-picked’ from a large fragment library to compounds that are as ‘drug-like’ as possible; In other words, these compounds would possess^{44,45}:

- Potency towards the drug target (measured by Biochemical assays)
- Desired physiological affect whilst interacting with the target
- Solubility in water
- Stable
- Non-cytotoxicity
- Selectivity versus other similar targets^{44,45}

Compounds possessing these properties would be optimal candidates for continued development by *in vivo* studies and eventual clinical trials.

1.5 Aims

Whilst TbrPDEB1 has been proven to be a valid drug target, the fact that the catalytic domain is soluble²² and amenable to crystallisation and Biophysical assay experiments is extremely advantageous and makes it an even more favourable target choice. A research paper published in 2017 by F. Svensson, A. Bender and D. Bailey²² reviews the use of FBS in identifying active PDE inhibitors. The paper reports a small selection of 6 fragment-sized PDE inhibitors advanced to clinical trials, developed from a library consisting of thousands of fragments following screening and optimization techniques like those previously mentioned²².

Hence in this project, a small set of 31 compounds put together by Dr. Fredrik Svensson from IOTA Pharmaceuticals⁴⁶ – a research company specialising in fragment-based drug discovery - will be used for several different screening methods against the validated TbrPDEB1 target enzyme. The small fragment library was put together by cross referencing commercially available compounds from vendor catalogues with compounds

possessing known human PDE activity as reported by ChEMBL⁴⁷ – a publicly available database which archives compounds and their bioactivity data against certain drug targets.

The main tasks of this project are therefore to:

- Transform gene expression host with TbrPDEB1 catalytic domain construct plasmid
- Grow cells containing TbrPDEB1 plasmid construct
- Induce TbrPDEB1 target protein production
- Purify target TbrPDEB1 protein
- Run a quality control by Proton 1D NMR on the purchased fragment library
- Confirm PDE enzyme activity by Biochemical Assays
- Screen fragment library against protein target by X-Ray Crystallography
- Screen X-Ray Crystallography fragment hits against protein target by Biochemical assays
- Screen X-Ray Crystallography fragment hits against protein target by NMR
- Analysis of screening experiment hits, ranking ligands by potency / activity and ‘cherry-picking’ fragments with multiple screening confirmations
- Hit-Lead follow up by performing a substructure similarity search on the determined best fragments in order to find chemotypically similar compounds to expand the parasitic PDE active fragment library

Additionally, the overall aim of this project is to identify novel compounds as active TbrPDEB1 inhibitors through the use of several biochemical and biophysical techniques such as those described in the list of tasks above.

2 Experimental

2.1 Chemicals and Materials

The chemicals, reagents and materials used in this research project were purchased from a number of different suppliers, listed as following:

Bacterial growth media reagents such as Tryptone, Yeast Extract, Agar and Agarose were purchased from Melford Laboratories Ltd, in addition to HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), Tris (tris(hydroxymethyl)aminomethane) and IPTG (isopropyl β -D-1-thiogalactopyranoside). Chloramphenicol and Kanamycin antibiotics, as well as buffer reagents such as Guanidine, Calcium Chloride, Zinc Chloride, Magnesium Chloride, PEG 3350 (polyethylene glycol), Sodium Formate, Imidazole, MES (4-Morpholineethanesulfonic acid) and BME (2-Mercaptoethanol) were purchased from Sigma Aldrich. Other buffer reagents and general chemicals such as Sodium Chloride, DMSO (Dimethyl sulfoxide), SDS (Sodium dodecyl sulfate), 100% (v/v) Glycerol, Acetic acid, Absolute Ethanol, Hydrochloric acid, NuPAGE MOPS SDS running buffer (20x) and EDTA (Ethylenediaminetetraacetic acid) were purchased from ThermoFisher Scientific. Coomassie Brilliant Blue G-250 dye, as well as PageRuler Plus Prestained Protein SDS-PAGE Ladder were both purchased from Bio-Rad and ThermoFisher Scientific, respectively. Premade NuPAGE 4-12% Bis-Tris Gels for electrophoresis were purchased from Novex. 10/20 μ l, 200 μ l and 1250 μ l XL Graduated Tips were purchased from TipOne. Hampton 24-well plates and crystallography cover slips were purchased from Hampton Research.

10 mg of 31 different compounds of varying chemotypes were purchased from Enamine, all of which were manually chosen due to their known binding affinity to human PDEs. These fragments can be seen listed below: (additional information can be found in *Appendix 1* in the appendices)

1. N-benzylthieno[3,2-d]pyrimidin-4-amine"
2. N-cyclopentyl-3,4-dimethoxybenzene-1-sulfonamide"
3. 7-butyl-3-propyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione"
4. 1,2-dihydrophthalazin-1-one"
5. 6,7-dichloro-1H,2H,3H,5H-imidazolidino[2,1-b]quinazolin-2-one hydrochloride"
6. 3-propyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione"
7. 6-methyl-2-oxo-5-(pyridin-4-yl)-1,2-dihydropyridine-3-carbonitrile"
8. 4-ethyl-5-(pyridine-4-carbonyl)-2,3-dihydro-1H-imidazol-2-one"

9. 3-ethyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione"
10. 1-(3,4-dimethoxybenzoyl)-1,2,3,4-tetrahydroquinoline"
11. 6-methoxy-1,2,3,4-tetrahydroquinoline-2,4-dione"
12. 2-methyl-5-(methylsulfanyl)-[1,2,4]triazolo[1,5-c]quinazoline"
13. 5-(ethylsulfanyl)-2-methyl-[1,2,4]triazolo[1,5-c]quinazoline"
14. 3-methyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione"
15. 3-ethyl-5-[(3-hydroxy-4-methoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one"
16. N-cyclopentyl-2-phenylquinazolin-4-amine"
17. N-benzyl-2-chloroquinazolin-4-amine"
18. 6-(3,4-dimethoxyphenyl)pyridazin-3-ol"
19. 3-phenyl-2-sulfanyl-3,4-dihydroquinazolin-4-one"
20. 5-[(3,4-dimethoxyphenyl)methylidene]-1,3-thiazolidine-2,4-dione"
21. 1,3-dimethyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione"
22. 4-chloroquinazoline"
23. 4,6-dichloro-2-cyclopropyl-5-methylpyrimidine"
24. N-(2-chlorophenyl)-3,4-dimethoxybenzamide"
25. N-[(2H-1,3-benzodioxol-5-yl)methyl]quinazolin-4-amine"
26. 5-[(3-ethoxy-4-methoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one"
27. 2-({2-methyl-[1,2,4]triazolo[1,5-c]quinazolin-5-yl} sulfanyl)acetonitrile"
28. 5-[(4-ethoxy-3-methoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one"
29. N,N,2-trimethyl-5-nitrobenzene-1-sulfonamide"
30. 5-[(3,4-dimethoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one"
31. 5-[(2,3-dimethoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one"

Various different screening experiments were performed using these ligands, as will be presented and discussed later in this report. Note that most of the experimental procedures found in this report were inspired by and adapted from several relevant research papers including a 2014 Fragment-based Screening report by A.R. Blaazer *et al*⁴⁸, and a similar report by Dr. Fredrik Svensson *et al*²² from IOTA Pharmaceuticals.

2.2 TbrPDEB1 Catalytic Domain Construct

The TbrPDEB1 catalytic domain plasmid construct was provided by the host laboratory at the University of Kent ligated in a pET28a bacterial expression vector. The plasmid map of this vector (generated by using the SnapGene software⁶⁹), can be seen in *Figure 10*. Some key features of this construct include the KanR kanamycin resistance gene at base pairs 560-1375 and the T7 promoter for bacteriophage T7 RNA polymerase, responsible for DNA transcription. The lac operator is repressed by the lacI repressor gene in order to inhibit the transcription in *E. coli* (Escherichia coli) cells. Upon addition of IPTG reagent, transcription is permitted as a result of the binding of IPTG and the lac repressor. The release of the repressor from the operon binding site allows T7 RNA polymerase to transcribe the DNA sequence. This transcription manipulation is beneficial as protein production can be induced whenever desired, allowing maximum cell growth prior to induction thus to increase yield as much as possible. Additionally, IPTG is not metabolized by *E. coli* cells and hence the rate of protein expression remains constant. The His₆tag and thrombin site can also be seen at base pairs 5083-5127, serving the purpose of characterising the target TbrPDEB1 protein with a long histidine tail which is essential for facilitating protein purification, as will be later discussed in this report.

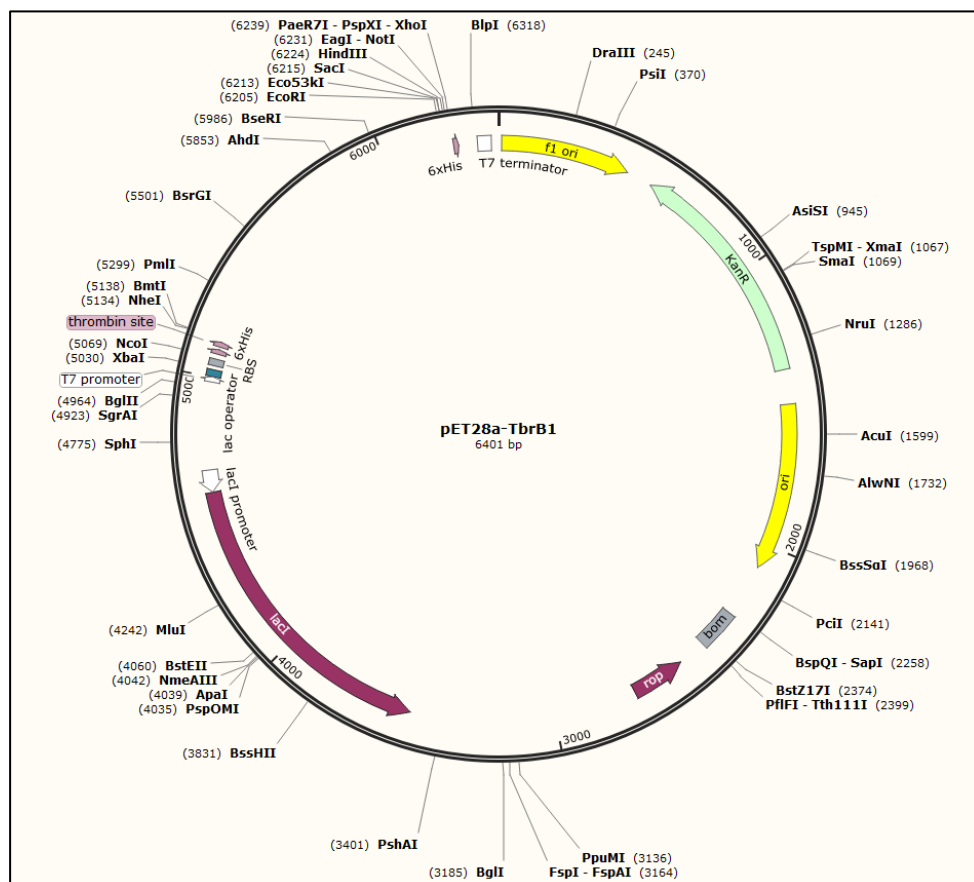


Figure 10:
pET28a-TbrPDEB1 plasmid map, generated by SnapGene

2.3 Preparation and Transformation of Competent cells

Note that all experiments involving bacteria handling were done using appropriate aseptic techniques e.g. in a sterile environment with the use of flames to kill bacteria and sterilizing all tools and media (by heating at 121 °C, pressure 1 bar in an autoclave) thus to reduce the likelihood of bacterial contamination. After each experiment, the workplace was cleaned using 70% ethanol and every container such as bottles and flasks were cleaned thoroughly using Virkon disinfectant.

Making *E. coli* cells 'competent' is a common technique used to make cells more likely to incorporate foreign DNA into their cytoplasm, thus allowing the manipulation of biological processes. There are several different ways of making *E. coli* cells competent; most commonly being by chemical or electroporation techniques. In this project, chemical methods were used to make *E. coli* cells competent. The BL21(DE3*) Codon-Plus Gram-negative *E. coli* bacteria strain is an optimal host in this experiment for the expression of recombinant protein as the DE3 lysogen hosts the T7 RNA polymerase gene which controls the T7 promoter in the lac-operon mechanism essential for the translation process as previously mentioned, whilst also providing resistance to chloramphenicol, allowing the antibiotic to be used as a selection marker. The thin cell wall of this Gram-negative bacteria also allows for easier cell lysis.

Hence, under sterile conditions, a sample of this *E. coli* strain (purchased from Agilent) was used to restreak an LB agar plate (10g tryptone, 5g yeast extract, 5g sodium chloride, 15g agar per litre of dH₂O, sterilised using an autoclave) along with 100µl of chloramphenicol. The plate was incubated at 37 °C for 24 hours. A single colony was used to inoculate 20ml of LB media (10g tryptone, 5g yeast extract, 5g sodium chloride per litre of dH₂O, sterilised using an autoclave) before the cells were left to incubate and grow to an optimal cell density of OD₆₀₀ = 0.3, measured on a Spectrophotometer apparatus. A graph tracking the growth of the competent cells can be seen in *Figure 11* on the following page:

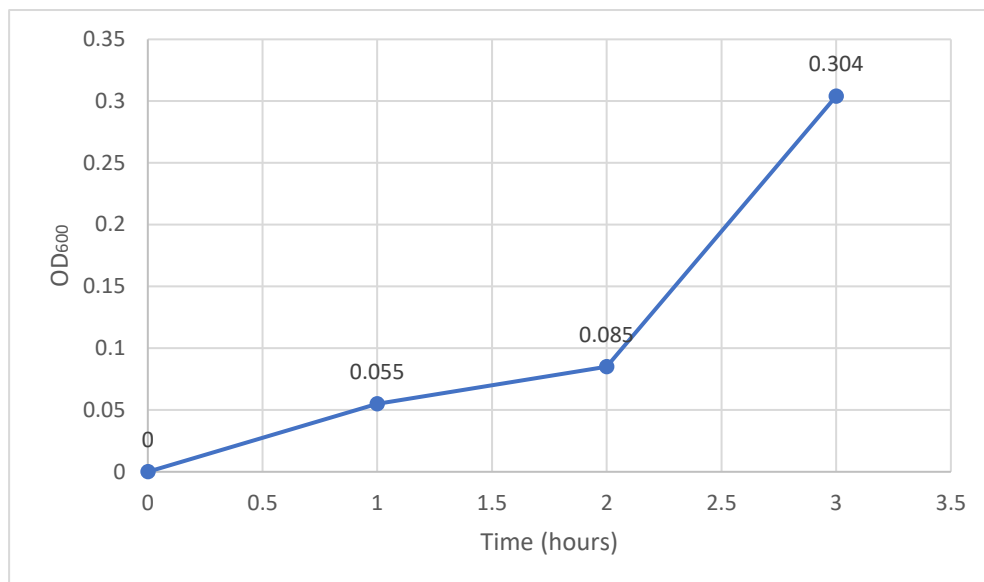


Figure 11:
Graph depicting BL21(DE3) Codon-Plus competent cell growth*

The cells, maintained under cold conditions to prevent further cell growth, were centrifuged and resuspended with 0.1M calcium chloride before being aliquoted out (100µl in 1.5ml Eppendorf tubes) and frozen at -80 °C. The purpose of aliquoting is to minimise freeze-thaw cycles which is known to reduce cell integrity. The calcium chloride treatment is key to chemically create cell competence, as the calcium ions increase cell wall permeability allowing for uptake of foreign DNA.

The following step is the genetic alteration, or transformation of the cells by introducing foreign DNA into the system. As previously stated, the catalytic domain of the TbrPDEB1 enzyme between residues 668 and 905, shown in *Figure 12*, was isolated and ligated into a pET-28a host vector creating the construct as seen in *Figure 10*.

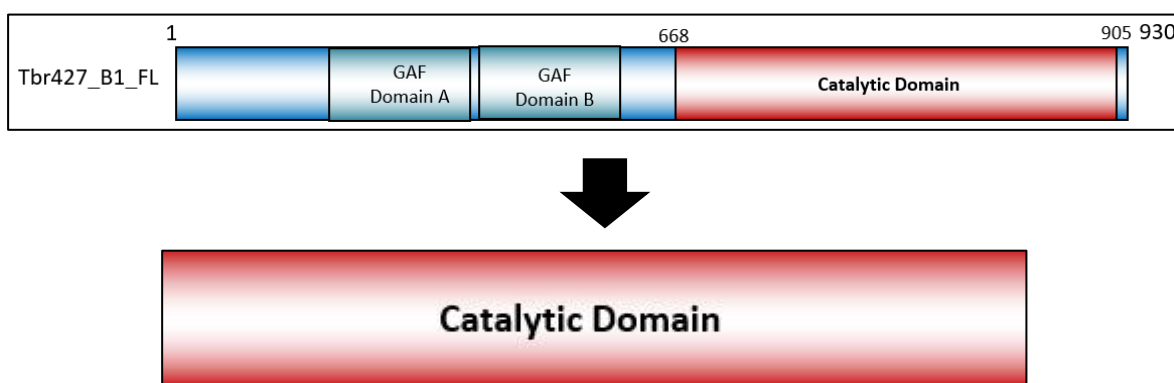


Figure 12:
TbrPDEB1 catalytic domain used for transformation

1µl of this plasmid construct was added to an 100µl aliquot of BL21(DE3*) Codon-Plus competent cells, which were left on ice for 15 minutes, following a 1-minute shock of heat at 42 °C in a water bath. The cells were then put back on ice for 2 minutes. 300µl of SOC

medium (2g tryptone, 0.5g yeast extract, 1M sodium chloride, 1M potassium chloride, 2M Mg^{2+} and 20% glucose, per 100ml of H_2O , sterilised using an autoclave) was added to the cells before they were left to incubate at 37 °C for 30 minutes to recover from the heat shock. The purpose of this heat shock is to alter the fluidity of the cell membrane, which is now permeable due to the cell being competent, by creating pores. These pores, created by the sudden increase in temperature, allow the pET-28a-TbrPDEB1 plasmid construct to pass through the pores and enter the cell. The pores once again close upon temperature decrease, thus to keep the foreign DNA within the cell, and prevent any other DNA from entering. The cells were then spread on a LB agar plate with 100mM chloramphenicol and kanamycin antibiotics, as the newly transformed cells have chloramphenicol resistance from the original *E. coli* strain, and kanamycin resistance from the original plasmid construct. The plate was left to incubate at 37 °C for 24 hours. A second plate was left to grow alongside as a control, with no plasmid added, hence no kanamycin resistance gene being present in the cells. As expected, there was no bacterial growth on the control plate as the cells were killed by the kanamycin antibiotic, whilst on the experimental plate colonies grew healthily indicating successful uptake and expression of the introduced recombinant genetic material.

2.4 Cell Growth and Expression of Target Protein

A single colony was picked out from the plate and used to inoculate a 250ml Erlenmeyer flask containing 50ml sterile LB media and 50 μ l of 100mM chloramphenicol and kanamycin antibiotics. This starter culture was left to grow overnight for a further 24 hours, incubated at 37 °C and 200rpm (rotations per minute). 25ml of starter culture was used to inoculate 2L of sterile 2YT media (16g tryptone, 10g yeast extract, 5g sodium chloride per litre of dH_2O , sterilised using an autoclave), in 2 x 2.5L flasks with the addition of 1ml of chloramphenicol and kanamycin antibiotics. The purpose of antibiotic addition is to maintain the function of the new genetic material by expressing the respective genes; in this case being antibiotic resistance, as without gene expression protein production cannot be induced. Hence, the cells were left to grow in sterile 2YT media at 37 °C up to an optimal OD_{600} reading of 0.6-0.8 as seen in a similar experiment in A.R. Blazeer's *et al* 2015 paper⁴⁸. An OD_{600} reading was taken hourly using a Spectrophotometer apparatus, thus to keep track of the rate of cell division. Note that a reference sample was taken from one flask directly prior to inoculation. The data recorded can be seen tabulated in *Table 3* on the following page.

Table 3:
Transformed competent cell growth

Time (hrs)	Flask 1 OD₆₀₀	Flask 2 OD₆₀₀
0	0	0
1	0.172	0.151
1 ½	0.354	0.314
2	<u>0.675</u>	0.597

Flask 2 was left to incubate for a further 10 minutes, reaching an ideal OD₆₀₀ value of 0.818.

Having reached optimal cell density, the flasks were left in a cold room at 4 °C for 20 minutes to prevent further cell division. Protein production was induced with the addition of 300mM IPTG, initiating transcription and translation of protein as previously discussed. The flasks were left to incubate at 18 °C for 48 hours, maximizing protein production. The TbrPDEB1 protein-containing cells were collected 48 hours later and harvested via centrifugation at 5000rpm for 10 minutes. The cells form a cluster, or a ‘pellet’ at the bottom of each flask due to the strong centripetal force created by the rapid spinning. The collected pellets weigh a total of 21.40g of protein-containing cells from a total of 2L of media. Note that from this stage, all work was done under cold conditions, either with ice or by setting any machinery used to 4 °C, thus to prevent denaturing or disturbing the protein.

The cell pellets were then resuspended and diluted in lysis buffer (10ml of buffer per gram of cells), which consists of:

Lysis buffer –

- 20mM Tris-HCl pH 7.5
- 200mM Sodium Chloride
- 10mM Imidazole
- 5% Glycerol
- 2mM BME
- 2 Protease Inhibitor Tablets

The resuspended cells were then lysed by a French press cell disruptor under 20,000 KPSI (kilo pounds per square inch) in order to release the TbrPDEB1 protein in the cells into a

suitable environment of the lysis buffer. The cell lysis solution was once again centrifugated at 17,000rpm for 30 minutes in order to separate the unwanted cell debris and the desired, TbrPDEB1 protein-containing supernatant.

2.5 Purification of Target Protein

The desired target protein acquired is contaminated with other unwanted proteins produced during the expression process. The following step is to separate the desired protein from the lysate mixture by means of several different chromatographic techniques. A Fast Performance Liquid Chromatography (FPLC) pumping mechanism is used to control the rate at which the mobile phase (in this case the cell lysate) passes through the stationary phase. The stationary phase varies depending on the separation technique. The machinery detects separation of proteins by using an ultraviolet detector, as proteins absorb UV light at 280nm mostly due to the aromatic rings present in amino acid residues. This data can be tracked and analysed using AKTA⁷⁰ – a liquid chromatography software. The purification stage was split up in a number of different techniques:

2.5.1 Immobilized Metal Affinity Chromatography (IMAC)

The principle of IMAC works by using the terminal histidine base pairs on the original vector as previously mentioned; Histidine has a high binding affinity to and is able to form a chelate-metal complex with several types of immobilized metal ions including nickel, cobalt, and copper under specific buffers. Following protein expression, the base pairs shown are translated to a chain of 6 His-tags attached to the desired TbrPDEB1 protein. Therefore, the protein lysate can be passed through a column packed with a Chelating Sepharose resin stationary phase charged with Ni²⁺ ions, which will selectively retain any proteins containing this histidine tag, thus isolating desired protein within the Nickel column. The column can then be washed several times with buffer, separating out undesired protein still remaining in the column, before the desired protein can be eluted out by washing through with imidazole – which has a greater binding affinity to Nickel than histidine, and will therefore displace the bound protein from the column, allowing for collection. Following elution, the collected protein is passed through an additional larger column for the purpose of desalting the solution by separating out the imidazole from the protein by size exclusion, as the larger salt molecules pass through the tightly packed resin beads of the desalting column much slower than the protein. Instead, the protein is buffer exchanged with desalting buffer. This step is required as the presence of imidazole in a protein solution for long periods of time can cause the stripping of divalent metal ions

present in proteins to occur, which in turn would cause the protein to lose activity. This will be discussed in greater detail later in this report.

Hence, three different buffers were made up –

Buffer A (Binding buffer) –

- 20mM Tris-HCl pH 8
- 200mM Sodium Chloride
- 20mM Imidazole
- 5% Glycerol
- 2mM BME

Buffer B (Eluting buffer) –

- 20mM Tris-HCl pH 8
- 200mM Sodium Chloride
- 1M Imidazole
- 5% Glycerol
- 2mM BME

Buffer C (Desalting buffer) –

- 50mM Tris-HCl pH 8
- 100mM Sodium Chloride
- 5% Glycerol

Note that in every purification experiment, buffers were made up in cold bottles, as the pH of Tris-HCl is known to be temperature dependant, and as the experiment is done in the FPLC machinery at 4 °C, bottles were pre-chilled thus to ensure pH remained constant. Additionally, all solutions used were filtered and degassed accordingly to remove any air and small particles in order to prevent blockage of the FPLC pumps. Pumps A and B on the machinery were initially washed with deionised H₂O and equilibrated in buffers A and B, respectively.

A 5ml HisTrap™ HP column prepacked with charged Ni²⁺ Sepharose beads was connected to the FPLC and washed with 10 column volumes (C.V.) of deionised H₂O (50ml), and equilibrated with 10 C.V. of Buffer A. Similarly, a 54ml HiPrep™ desalting column prepacked with Sephadex G-25 Fine, cross-linked dextran resin was washed with 2 column

volumes of deionised H₂O (108ml) and equilibrated with 2 C.V. of desalting buffer. The cell lysate was loaded into the Nickel column through pump A, before contaminants were separated out by washing the column with a further 10 C.V. of Buffer A. The wash through and the flow through were collected. The his-tagged protein bound to the Nickel column was then eluted with an increasing imidazole concentration (using Buffer B) at a linear gradient and collected in fractions. Fractions yielding a large UV peak on the AKTA software (as can be seen in the Results and Discussion section) indicating protein elution were sampled and used for an SDS-PAGE experiment in order to visually evaluate the purity and basic characteristics of the yielded protein following the first purification step by inspection of a polyacrylamide gel. The appropriate fractions were combined and passed through a desalt column equilibrated in Buffer C as described previously in order to remove imidazole and buffer exchange the yielded protein. The protein concentration was also measured using a Spectrophotometer apparatus.

2.5.1.1 SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis, or SDS-PAGE is an analytical method used in biochemistry with the principle of separating charged molecules in an electric field based on their molecular weight, or size. SDS is a detergent that denatures and unfolds protein structures into its original subunits by binding to the polypeptide chains to form complexes with a constant charge to mass ratio which negatively charges proteins. Hence, 30µl samples of the loaded lysate, flow through, wash through, and eluted protein fractions were taken and mixed with 30µl of 2x SDS Sample Buffer, consisting of:

SDS Sample Buffer (2x) –

- 10mM Tris-HCl pH 6.8
- 4% (w/v) SDS (sodium dodecyl sulfate; electrophoresis grade)
- 0.2% (w/v) Bromophenol blue
- 20% (v/v) Glycerol
- 200mM DTT (Dithiothreitol)

The mixed samples were then heated to 150 °C on a heat block to further denature the protein by reducing disulphide bonds. 20µl of each heated mixture, and 5µl of a Precision Plus Protein Dual Colour standard are pipetted into the wells of a gel cassette. An electric current is applied across the gel, which is immersed in NuPAGE MOPS SDS Running Buffer (1x), causing proteins with a negative charge to migrate towards the positive anode

at the bottom of the cassette. The rate at which a certain protein travels across the gel is determined by its size – smaller proteins being able to pass through the pores of the gel matrix at a faster rate are expected to appear more towards the base of the cassette, and vice versa. This way, every protein in the samples taken can be analysed on the gel by their molecular weight, which can be determined by comparing to the known standard also applied. Hence, the desired TbrPDEB1 protein with a calculated molecular weight of 40.57 kDa (kilodaltons) can be seen in the Results and Discussion section.

2.5.2 Secondary IMAC

The protein sample is seen to contain additional contaminants, hence further purification steps are required. Prior to the next purification step, the protein sample was diluted with a further 50ml of desalting buffer to prevent protein aggregation and eventual crashing. 75µl of Thrombin protease enzyme (2 units/microlitre) was added to the sample (determined by adding 10 units of enzyme per mg of protein) in order to cleave the histidine tag present on the desired protein, before being left at 4 °C for 24 hours. The columns used were cleaned accordingly.

A small, table-top 5ml column with a sieve and Chelating Sepharose Fast Flow resin was charged with 2.5ml of a 0.1M solution of NiSO₄, pH 6. The column was prepared appropriately for a manual IMAC experiment, adapted from the example previously discussed. The protein sample was loaded into the column, which was equilibrated in 10 C.V. of desalt buffer (20ml). Assuming successful thrombin his-tag cleavage, the desired, untagged TbrPDEB1 protein can be collected in the flow through. Any remaining untagged protein is washed through with an additional 10 C.V. of desalt buffer, leaving any contaminants with remaining histidine tags bound to the Nickel column. Note that the flow through and the wash through, assumed to contain desired protein, are combined. The remaining his-tagged contaminants are eluted with 10 C.V. of desalting buffer containing 300mM imidazole. SDS-PAGE samples were taken after each step, in order to observe cleavage quality. The column used was cleaned accordingly. The combined wash and flow solution was diluted with 100ml of desalt buffer and was prepared for the next purification step.

2.5.3 Ion Exchange Chromatography (IEX)

The method to IEX-Q, or Ion Exchange Chromatography-Quaternary Ammonium is similar to the IMAC procedure; however instead of separation by metal-chelate affinity,

separation is based on charge difference. The principle behind this method works by exploiting the charge difference between different molecules or proteins in a sample. A desired protein or molecule in a sample can ionically bind to an oppositely charged stationary phase; Similarly charged contaminants are repelled by the stationary phase, or contaminants with varying charges bind very weakly and hence elute much easier – in any case, separation occurs, once again isolating the desired protein from any remaining impurities. In order for a target protein to bind to the stationary phase, its isoelectric point (pI) (the point at which the protein carries a net zero charge) has to be determined. Once this is known, the pH of the buffers and hence charge of the stationary phase can be manipulated to give ideal ionic binding to the desired protein.

The isoelectric point of a protein is determined by working out the isoelectric point of every amino acid in the protein structure and taking the average. In turn, the pI of amino acids can be calculated by averaging the total pK_a of all the functional groups in the residue – giving a value of pH at which the respective amino acid is zwitterionic.

Using the amino acid sequence of the TbrPDEB1 protein, the pI is calculated to be 5.15 using a protein isoelectric point calculator (IPC)⁸⁵. Therefore, by increasing the pH past the pI, the protein becomes less protonated due to increasing OH⁻ ion concentration as seen in simple Lewis acid-base reactions, making the protein much more prone to ionically bind to a positive cationic stationary phase; in this case, the positively charged quaternary ammonium, NH₄⁺.

Hence, two different buffers are made up –

IEX-A (QA buffer) –

- 20mM Tris HCl pH 8
- 100mM Sodium Chloride
- 5% Glycerol
- 1mM BME

IEX-B (QB buffer) –

- 20mM Tris HCl pH 8
- 1M Sodium Chloride
- 5% Glycerol
- 1mM BME

Pumps A and B of the FPLC machinery were equilibrated in buffers QA and QB, respectively. A 5ml HiTrap™ Q HP column prepacked with Sepharose beads was washed with 10 C.V. of deionised H₂O, before being equilibrated with 10 C.V. of QA buffer. The protein sample was loaded into the column through pump A. The flow through was collected and sampled for an SDS-PAGE. Any weakly bound contaminants are washed through with an additional 10 C.V. of QA buffer. The desired protein, assumed to remain bound to the stationary phase, is eluted out by an increasing salt concentration (with QB buffer) at a linear gradient. Elution fractions were collected as described in the IMAC purification step. The increasing salt concentration causes bound proteins to desorb from the stationary phase. The rate of protein desorption is dependent on ionic interaction strength between the protein and the stationary phase. In terms of the desired protein, if pH is too low, causing the protein to be closer to net zero charge, the ionic interaction would be very weak and displacement would easily occur (thus failing to separate from protein contaminants). Having a low pH also risks protein aggregation and crashing. On the contrary, if pH is too high, it would be difficult to displace bound protein from the column, even at high salt concentrations. Hence, an optimal pH of 2 above the pI is used. The recorded AKTA graph can be seen in the Results and Discussion section. SDS-PAGE samples were taken from the appropriate fractions. The TbrPDEB1 protein containing fractions – as determined by the SDS-PAGE experiment – were combined. The concentration and yield were recorded. The next purification step, Size Exclusion Chromatography, was not required as the protein has reached ideal purity, indicated by no contamination bands being present on the SDS-PAGE gel (as can be seen on the IEX SDS-PAGE gel in the Results and Discussion section).

The combined fraction contents were buffer exchanged with Storage Buffer and concentrated using an Amicon Ultra-15 10k Centrifugal Filter Device, or a ‘concentrator’.

Storage Buffer consists of:

- 20mM Tris HCl pH 7.5
- 5% Glycerol
- 50mM Sodium Chloride
- 2mM BME
- 1mM MgCl₂ (Magnesium Chloride)

Once the QB buffer was removed, the concentrator was topped up with Storage Buffer and spun again. The protein precipitant was discarded. The protein sample was separated in 100µl aliquots and cryogenically frozen at -80 °C. The concentration was measured using a Spectrophotometer apparatus, using an extinction coefficient of 24,420 and molecular weight of 40.57 kDa as characteristics for the TbrPDEB1 protein.

2.6 Biochemical Assays of Target Protein

Note that following purification of the target protein, the 31 solid 10mg compounds purchased were used to prepare 100mM and 400mM compound super stocks in 100% DMSO which were used for several experiments as will be later discussed.

A Lonza PDELight HTS cAMP phosphodiesterase kit⁵⁰ was used to perform several experimental assays in order to:

- Confirm TbrPDEB1 enzyme activity
- Optimise concentrations of experimental reagents
- Assess inhibition of purchased compounds against TbrPDEB1
- Determine the Km of the purified enzymatic TbrPDEB1 protein
- Produce Dose-Response curves of active compounds against TbrPDEB1

This bioluminescent detection assay works by exploiting the previously discussed reaction pathway in *Figure 4*. Assuming a working PDE, the amount of AMP converted from cAMP as a result of a hydrolysis reaction catalysed by PDE enzyme can be indirectly quantified through this assay. This can be seen as an equation in *Figure 13*:

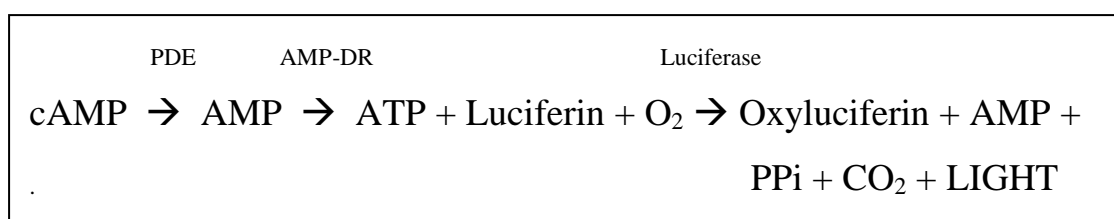


Figure 13:
Lonza PDELight assay kit reaction scheme

As can be seen, the AMP detection reagent used in this assay kit converts AMP to ATP (adenosine triphosphate). The luciferase also present in the kit is involved in the production of light over time from ATP, luciferin and free oxygen, as can be seen in the reaction scheme in *Figure 13*. The light emitted can be measured by a FLUOstar Omega

Luminescence plate reader (from BMG LABTECH) and is directly proportional to the quantity of AMP broken down from cAMP. Hence, this is beneficial for PDE activity studies.

2.6.1 Verifying TbrPDEB1 activity and determining working concentrations

Note that this protocol was adapted from the official Lonza PDELight protocol⁵⁰ as will be discussed further in greater detail.

20µl of TbrPDEB1 protein at a concentration of 1mg/ml was used to prepare a 200µl serial dilution as follows: 1:100, 1:300, 1:500 and 1:1000, 1:10,000 with Stimulation Buffer, consisting of:

- 50mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) pH 7.8
- 100mM NaCl
- 5mM MgCl₂
- 0.1mM ZnCl₂ (Zinc Chloride)
- 5% Glycerol

10µM aliquots of cAMP and AMP were also prepared and diluted with Stimulation Buffer. 80% Detection Reagent (diluted with Reconstitution Buffer) was also prepared.

17.5µl of each PDE dilution was pipetted into a white Corning 96-well assay microplate, followed by the addition of 17.5µl of Stimulation Buffer consisting of 6% DMSO (dimethyl sulfoxide). This was repeated 3 times for each dilution (in triplets). The plate was covered and incubated at room temperature for 20 minutes following this step. Also note that the plate was spun for several seconds after each step, in order to mix the added reagents and prevent protein aggregation. 35µl of detection reagent was then added, followed by 35µl of cAMP. In addition to the triplets of dilutions, positive and negative controls were also ran simultaneously in triplets. 35µl of AMP was directly added to the positive control instead of cAMP, whilst for the negative control the 17.5µl of TbrPDEB1 was replaced with 17.5µl of Stimulation Buffer. The plate was read as previously described for a duration of 55 minutes.

A similar assay was also performed with a constant optimal TbrPDEB1 concentration and a similar serial dilution of cAMP, thus to determine enzyme kinetics by working out the K_m through a Michaelis-Menten graph. These graphs are seen plotted in the Results and

Discussion section. A variation of this protocol was also used to run the fragment screen following determination of a suitable TbrPDEB1 concentration, with the addition of 17.5µl of fragment in solution at a concentration of 200µM (made by dilution of a 100mM super stock with Stimulation Buffer). The plate was once again read for 55 minutes, and a suitable graph was plotted, as will be later discussed in the Results and Discussion section.

2.7 X-Ray Crystallography

A crystallisation protocol was adapted from a relevant 2013 study by C. Jansen *et al*⁴⁹. As a brief overview, X-ray Crystallography work was conducted in two parts:

The first being crystallisation via XCHEM – a semi-automatic lead Fragment-based discovery approach, allowing the user to quickly set up nanolitre crystallisation drop plates and rapidly screen fragment libraries consisting of large numbers of varying ligand chemotypes using advanced machinery at the Diamond Light Source⁷⁸ in Oxfordshire; further details will be discussed later on in this report.

The second part being manual crystallisation – which included manually setting up crystal drops and crystal soaks at the University of Kent, a much slower and time demanding method, however more suitable for this project considering that the fragment library only consists of 31 ligands.

Nevertheless, despite using two varying crystallisation techniques, both methods use the same crystal soaking principle:

The aim of a crystal soaking experiment is to make highly concentrated super-stock solutions of each compound and use them to ‘soak’ crystallised TbrPDEB1 protein. Crystal soaking is a technique where yielded crystals are introduced to an environment containing the original crystallisation conditions in solution with a concentration of a given compound – dissolved in DMSO. In theory, the compound would be taken up by the PDE target as a result of the soak, depending on the binding affinity. The duration of the soak, concentration of compound and quality of crystal should all be considered – the longer the soak duration or the higher the soak concentration – the more likely it is that the compound is taken up by the PDE target. However, long exposures and high concentrations of DMSO can damage the crystal lattice and yield poor crystallographic data or completely shatter the crystal altogether. Therefore, a range of variables were tested to aid in optimising the conditions most suitable to the yielded crystals.

Following soaking, the compound containing crystals underwent X-Ray diffraction at the Diamond Light Source thus to yield crystallographic data and diffraction patterns which were used to visualise the structure three-dimensionally with the use of softwares such as CCP4i⁶⁶ and COOT⁶⁷.

2.7.1 Crystallisation by XCHEM

A 100µl aliquot of purified protein in Storage Buffer was buffer exchanged to Crystallisation Buffer, made up of:

- 20mM Tris HCl pH 7.5
- 5% Glycerol
- 2mM BME

The buffer exchanged protein was centrifuged down to a concentration between 7-7.5mg/ml, interpreted from a similar study by C. Jansen *et al*⁴⁹ as an optimal concentration for crystallisation, however it was found that crystals would form even at concentrations as low as 6mg/ml. The crystallisation conditions were made up as following:

- 100mM MES (2-(N-morpholino) ethane sulfonic acid pH 6.5
- 18% PEG (polyethylene glycol) 3350
- 400mM Sodium Formate
- 300mM Guanidine

30µl of crystallisation conditions were pipetted into each well of a 96-well 3-drop SwissCI dispensing plate purchased from Molecular Dimensions. 0.2µl + 0.2µl drops of protein were laid down on the wells of the dispensing plate using a Mosquito, allowing vapour diffusion crystallisation to occur via a sitting drop technique. This quickly super saturates the protein drops, yielding TbrPDEB1 crystals at 4 °C within 48 hours.

The crystal plate was analysed using a Formulatrix imager at the Diamond Light Source in order to select the best crystals and filter out poor drops. The wells containing the best crystals were selected and soaked with a similar dispensing plate containing the entire fragment library at a 100mM concentration, using an ECHO 555 dispenser – a precise nanolitre volume Liquid Handler. Soaks were done for 3 hours, with 20mM and 30mM compound concentrations. The soaked crystals were then mounted on loops using the aid

of a Shifter, before being hit with X-Ray beams. The crystallographic data was acquired from various different auto-processing pipelines such as fast-dp and autoPROC. An overview of the XCHEM workflow can be seen presented in *Figure 14* on the following page:

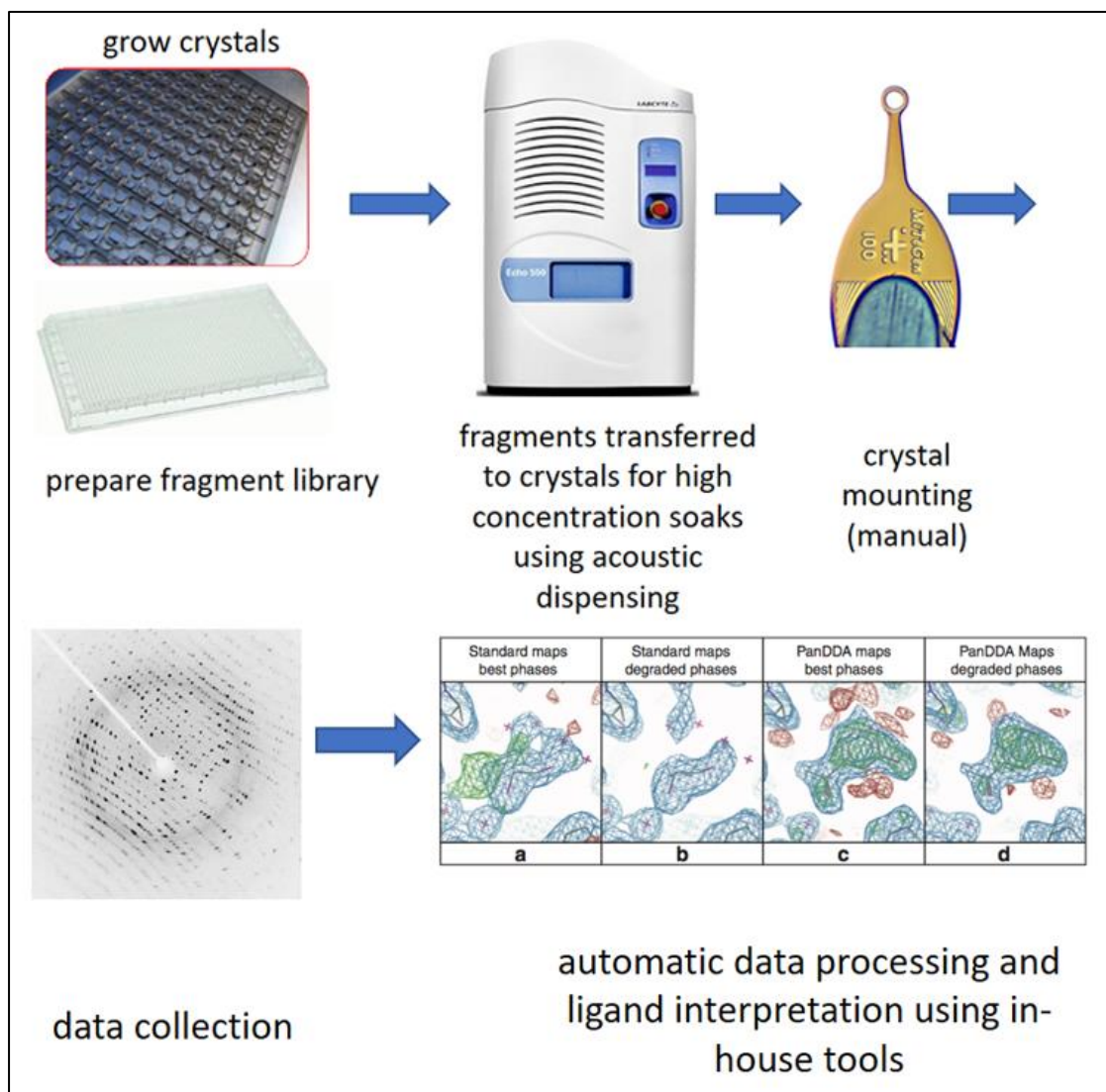


Figure 14:
XCHEM platform workflow A.K. Singh 2017

2.7.2 Manual Crystallisation

Manual crystallisation was set up in a similar way, with the exception of a 24-well Linbro plate being used. 500 μ l of crystallisation conditions were pipetted into all 24 wells of the dispensing plate. 1 μ l protein + 1 μ l crystallisation condition drops were combined on a circular glass slide, which was placed upside down on each well on the plate, creating a similar vapour diffusion with a hanging drop technique. This technique also yielded TbrPDEB1 crystals at 4 °C within 48 hours.

400mM compound super stocks were used to prepare compound soak solutions of varying concentrations. 1µl drops of compound solutions were laid out on circular glass slides, which would later be placed on a fresh dispensing plate as previously described. TbrPDEB1 crystals were manually 'fished' out of the original drops using circular 100-250-micron loops under a microscope and placed into the compound solution drops to allow soaking. Multiple crystals were placed in each drop to increase the chance of crystal survival whilst also varying soak time. Following soaking, the crystals were once again manually fished out and were briefly placed in a drop of Cryogenic Solution, containing 80% crystallisation conditions and 20% glycerol. The crystals were then archived in a puck which was kept immersed in liquid nitrogen at all times. The pucks were stored in a transportation Dewar decanted of liquid nitrogen and shipped to the Diamond Light Source to yield crystallographic data by undergoing X-Ray diffraction as previously described.

2.8 Nuclear Magnetic Resonance (NMR)

NMR experiments were conducted in a 600MHz Avance III spectrometer with Cryoprobe throughout this research project to analyse and verify the purity of the purchased compounds whilst also using more complex sophisticated NMR experiments to analyse the TbrPDEB1 protein itself and observe binding affinity to compounds of interest determined by X-Ray Crystallography. All NMR data was analysed and interacted with using Bruker's Topspin 3.5p17 NMR processing software⁷¹.

2.8.1 Compound Quality Control (QC)

100mM compound super stocks were used to make 1mM fragment solutions made up to 650µl of DMSO in a 5mm NMR tube. Proton 1D experiments were ran on all 31 fragments. The yielding spectra were analysed accordingly, identifying any impurities present. The tubes were cleaned appropriately.

2.8.2 Protein NMR

A 100µl aliquot of purified protein in Storage Buffer was concentrated down to 3mg/ml and buffer exchanged to Phosphate Buffer, made up of 100mM Sodium Chloride and 20mM phosphate pH 7.5. The protein in Phosphate Buffer solution was made up to 250µl in a 3mm NMR tube, which was spiked with 5% D₂O. A 1D Proton NMR experiment yielded a ¹H spectrum for the TbrPDEB1 protein, as can be seen later in this report.

2.8.3 Protein-Ligand Binding Affinity NMR

Compounds that gave the best crystallography results were brought forward for biophysical screening experiments to test for binding affinity. Two 3mm NMR tubes were made up as follows (for each compound tested):

Tube 1 – 0.5mM Ligand, 0.01mM Protein in 250µl of Phosphate Buffer, spiked with 5% D₂O

Tube 2 – 0.5mM Ligand in 250µl of Phosphate Buffer, spiked with 5% D₂O

The above tubes were used for STD (Saturation-Transfer Difference), CPMG (Carr-Purcell-Meiboom-Gill) and WaterLOGSY (Water-Ligand Observed via Gradient Spectroscopy) NMR experiments to test for binding affinity.

2.8.3.1 STD NMR

¹H STD experiments were acquired on a Bruker Avance III 600 MHz spectrometer at 25°C, using 16,384 points over a spectral width of 9615 Hz. A total number of 256 scans were acquired with an acquisition time of 0.85 seconds. Saturation was achieved using EBURP2⁸⁴ pulses over a 50ms saturation period; Saturation was applied at -0.3 / -30ppm for the on and off resonance experiments respectively. Water suppression was achieved using excitation sculpting⁸¹ and the spectra was referenced against water. Data were zero filled and Fourier transformed using topspin 3.5p7. A convolution filter⁸² was used to further suppress water the residual water signal (sfil, bcfw 0.2ppm) and an exponential line broadening of 3 Hz was applied and the spectrum was referenced against water^{74,75}.

2.8.3.2 waterLOGSY NMR

¹H WaterLOGSY⁷³ experiments were acquired on a Bruker Avance III 600 Hz spectrometer at 25°C, with 32,768 points (spectral width 9615 Hz) with a total number of 256 scans measured within an acquisition time of 1.70 seconds. Data was processed using topspin 3.5p17 and line broadening of 3.0 Hz was applied. Acquired data was processed in topspin 3.5p17. A convolution filter⁸² was used to further suppress residual water signals (sfil, bcfw 0.2ppm) with an exponential line broadening of 3.0 Hz. The spectrum was referenced against water. Water inversion was achieved using a 6.4ms eSNOB inversion pulse. The NOE mixing period was 1.2s during which a lower power of

10G/cm Z gradient pulse was applied. Further suppression of the protein signals was achieved using 20 cycles of a CPMG sequence with an inter-pulse delay of 1ms.

2.8.3.3 CPMG NMR

^1H CPMG⁷⁷ experiments were acquired on a Bruker Avance III 600 Hz spectrometer at 25°C. Data was collected with 16,384 points and a spectral width of 9615 Hz, with a total number of 128 scans and an acquisition time of 0.85s. The CPMG element used 600 cycles with a delay of 1ms between 180-degree pulses. Water suppression was achieved using a 3:9:19 Watergate sequence⁸³ and 100ms of pre-saturation. Acquired data was processed in topspin 3.5pl7, a convolution filter⁸² was used to further suppress residual water signals (sfil, bcfw 0.2ppm) with an exponential line broadening of 3.0 Hz and the spectrum was referenced against water.

3 Results and Discussion

3.1 Mass Spectrometry

A mass spectrometry experiment was performed on a 25 μ M sample of purified TbrPDEB1 protein following purification thus to ensure the 40.57kDa protein mass. This can be seen in *Figure 15* below:

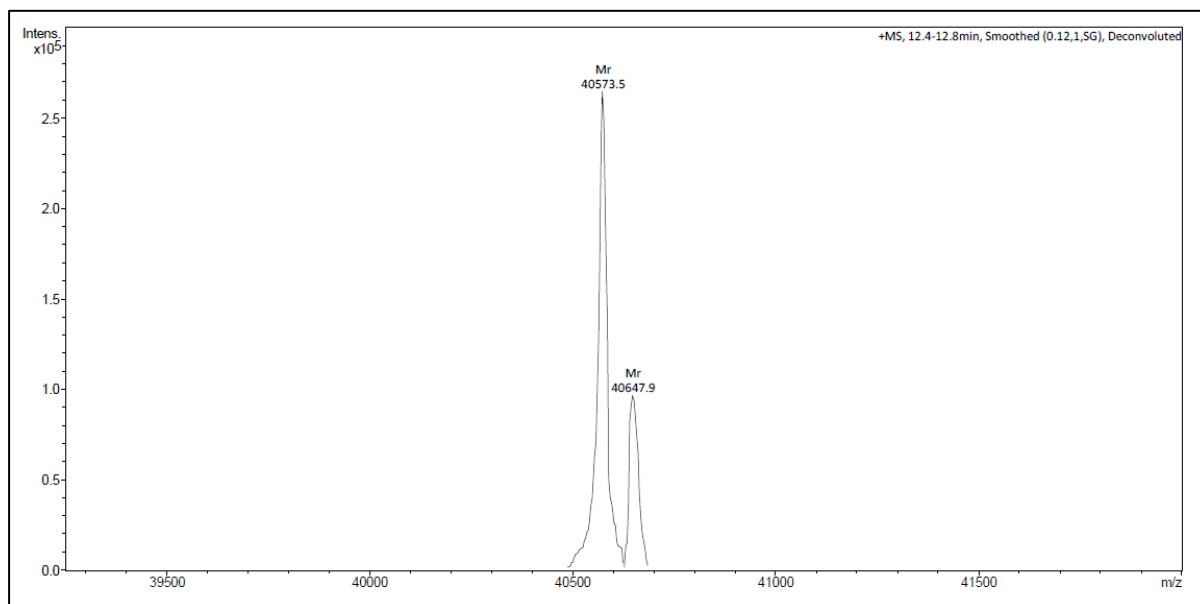


Figure 15:
TbrPDEB1 deconvoluted mass spectrum

As can be seen, the deconvoluted mass spectrum shows a large abundant peak of $\times 10^5$ intensity at mass 40573.5Da, the correct mass for the TbrPDEB1 protein target.

Additionally, a smaller peak is seen at 40647.9Da. The 74.4Da mass difference between the two peaks is due to a mass 74.4Da glycerol peak being present and detected in the sample due to remaining in the protein sample following storage buffer exchange. Thus, the presence of target protein within the purified sample of correct mass is confirmed.

3.2 Purification Analysis

This chapter is dedicated to discussing the results of the protein purification methods to further characterise the TbrPDEB1 protein and present the yield.

3.2.1 IMAC AKTA Graph

The AKTA graphs overviewing the IMAC purification process can be seen below:

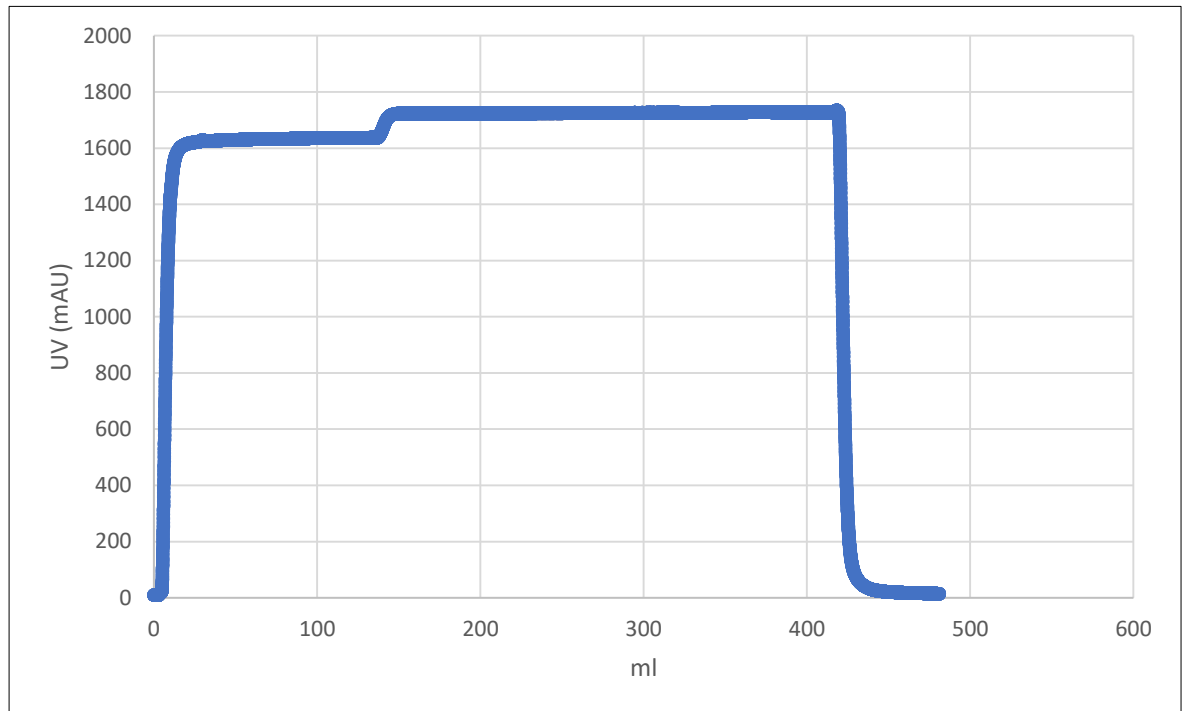


Figure 16:
TbrPDEB1 IMAC Sample Load and Wash graph

The IMAC purification process can be split up into three stages: Load, Wash, and Elution. The load and wash stages of the protein sample can be seen plotted on a graph in *Figure 16* above. As depicted, the ultraviolet detector reading is seen to spike directly after loading, quickly reaching the detection limit. The reason for this is due to the large quantity of protein present within the cell lysate mixture acquired after protein expression and induction – consisting of desired protein, but mainly contamination. After loading the 400ml of lysate mixture, the UV reading is seen to quickly decline as a result of the wash stage. Once the lysate has passed the detector and is fully loaded into the column, a further 10 C.V. of binding buffer is pumped through the system as previously discussed in order to wash any remaining weakly-bound protein contaminants out of the column. This stage continues until the UV reading plateaus as seen at 420ml, indicating binding buffer with no protein present and hence a successful wash, leaving the desired histidine-tagged protein bound to the Nickel column. This protein is then eluted as seen in *Figure 17*:

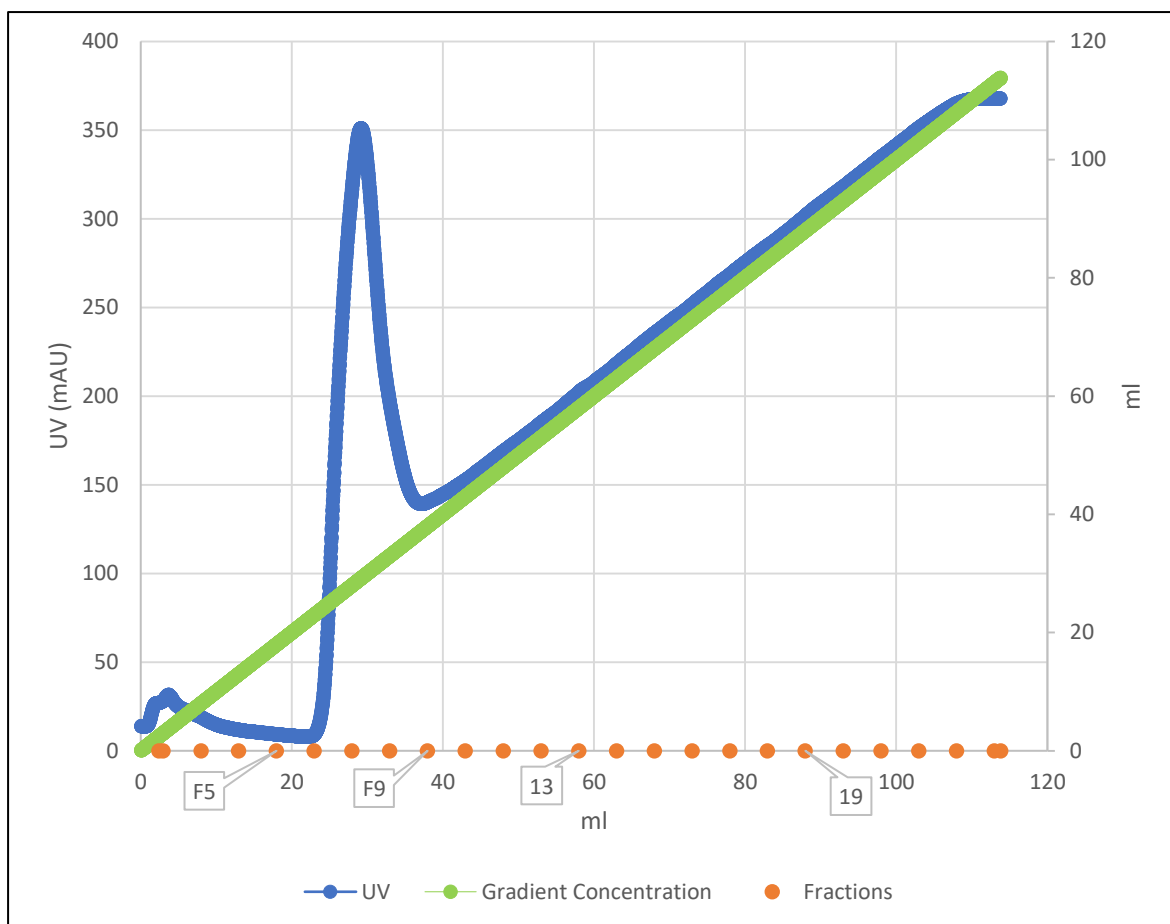


Figure 17:
TbrPDEB1 IMAC Elution and Fractionation graph

The protein is eluted with an increasing imidazole concentration at a linear gradient, shown in green in *Figure 17*. The increasing imidazole concentration displaces the bound histidine-tagged protein as indicated by the large UV peak seen at 22ml. The elution samples were collected in 5ml fractions until everything in the column was displaced, guaranteed by a 100% (1M) imidazole concentration. The UV is seen to continue to increase gradually in parallel with the gradient concentration; This however is not due to any protein elution; On the contrary, imidazole also has an ultraviolet absorbance at 280nm, explaining the linear UV trend. For a similar reason, concentration measurements at this stage are not accurate due to the imidazole present in the sample which is likely to give incorrect readings. Fractions 5, 6, 8, 9, 13 and 19 were collected and sampled to be analysed on a gel using the SDS-PAGE analytical technique.

3.2.2 IMAC SDS-PAGE Gel

An interpreted image of the SDS-PAGE Gel following IMAC purification can be seen in *Figure 18*. The PageRuler Plus Prestained Protein Ladder can also be seen in *Figure 19*⁵¹. The columns are labelled as follows (from left to right): ladder, supernatant, flow through, wash through, Fraction 5, Fraction 6, Fraction 8, Fraction 9, Fraction 13, Fraction 19.

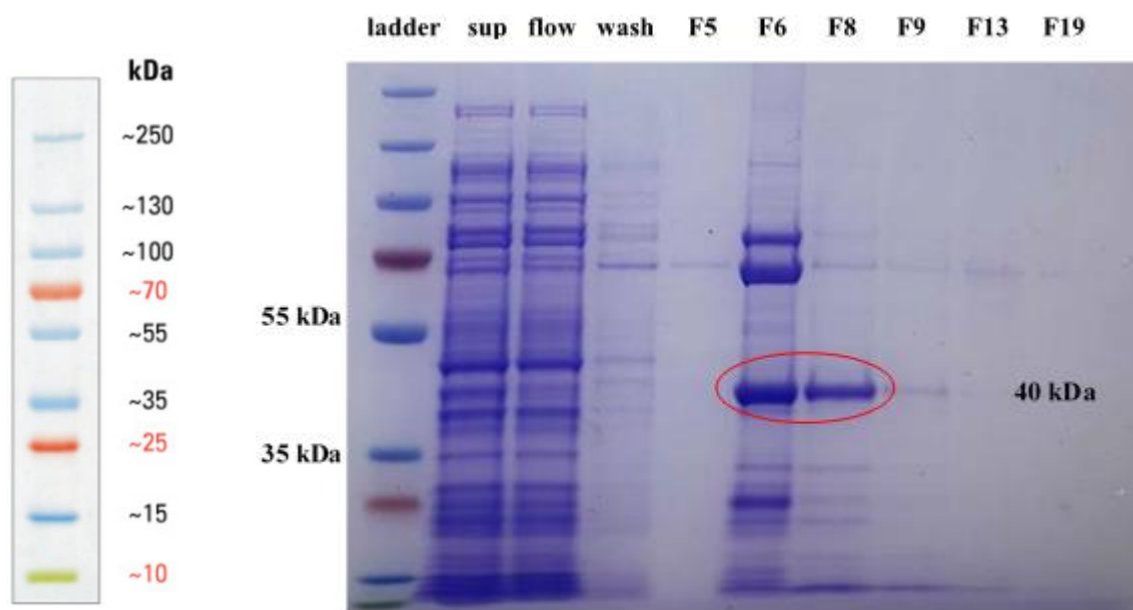


Figure 19:
*Protein Ladder*⁵¹

Figure 18:
TbrPDEB1 IMAC SDS-PAGE Gel

As expected, the supernatant and flow through samples yielded a large number of thick bands on the gel due to a large quantity of contaminant protein being present. The TbrPDEB1 protein of interest is seen present in Fractions 6 and 8, indicated by thick bands at 40kDa (as seen circled), deduced by comparing to the adjacent protein ladder figure. This is the expected position of TbrPDEB1 protein on the gel as the determined molecular weight by Mass Spectroscopy was measured to be 40.57kDa. This implies successful gene expression and confirms the presence of desired protein in the elution sample. Hence, fractions 6-8 were combined and buffer exchanged using a desalting column as previously discussed.

The yield was recorded as follows: (assuming 40.57 kDa molecular weight and 28,420 extinction coefficient parameters)

0.340mg/ml of protein in 44ml of elution sample = 14.96mg

The A260/A280 ratio was recorded as 1.100 (much higher than a 0.57 reading for pure protein⁵³, due to excessive contamination being present).

3.2.3 IEX AKTA Graph

The AKTA Graph, overviewing the elution stage of the IEX purification process can be seen in *Figure 20* below:

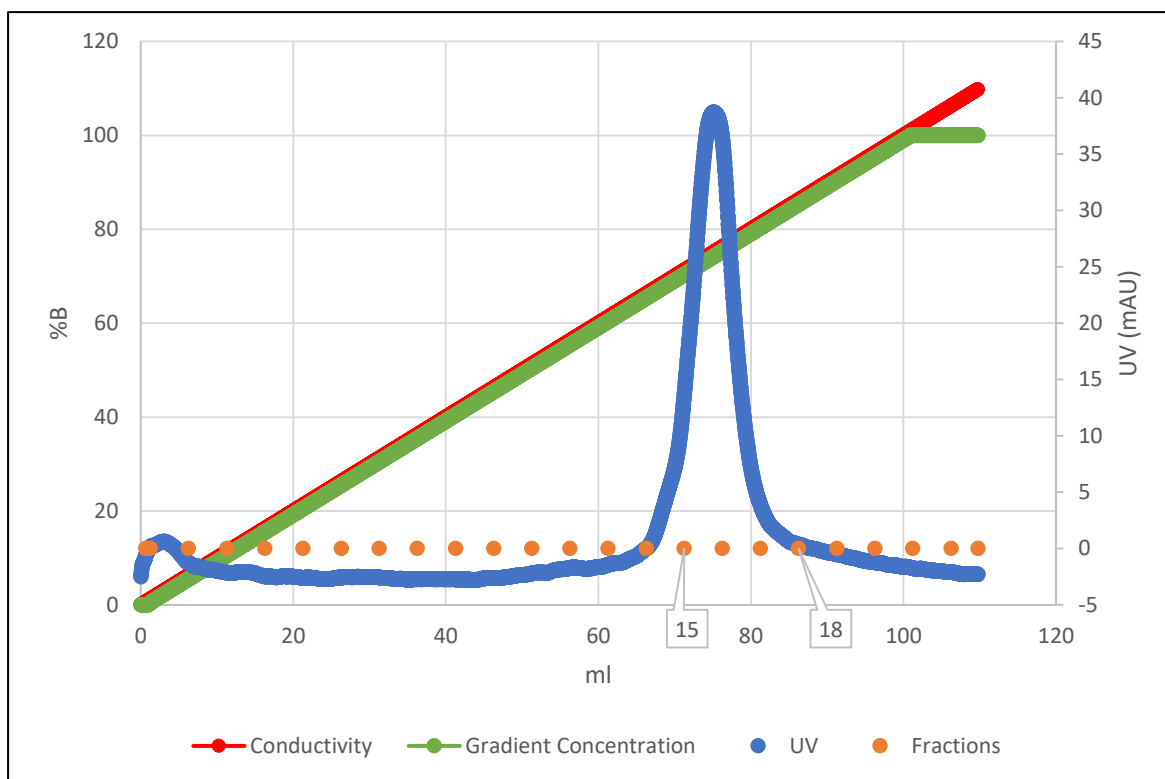


Figure 20:
TbrPDEB1 IEX Elution and Fractionation graph

In a similar method to the IMAC elution stage, the protein bound to the ion exchange column was eluted with an increasing salt concentration at a linear gradient, shown in green in *Figure 20*. The increasing salt displaces any bound protein as previously discussed. The gradient is seen to plateau at 100ml, having reached 100% salt concentration (1M NaCl). Additionally, the conductivity is also seen to increase at a linear gradient (as seen in red) due to being directly proportional to the salt concentration. A large UV peak is seen at 65ml, implying successful elution at a concentration of 650mM NaCl. Whilst this is high, it is expected as protein at pH = 8 is likely to bind tightly to the ion exchange column. The elution was collected in 5ml fractions. Fractions 15, 16, 17 and 18 were collected and sampled for an SDS-PAGE experiment, as seen on the following page.

3.2.4 IEX SDS-PAGE Gel

An interpreted image of the SDS gel following secondary Nickel and IEX purification can be seen in *Figure 21*, with columns annotated as follows (from left to right): ladder, IMAC elution, untagged protein (secondary Nickel flow through + wash through), Tagged protein (secondary Nickel elution), IEX flow through, Fraction 15, Fraction 16, Fraction 17, Fraction 18.

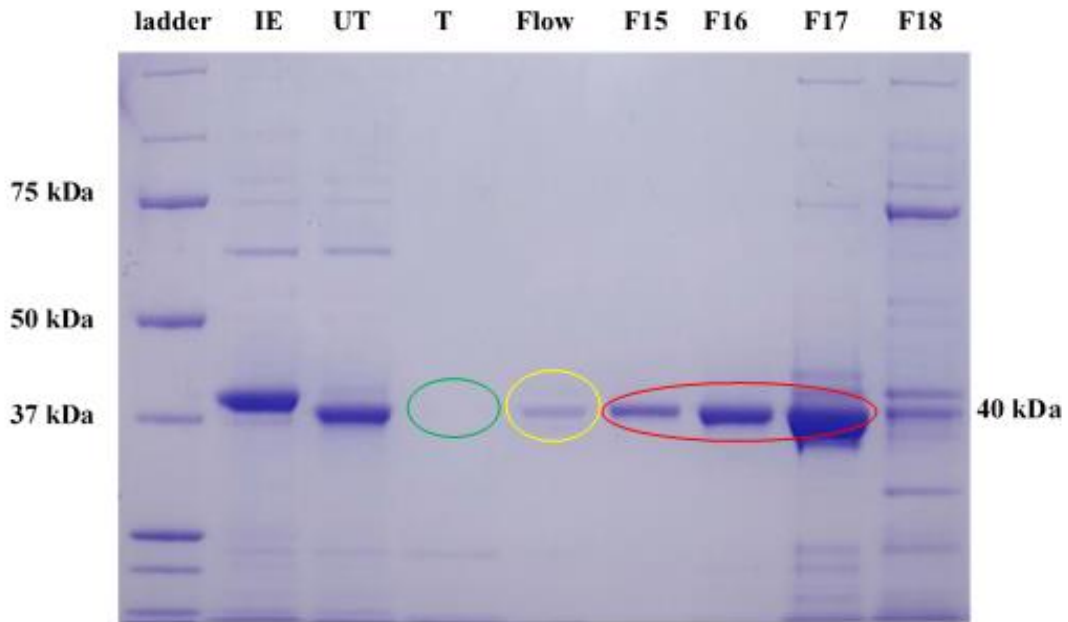


Figure 21:
TbrPDEB1 IEX SDS-PAGE Gel

As observed in the IMAC SDS-PAGE Gel, the 40kDa value indicating desired TbrPDEB1 protein persists throughout the next stages of purification (seen circled in red), implying the correct protein conditions have been met, as the protein was successfully isolated. This is confirmed by the lack of contamination peaks seen in bands F15 and F16. This also confirms the larger peak seen in the AKTA graph corresponding to desired protein. A lack of band circled in green suggest no TbrPDEB1 protein was present in the elution solution following secondary Nickel purification. This is due to no binding occurring between desired protein and the Nickel column as a result of cleavage of the histidine tag. In this case, the thrombin cleavage of the histidine tag is optimal. Despite some faint contamination bands being present in F17, the prominent 40kDa band indicates mainly TbrPDEB1 being present in the fraction, therefore the bands circled in red containing fractions F15, F16 and F17 were combined as the final purified protein sample. Sample F18 was discarded due to little desired protein present, and contaminants being present, especially at 75 kDa. The combined 15ml sample was concentrated and buffer exchanged accordingly, as previously discussed. The flow through is also seen to contain some

desired protein, as circled in yellow. Hence, the collected flow through was also concentrated down and buffer exchanged, further adding to the overall yield, which was recorded as follows:

A concentration of 0.71mg/ml after combining the appropriate fractions mentioned above, in 15ml volume. This results in:

$0.71 \times 15 = \underline{10.65\text{mg}}$ of purified TbrPDEB1 protein.

Additionally, 0.02mg/ml of TbrPDEB1 protein was found present within 98ml of flow through, resulting in an additional:

$0.02 \times 98 = 1.96\text{mg}$ of desired protein.

The flow through was extensively concentrated down and buffer exchanged and combined with the 10.65mg sample to yield:

$10.65 + 1.96 = \underline{\mathbf{12.61\text{mg}}}$ of purified TbrPDEB1 protein with an A260/A280 ratio of 0.6.

This purified protein yielded from a 2L culture was aliquoted appropriately as previously discussed and used for a number of screening experiments against the established fragment library, as will be presented in greater detail in the following chapters of this report.

3.3 Quality Control (QC)

As previously mentioned, the fragment library put together by Dr. Fredrik Svensson from IOTA Pharmaceuticals consisting of 31 compounds (purchased from Enamine) first underwent quality control by 1D Proton (^1H) NMR spectroscopy in order to ensure and verify optimal purity before undergoing experimentation with the purified TbrPDEB1 protein. A list of this fragment set can be seen tabulated in *Appendix 1* in the Appendices section at the end of the report. Hence, the 1D Proton NMR spectra and tabulated annotations can be seen for Ligands 6, 29 and 27. Similar information for the rest of the fragment library can also be found in the Appendices.

3.3.1 Ligand 6 QC

Figure 22 below shows the proton NMR spectrum for Ligand 6:

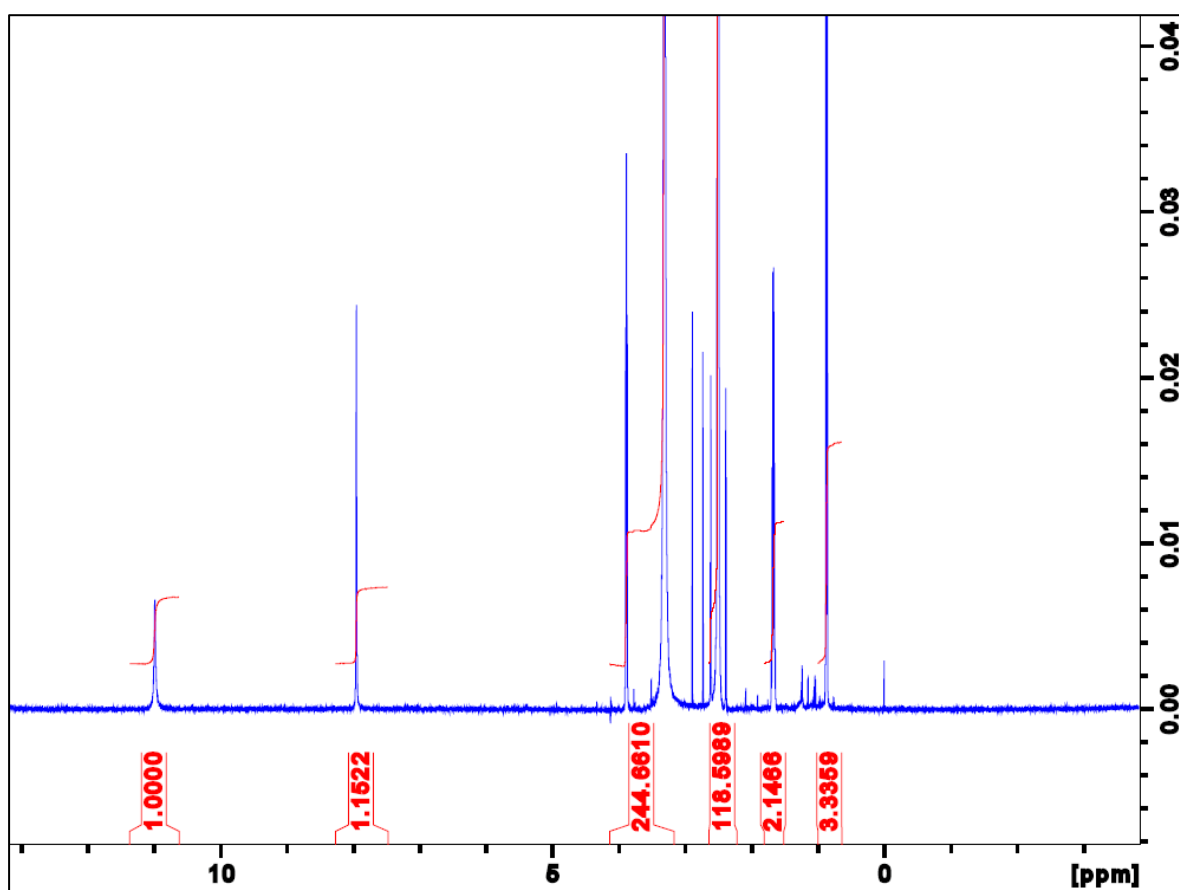


Figure 22:
Ligand 6 Proton NMR QC Spectrum

Additionally, an annotated table can be seen in *Table 4* with assigned peaks seen in *Figure 23* on the following page.

Table 4:
Ligand 6 Proton NMR annotated table

Peak assignment	Chemical shift (ppm)	Splitting	Corresponds to
/	0	Singlet	TMS
B	0.9	Triplet	CH ₃
C	1.75	Multiplet	CH ₂
/	2.4	Triplet	DMSOD ₆
/	3.4	Singlet	Solvent peak
F	3.9	Triplet	CH ₂
G	8.0	Singlet	CH
H	11.0	Singlet	NH

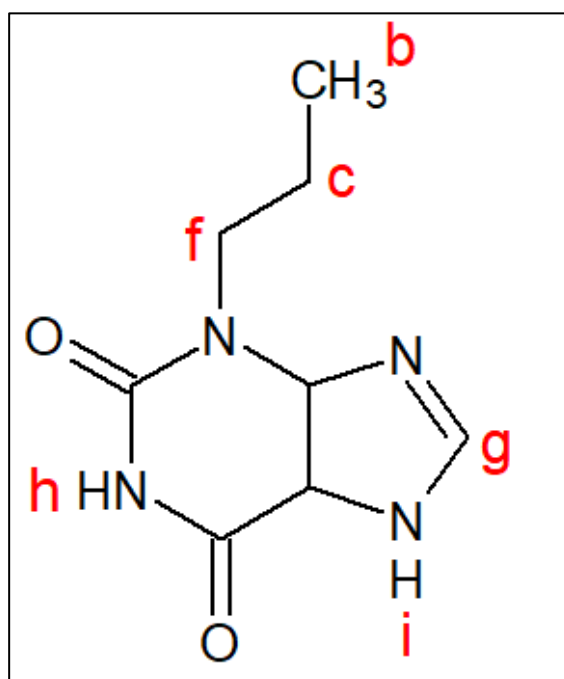


Figure 23:
Ligand 6 assigned peaks

3.3.2 Ligand 29 QC

Similarly, for Ligand 29:

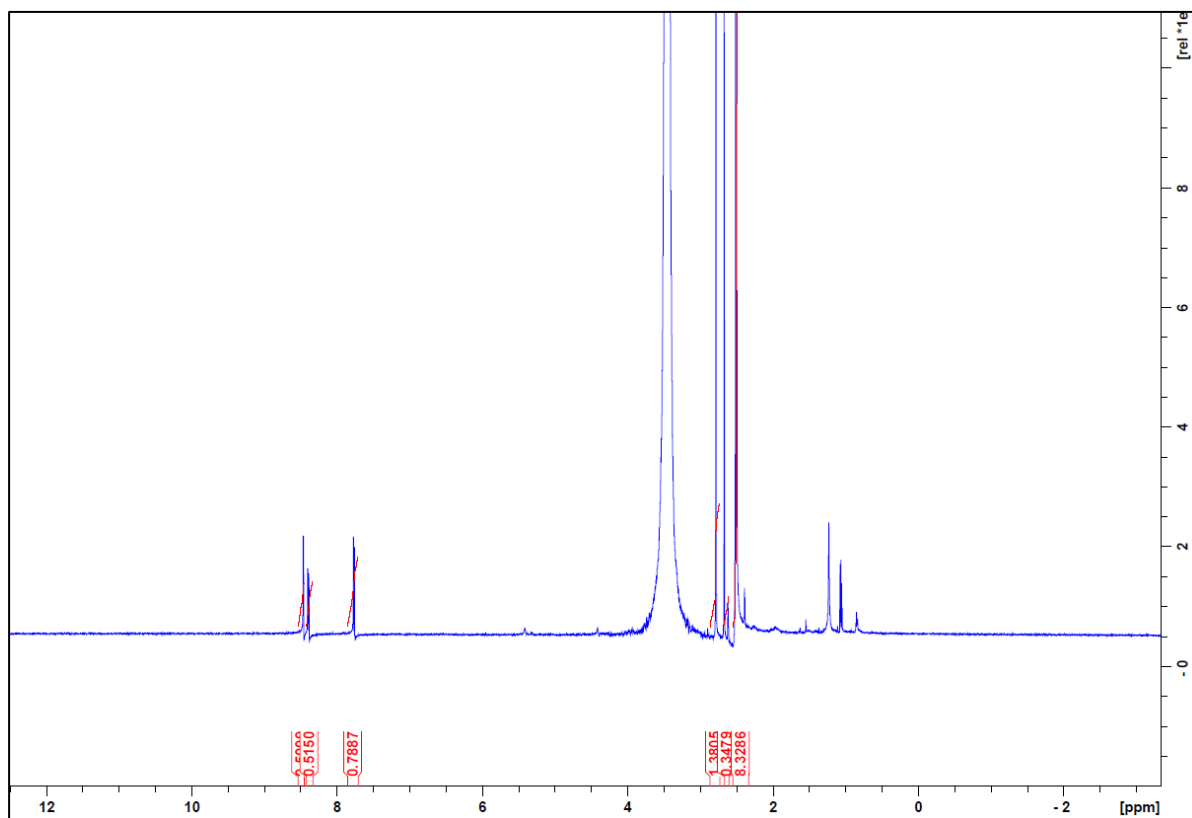


Figure 24:
Ligand 29 Proton NMR QC Spectrum

Table 5:
Ligand 29 Proton NMR annotated table

Peak assignment	Chemical shift (ppm)	Splitting	Corresponds to
/	2.4	Singlet	DMSO ₆
/	1.0	Singlets	Impurity
C	3.4	Singlet	Solvent peak
D	2.60	Singlet	CH ₃
E	2.79	Singlet	CH ₃
F	7.95	Doublet	CH
G	8.4	Doublet	CH
H	8.45	Singlet	CH

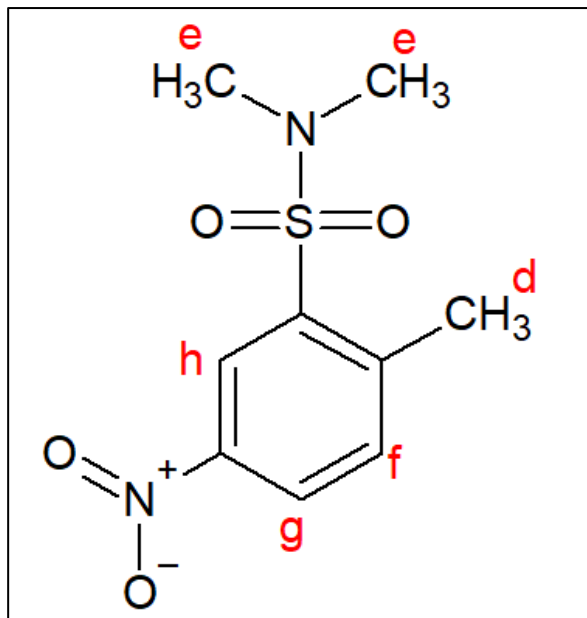


Figure 25:
Ligand 29 assigned peaks

3.3.3 Ligand 27 QC

Finally, for Ligand 27:

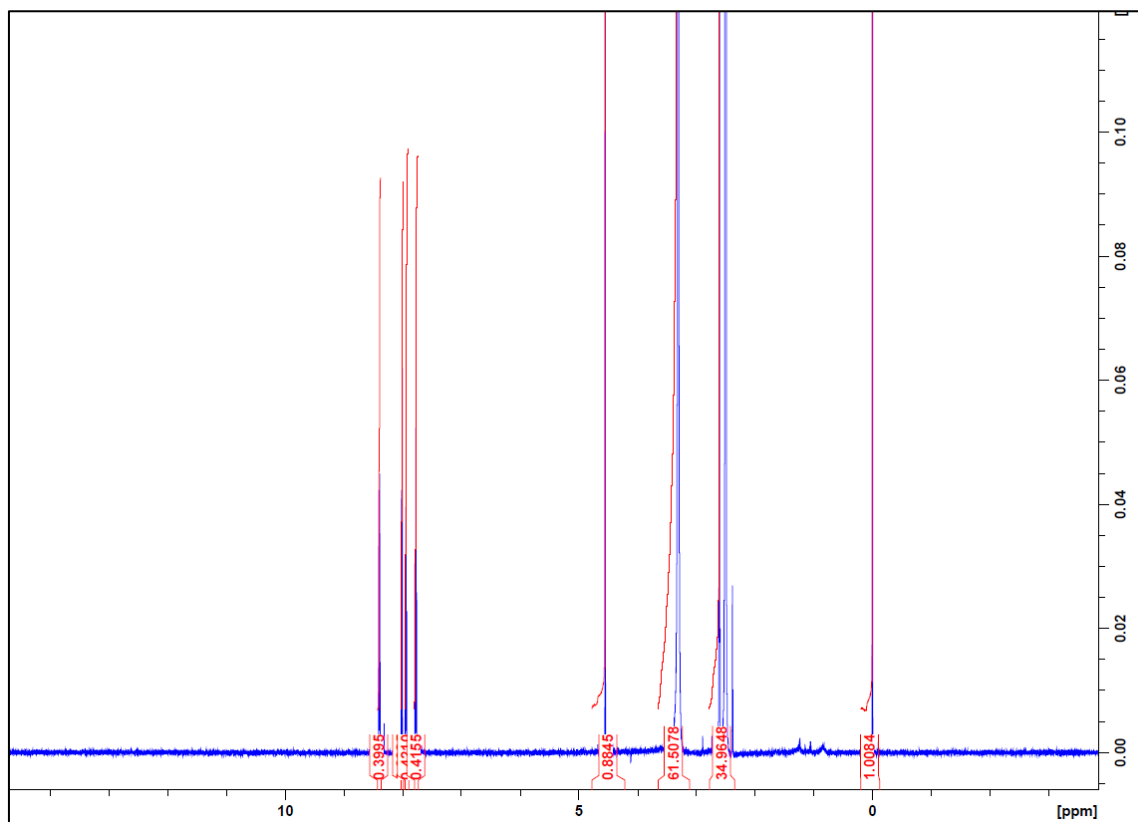


Figure 26:
Ligand 27 Proton NMR QC Spectrum

Additionally, the manually acquired QC spectra agree with those provided by Enamine for each compound (unless specified with a corresponding impurity or solvent peak in the annotated tables), as can be seen in the case of Ligand 27 by comparing *Figure 26* and *Figure 27*.

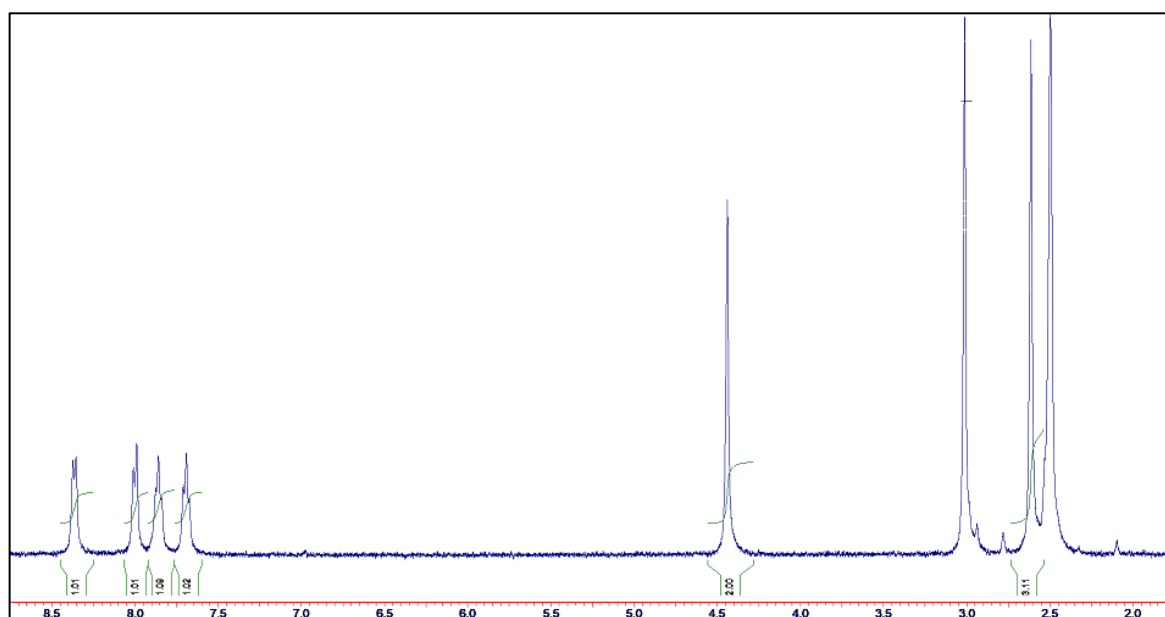


Figure 27:
Ligand 27 Proton NMR QC Spectrum provided by Enamine

Table 6:
Ligand 27 Proton NMR annotated table

Peak assignment	Chemical shift (ppm)	Splitting	Corresponds to
/	0	Singlet	TMS
/	2.5	Singlet	DMSOD ₆
C	2.6	Singlet	CH ₃
D	3.4	Singlet	Solvent peak
E	4.5	Singlet	CH ₂
F	7.6	Triplet	CH
G	8.0	Triplet	CH
H	8.1	Doublet	CH
I	8.3	Doublet	CH

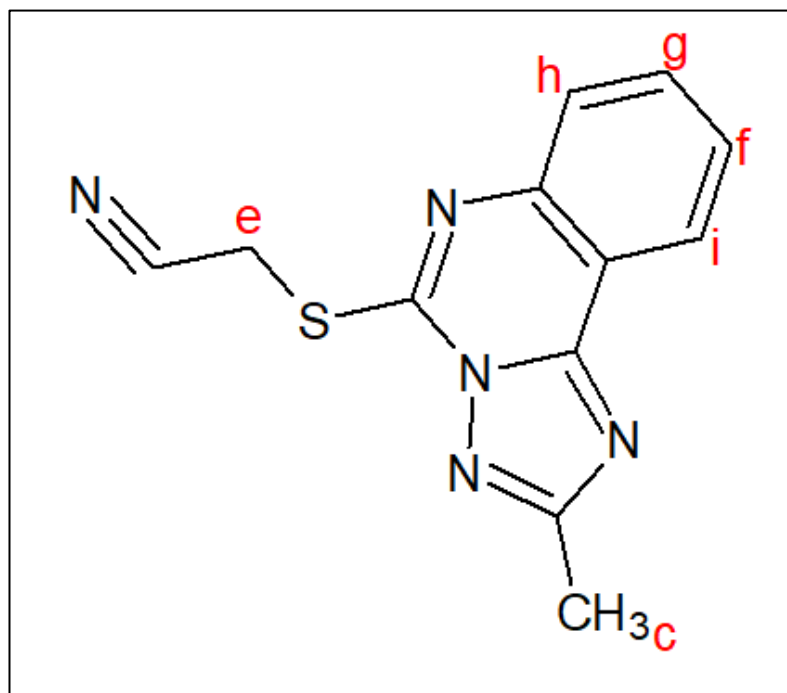


Figure 28:
Ligand 27 assigned peaks

As can be seen, a consistent solvent peak can be found in every Proton 1D spectrum at 3.4ppm. The DMSO₆ solvent in which all compounds were dissolved in as previously mentioned was also analysed in order to determine all solvent peaks still present in compound spectra. This can be seen in *Figure 29* below:

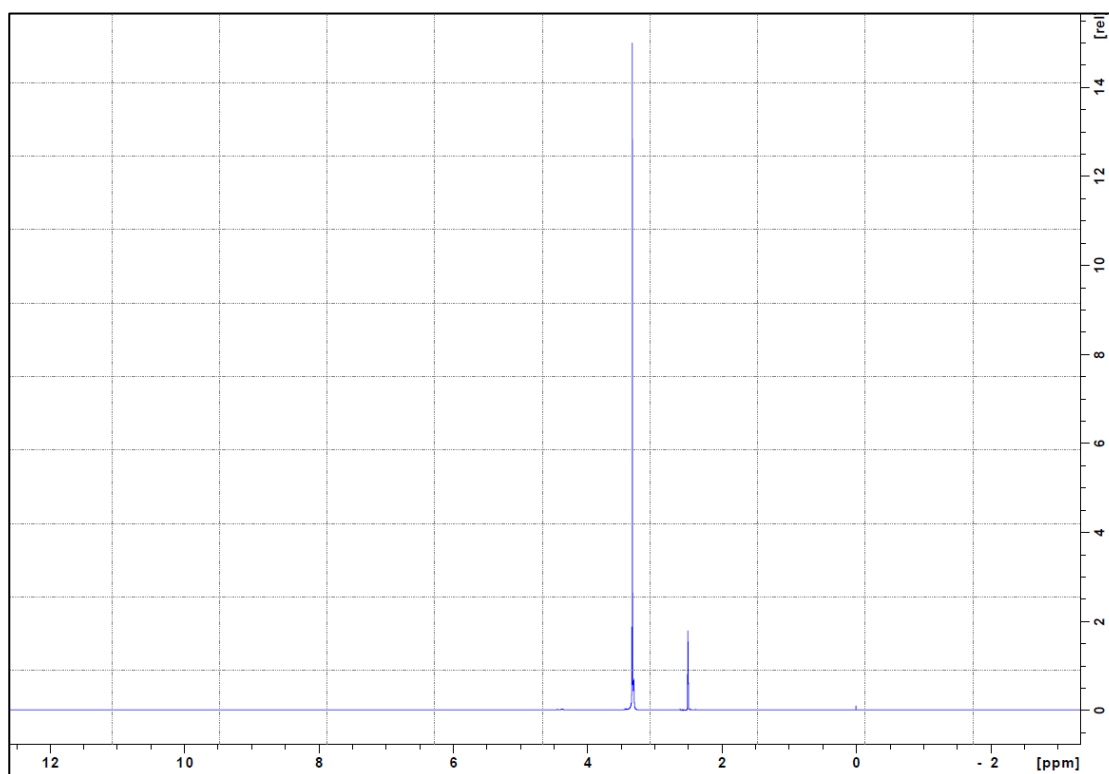


Figure 29:
DMSO₆ Solvent Proton NMR Spectrum

As can be seen, the solvent peak is present at 2.4ppm in all compound spectra presented above. Additionally, a large water peak is also seen at 3.4ppm – the same chemical shift as the consistent unknown peak. Additionally, Ligand 29 appears to be the only fragment in the set to contain any slight impurity, as shown in *Figure 24* and *Table 5* indicated by small singlet peaks at 1.0ppm. Ligand 29 was the only fragment to have undergone Proton 1D experimentation several months following the rest of the fragment set, so perhaps this impurity is due to degradation of the solution stock. Nevertheless, the purchased fragment set is of ideal purity appropriate for use in fragment-based screening experiments, as determined by comparing against the provided NMR spectra by Enamine, characterising and interpreting the peaks, and identifying any impurities or solvent/water peaks present. All peak assignments were also checked and verified using ACD/C+H NMR Predictor⁷², an NMR analysis software.

As previously mentioned, additional QC information can be found for the remaining ligands in the Appendices section (*Appendix 2-29*).

3.4 X-Ray Crystallography

Following purification, the yielded TbrPDEB1 protein was crystallised as previously discussed, producing robust crystals 20µm in length, as can be seen in *Figure 30* below. The protein crystals were used for X-Ray Crystallography experiments and were manually soaked with the library of 31 fragments used in this project (validated by QC NMR). Additionally, the XCHEM fragment screening approach previously mentioned was also done in parallel to manual crystal soaking in order to yield a wide variety of diffraction data used to assist in characterising, analysing and expanding the fragment library of known TbrPDEB1 binders used in this project by observing how each ligand interacts with the TbrPDEB1 protein at an atomic level, whilst also allowing comparison between the two methods of crystallisation.

Understanding how fragments are able to bind to the target protein is the first vital step in drug discovery by X-Ray Crystallography; Hence, before ligand evaluations can be made, it is first necessary to solve the unbound target protein apo structure.

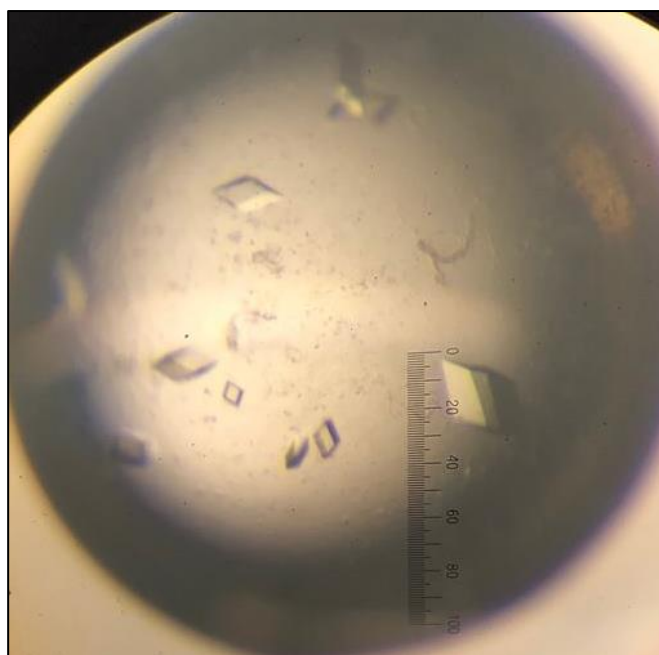


Figure 30:
Crystallised TbrPDEB1 protein – 20µm in length

3.4.1 Target Protein Apo Structure

In the case of TbrPDEB1 for this project, the protein crystal apo structure is already known and publicly available in the Protein Data Bank⁵⁴ (PDB, published by Wang, H and Ke, H), with the PDB code '4i15', and hence does not need to be solved. The PDB file displaying the 3D atomic structure in the case of the TbrPDEB1 catalytic domain protein structure, acquired from diffracting TbrPDEB1 crystals at 1.65Å resolution⁵⁴, can also be acquired from the PDB and used for 3D structure visualization using a number of different softwares as previously mentioned. The quoted total structure weight of the full-length protein is reported as 77.628kDa⁵⁴. The catalytic domain which hydrolyses cAMP to AMP as previously seen in *Figure 4* is present between residues 668-905 and contains 16 alpha-helices and 7 3_{10} -helices (types of secondary structures in polypeptides proteins). Note that the catalytic domain does not contain any beta-pleated sheets⁴⁹.

Hence, the published⁵⁴ 3D model is seen presented in *Figure 31*, created using PyMOL by Schrödinger – a molecular visualization software⁵⁵.

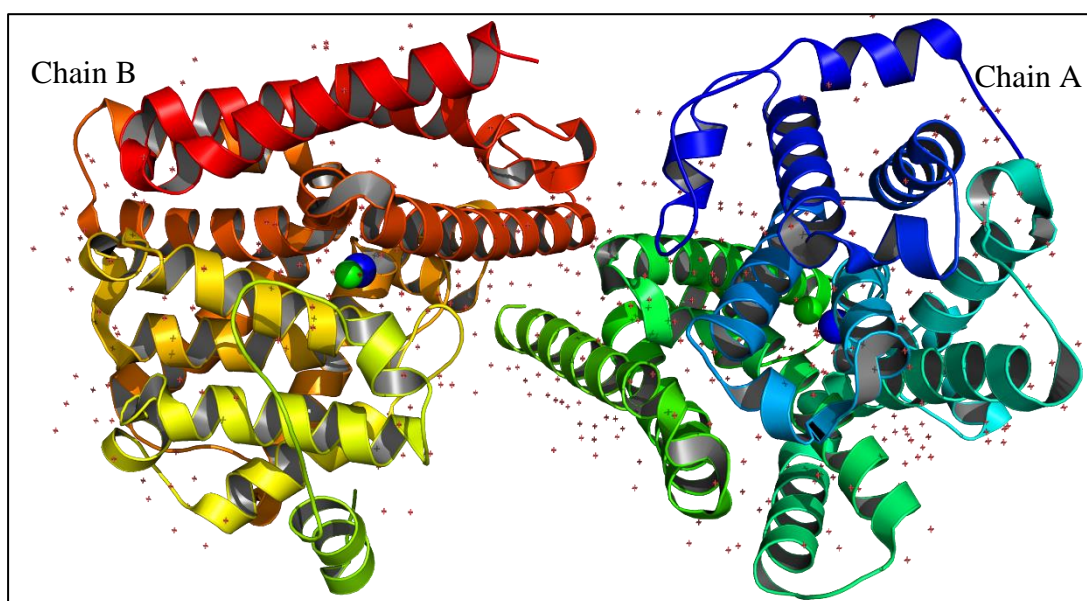


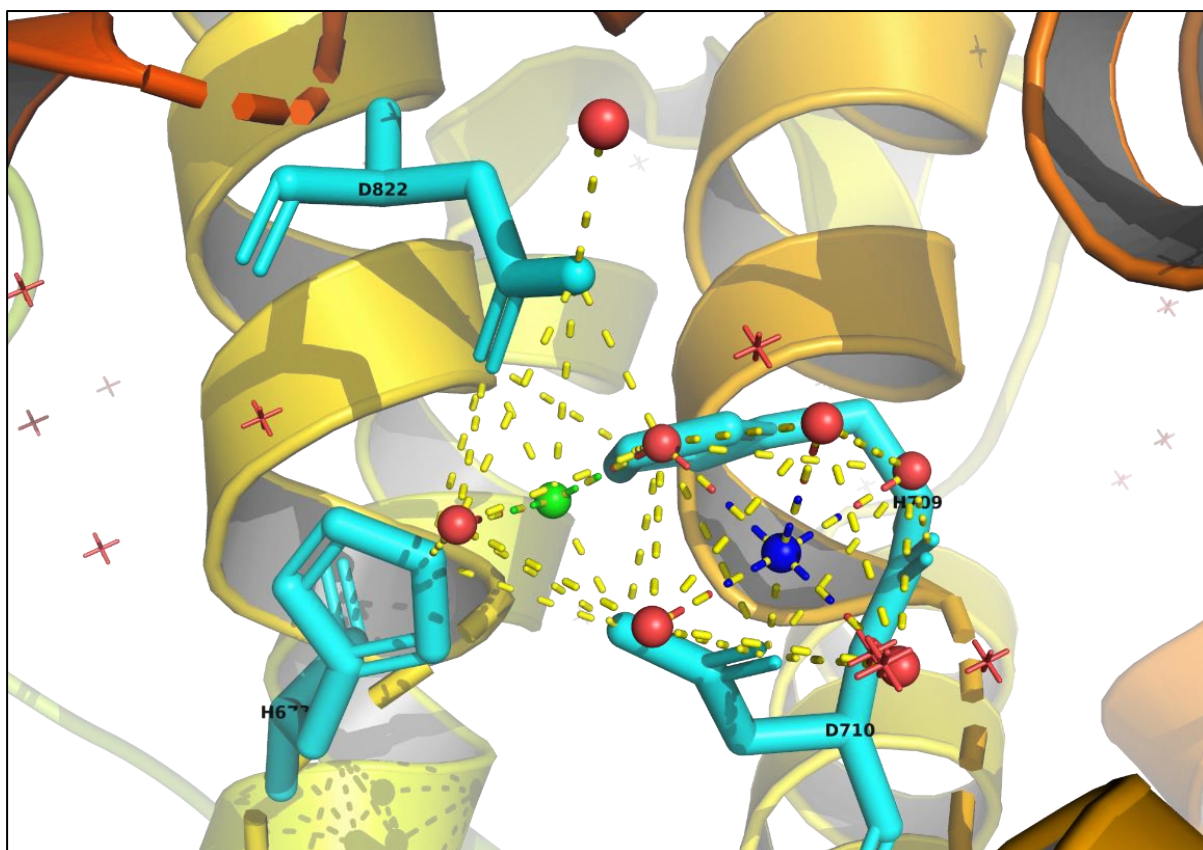
Figure 31:
TbrPDEB1 PDB Crystal structure overview – 1.65Å – Made with PyMOL

As can be seen, the asymmetric unit in *Figure 31* contains two TbrPDEB1 molecules labelled as Chain A (right) and Chain B (left) which form a dimer in the crystal lattice⁴⁹. Additionally, each chain is seen to contain two divalent metal ions – Zinc (shown as a green sphere) and Magnesium (shown as a blue sphere). Unbonded water molecules in the structure are presented as red dots.

Furthermore, specific regions of interest in the TbrPDEB1 crystal structure include:

1. Divalent Magnesium and Zinc metal ions
2. M-loop
3. Substrate binding pocket
4. Parasite specific 'P-pocket'

These regions of interest are presented in a number of images in the case of Chain B as follows:

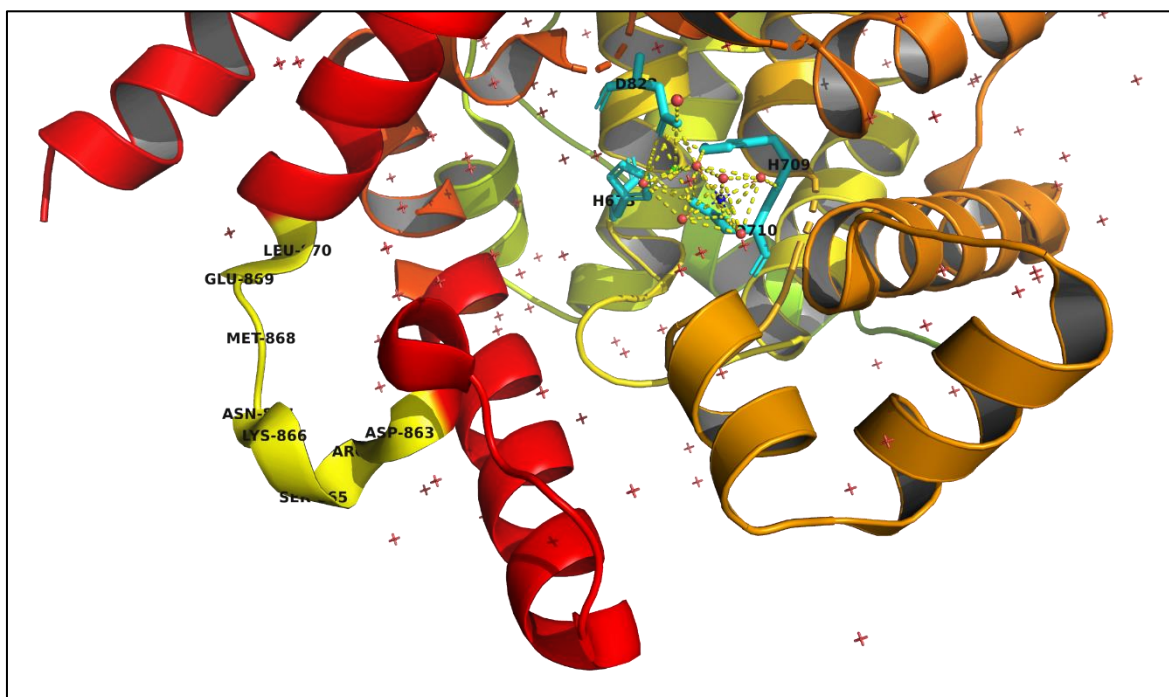


1 - Figure 32:
*TbrPDEB1 PDB Crystal structure – Divalent metal ions and water network –
Made with PyMOL*

The Magnesium and Zinc divalent metal ions present in the catalytic pocket of the TbrPDEB1 crystal structure serve an important role in enzyme biological function and activity. Metalloenzymes, in this case Zinc and Magnesium aid in catalysis by acting as Proton acceptors or Lewis acids. This aids in increasing polarity of chemical bonds and hence promote atom transfer⁵⁶, particularly in the case of Zinc due to having a stable d^{10} electron subshell – making it stable in the Zn^{2+} electrophilic state and readily able to form metal complexes⁵⁷. Being able to form water complexes, metal ions therefore also ‘free up’ the binding site for substrates^{57,58}. In a 2009 metalloenzyme study by S. Maric *et al*⁵⁹,

stripping a Malaria enzyme protein of Zinc by use of chelating agents such as EDTA causes a reported loss in enzyme activity by up to 57%. Upon addition of Zn^{2+} ions, an increase of enzymatic activity back to the original 100%⁵⁹ was reported. In addition to aiding in enzymatic catalytic activity, metal ions are also important nutrients; Magnesium is required for DNA replication and other biological processes and aids in protein folding and interactions⁶⁰.

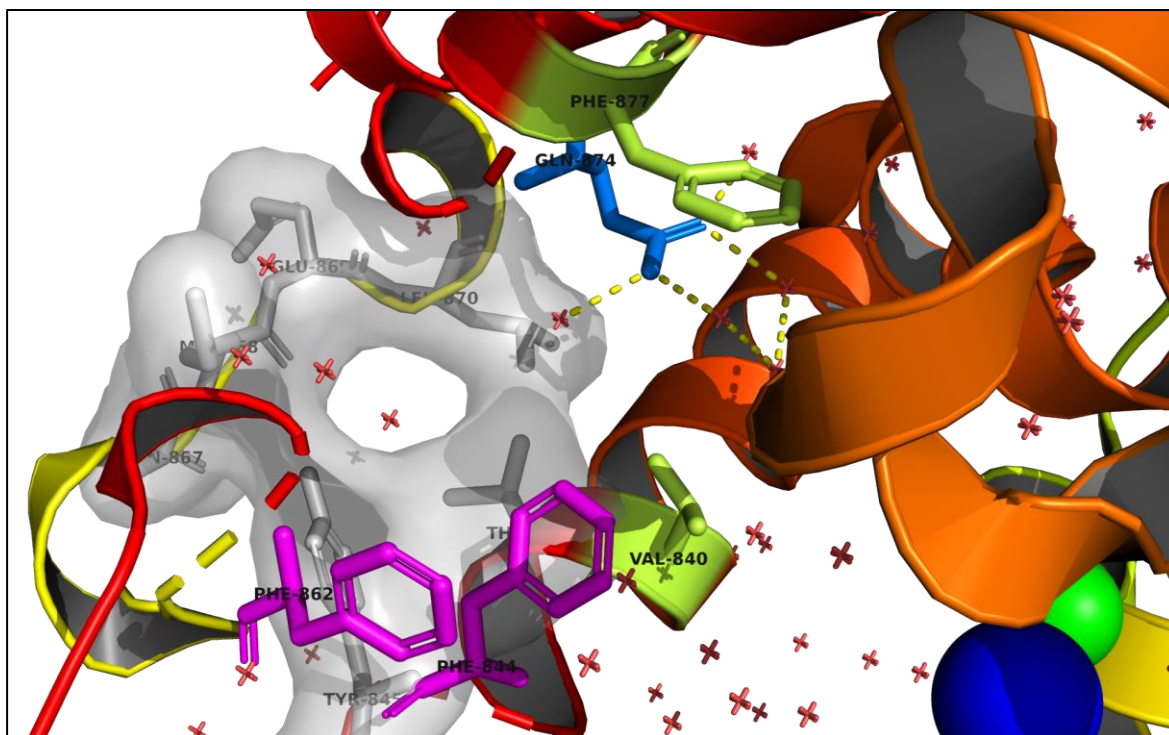
Looking at *Figure 32* on the previous page, the divalent metal ions both form octahedral complexes; Zinc (green sphere) is seen to coordinate 2 Histidine (H) residues at positions 673 and 709, 2 Aspartic acid (D) residues at positions 710 and 822 (shown in cyan) and 2 water molecules (red spheres). Magnesium (blue sphere) is seen to also coordinate with Aspartic acid 710 and 5 water molecules⁴⁹. These metal complexes are tightly bound and ensure enzymatic activity.



2 - *Figure 33:*

TbrPDEB1 PDB Crystal structure – Labelled M-loop – Made with PyMOL

The Aspartic acid, Arginine, Serine, Lysine, Asparagine, Methionine, Glutamic acid and Leucine amino acids ranging from residues 862-869 form the M-loop of the TbrPDEB1 crystal structure, as seen presented in yellow in *Figure 33*. The M-loop plays a part in the formation of the parasite-specific ‘P-pocket’ present in the crystal structure, which will be later discussed. Recent reports⁴⁹ claim that the length of the M-loop influences the appearance of the P-pocket in *Trypanosoma brucei*, *Trypanosoma cruzi* and also in *Leishmania major* parasitic Phosphodiesterases⁴⁹.



3 & 4 - Figure 34:
TbrPDEB1 PDB Crystal structure – Substrate binding pocket, Hydrophobic clamp, Glutamine + water network and Parasite specific ‘P-pocket’ – Made with PyMOL

Figure 34 shows the key points of interest in the TbrPDEB1 crystal structure. Firstly, the formation of M-loop (shown in yellow) is stabilized by the π - π stacking (caused by atomic orbital overlapping between atoms) non-covalent interactions between the benzene rings of Phenylalanine 844 and Phenylalanine 862, as shown in magenta⁴⁹. This is similarly seen upon addition of a conjugated ring system substrate within the binding pocket, as will be discussed shortly.

The binding pocket is present in-between the Phenylaniline 877 and Valine 840 residues, as shown in lime green in Figure 34. These moieties form a hydrophobic clamp, which is key for creating a preferential binding pocket. Due to the largely hydrophobic nature of phenylaniline and valine (according to the hydrophobic scales⁶¹), the two molecules have a tendency to attract to one other due to hydrophobic effect – explained by the attraction of nonpolar molecules as a result of entropically favourable conditions. Additionally, the presence of an aromatic substrate within the clamp further ensues non-covalent interactions as a result of π - π stacking between the aromatic substrate and the PHE-844, and possibly PHE-877 residues⁶². This is the first key principle of ensuring ligand potency and binding to the protein structure, hence all ligands purchased and used for crystallography

experiments are seen to have an aromatic scaffold as they are considered ‘known PDE binders’.

Another essential key principle for increasing ligand potency and binding strength is due to the presence of the Glutamine 874 ‘Q-switch’ residue within the binding pocket, shown in navy blue in *Figure 34*. A water network is also present around this glutamine residue, as shown. Upon the addition of a substrate with a polar side chain, the glutamine is able to interact with the added ligand by hydrogen bonding, further increasing the binding strength of the compound within the binding pocket. The water molecules forming this additional network are displaced upon ligand uptake thus to allow this hydrogen bonding to occur⁴⁹. Additionally, polar side chains being able to form multiple hydrogen bonds to this glutamine residue would increasingly strengthen protein-ligand interactions. The main point of interest presented in *Figure 34* is the Parasite-specific ‘P-pocket’ region shown in grey between the M-loop and Glutamine 874, formed by Alanine 837, Threonine 841, Tyrosine 845, Asparagine 867, Methionine 868, Glutamic acid 869 and Leucine 870 residues³³. The existence of this pocket may be exploited to give rise to phosphodiesterase inhibition specificity when targeting the parasitic enzyme. As previously discussed, due to the similarity in the binding pockets of human PDEs⁶³ and parasitic PDEs, inhibition would be non-specific and would therefore be non-ideal and anti-therapeutic. *Figure 35*, adapted from a similar 2013 novel study on PDE crystal structures by Jansen *et al*⁴⁹ depicts the differences between this pocket region in parasitic PDEs and hPDE4B (human phosphodiesterase 4B).

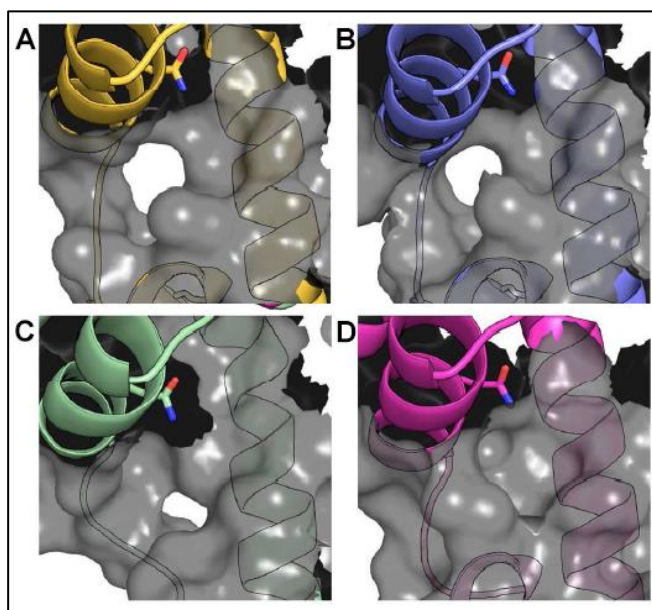


Figure 35:
P-pocket region differences in TbrPDEB1 (A), LmjPDEB1 (B), TcrPDEC (C) and hPDE4B (D), adapted from a study by Jansen et al⁴⁹

As can be seen, this pocket is seemingly only present in parasitic PDE structures, shown by the open density in *Figure 35A-C*, which is not present in the case of human PDE4B as seen in *Figure 35D*. Therefore, optimal compounds resulting in specific parasite PDE inhibition should contain an aromatic group and polar side chain scaffold in order to obtain ideal binding to the binding pocket, as well as large side chains capable of penetrating the empty density present in the P-pocket. In similar studies, such as the review by K. Orrling *et al* on Catechol Pyrazolinones as Trypanocidals⁶² and the Fragment-Based Drug Discovery of PDEs article by F. Svensson *et al*²², known human PDEs such as analogues of Rolipram and Sildenafil (otherwise known as Viagra) were used as starting scaffolds for parasitic PDE inhibitors. Slight variation in inhibitor chemotype was derived by methods such as virtual screening and molecular docking³³ of suggested derivatives⁴⁹. These were the main methods used to identify and evaluate parasite-specific inhibitors. Being the main structural difference between parasitic and human PDEs, this P-pocket cavity has been the main target of focus in this particular field of research^{33,49,62,22}.

As previously discussed in chapter 1.3, NPD-008³³ was found to be a highly potent and partially selective TbrPDEB1 inhibitor. *Figure 36* below, adapted from the mentioned 2018 study by Blaazer *et al*³³, shows the crystal structure of TbrPDEB1 with NPD-008 present in the binding pocket at a resolution of 1.8Å.

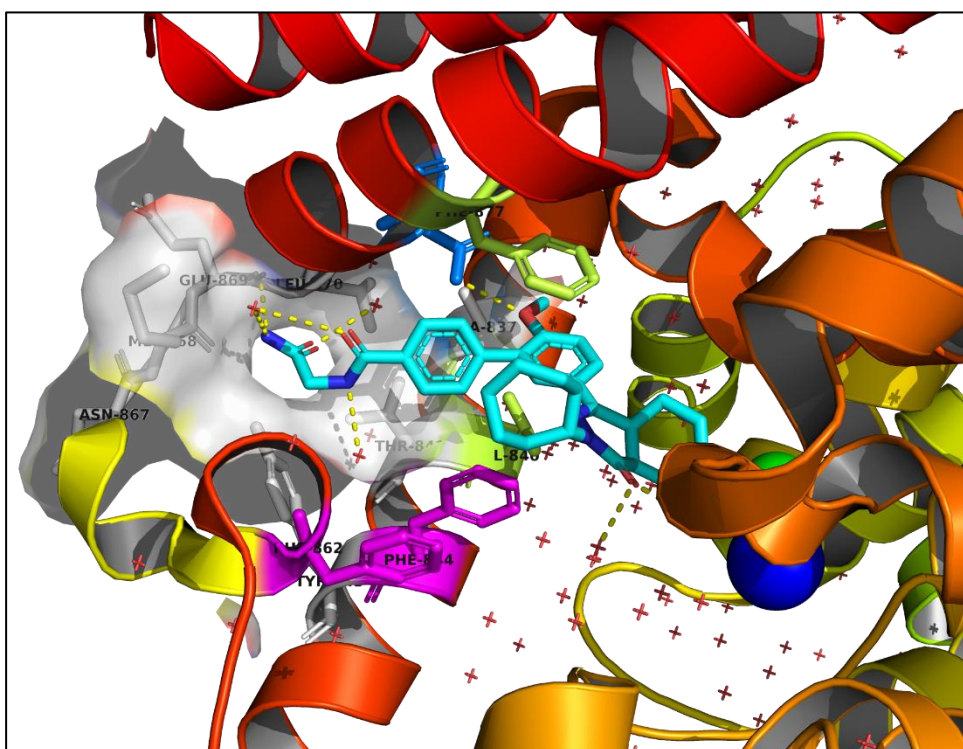


Figure 36:
NPD-008 bound to TbrPDEB1 protein structure, made with PyMOL using the 5G2B PDB Structure code³³

This inhibitor's potency is immediately apparent, as a result of multiple hydrogen bonds formed between the bound ligand, Glutamine 874, and other polar residues, as well as the previously discussed π - π stacking interactions in the substrate binding pocket.

Additionally, the glycinamide tail of NPD-008 is also seen to form a hydrogen bond with the Glutamic acid 869 residue present in the P-pocket, successfully targeting the desired cavity and also achieving the desired therapeutic effect as previously shown in *Figure 7* and *Figure 8*. However, only a narrow inhibition selectivity was reported in TbrPDEB1 over hPDE4, with clear cytotoxic effects of NPD-008 still being present against human MRC-5 cells³³ (Medical Research Council cell strain 5).

Hence, the fragment library used in this project will similarly be evaluated in terms of potency in order to identify any potential novel fragment binders to TbrPDEB1.

3.4.2 Inhibitor Confirmation by X-Ray Crystallography

Table 7 on the following page shows an overview of the ligand hits by X-Ray crystallography. As previously mentioned, varying ligand soak solution concentrations and soak durations were used throughout these experiments thus to have a wide variety of hit data and be able to best determine ideal binders by X-Ray Crystallography. Additional information is present in *Appendix 30* for manual soak experiments and *Appendix 31* for the XCHEM approach experiments.

Note that the XCHEM crystallography experiments were only attempted once, and the entire library screen was not complete. However as can be seen in *Table 7*, double confirmations in the case of Ligands 1, 4 and 12 show promise in those particular structures as binders by X-Ray crystallography. The damaging of the crystal lattice or a poor resolution map result was very common in the case of crystallisation using the XCHEM approach; Reasons for this will be discussed later in this chapter. As previously mentioned, the major advantage of this approach is to screen large fragment libraries quickly and efficiently; Due to the small size of the fragment library used in this project, it is evident that the manual soaking technique – despite being more time consuming – overall served a better purpose for this particular task as it yielded more hits that were not apparent following the XCHEM approach. Nevertheless, the ability to compare and overlap the two techniques is undoubtedly beneficial in the case of this research project.

Table 7:
Ligand hits by X-Ray Crystallography overview

Ligand #	Manual Soak Hit	XCHEM Hit	Ligand #	Manual Soak Hit	XCHEM Hit
1	✓	✓	17	✓	/
2	✗	/	18	✗	/
3	✓	/	19	✗	✗
4	✓	✓	20	✗	/
5	/	✗	21	✗	/
6	✓	/	22	✓	/
7	✗	/	23	✓	/
8	✗	/	24	✗	/
9	✓	✗	25	✗	✗
10	✗	/	26	✗	/
11	✗	/	27	✓	/
12	✓	✓	28	✗	/
13	/	/	29	✓	/
14	/	/	30	✗	✗
15	✗	✗	31	/	/
16	✗	/			

Key:

✓ = Binder by X-Ray crystallography

✗ = Non-binder by X-Ray crystallography

/ = Crystal damaged

Hence, out of the 11 X-Ray Crystallography ligand hits, a number of successful binders are seen presented and discussed in greater detail in the following subheadings. The remaining hits are presented in *Appendix 32-39*.

3.4.2.1 Ligand 29 Data Analysis

A TbrPDEB1 crystal soaked with Ligand 29 (ID Z45527541) at 40mM concentration for 48 hours is seen as collected in a crystal loop prior to diffraction in *Figure 37*:



Figure 37:
TbrPDEB1 crystal soaked with Ligand 29 snapshot

This crystal was diffracted at a resolution of 1.52Å and can be seen as collected on a Pilatus 6m image at the Diamond Light Source in *Figure 38*:

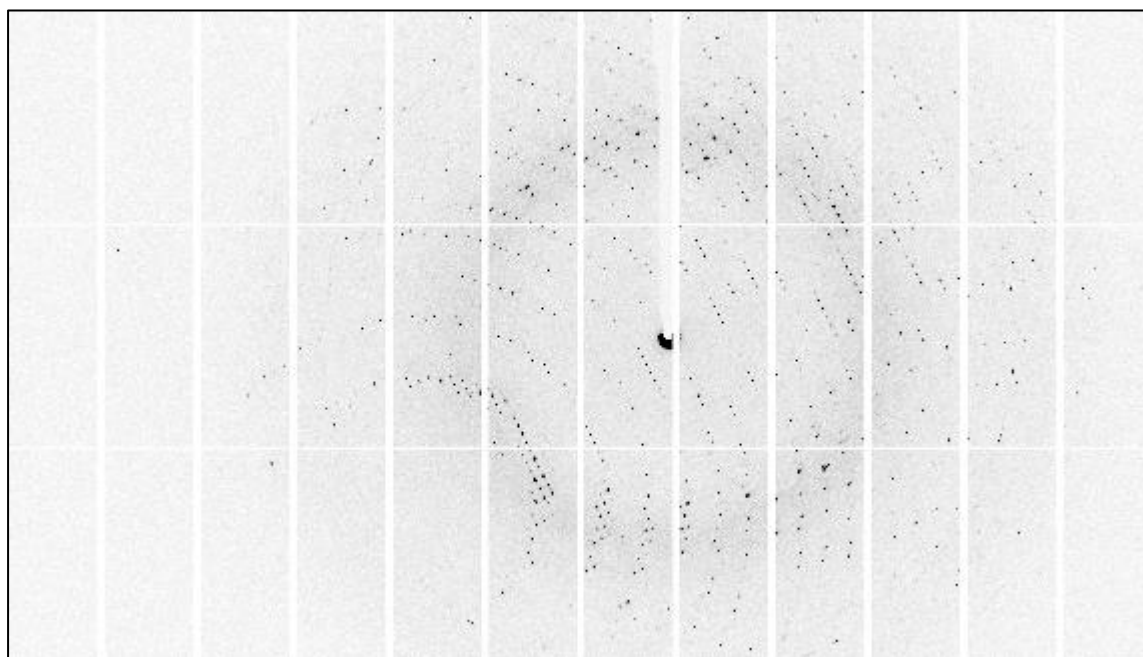


Figure 38:
TbrPDEB1 crystal soaked with Ligand 29 1.52Å diffraction

Additionally, crystal details are further presented in *Table 8* on the following page:

Table 8:
X-Ray data collection and refinement statistics for *TbrPDEB1-Lig29* complex

Data collection -	
Synchrotron and Beamline	Diamond Light Source, io4-1
Space group	C 1 2 1
Molecules	2
Cell dimensions	
a, b, c (Å)	115.70, 115.30, 68.60
α , β , γ (°)	90.00, 108.30, 90.00
Resolution (Å)	1.46 – 65.04
R_{merge}	0.039
$I/\sigma I$	17.1
CC (1/2)	0.999
Completeness (%)	98.86%
Multiplicity	3.3
Refinement -	
Resolution (Å)	79.45 – 1.46
No. reflections	7532
$R_{\text{work}} / R_{\text{free}}$	0.189 / 0.217
No. atoms	5130
R.m.s deviations	
Bond lengths (Å)	0.0186
Bond angles (°)	1.782

The crystallographic data collected for Ligand 29 was processed through the autoPROC processing pipeline at the Diamond Light Source, giving a reflection ‘mtz’ file. This reflection file was used to overlay the 4I15 PDB crystal structure for TbrPDEB1, and the map was solved accordingly. A green electron density ($F_o - F_c$ omit difference map) in the preferential binding pocket of the crystal structure (present between Phenylalanine 877 and Valine 840, as previously discussed) implies a structural difference in the collected diffraction data in comparison to the model used; In other words, successful ligand binding is indicated as can be seen in *Figure 39* at a 0.9 contour level, analysed with and created by using the CCP4i2 and COOT Crystallography and 3D Visualisation softwares previously mentioned.

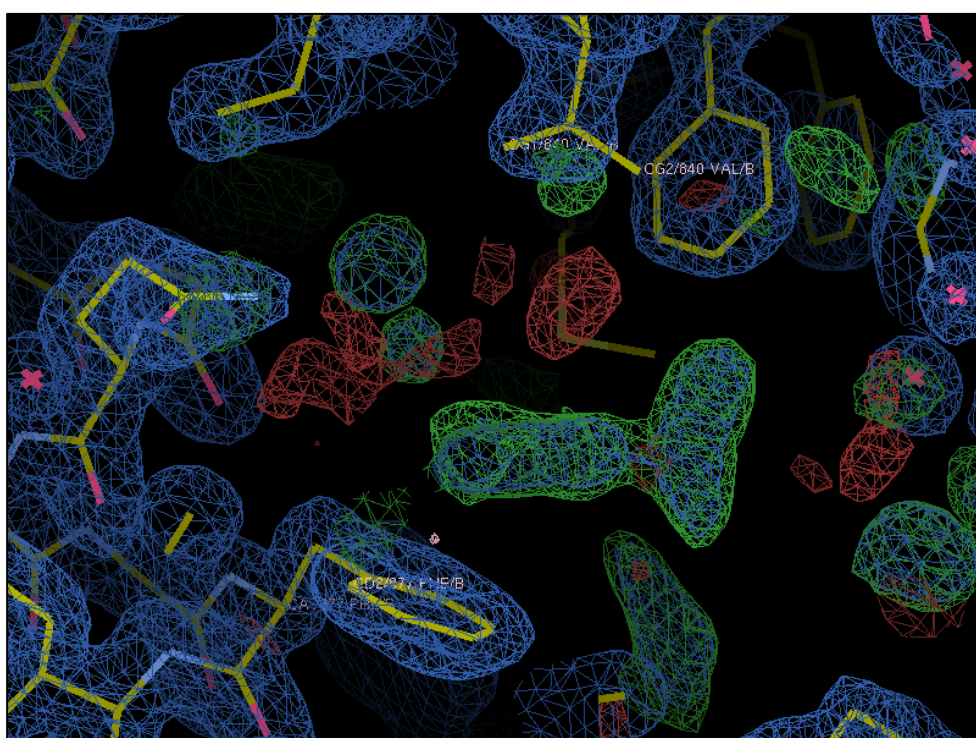


Figure 39:
TbrPDEB1-Lig29 successful ligand binding, indicated by green electron density

Following initial structural refinement, the ligand was modelled into the binding site using the appropriate smiles string shown in *Appendix 1*. Further refinement of the structure gave the final 3D interpretation of this protein-ligand complex as seen in *Figure 40* on the following page. Note that structural information for Ligands 6, 27 and 29 can be seen in *Appendix 1* or in *Figures 23, 28 and 25*, respectively.

Figure 40 depicts the binding mode of Ligand 29 in the TbrPDEB1 crystal structure binding pocket. As can be seen, the electrophilic sulphonamide tail group forms a surrounding water network by hydrogen bond formation. Surprisingly, the nitro group does not form a hydrogen bond with the invariant Glutamine 874. The likelihood of this is due to the ligand having a conformation pointing away from this glutamine residue; The distance between the polar oxygen on the mentioned nitro group and the Glutamine residue is measured (using COOT) to be 5.74Å in length – too large for the typical 1.5-2.5 angstrom range of hydrogen bonds.

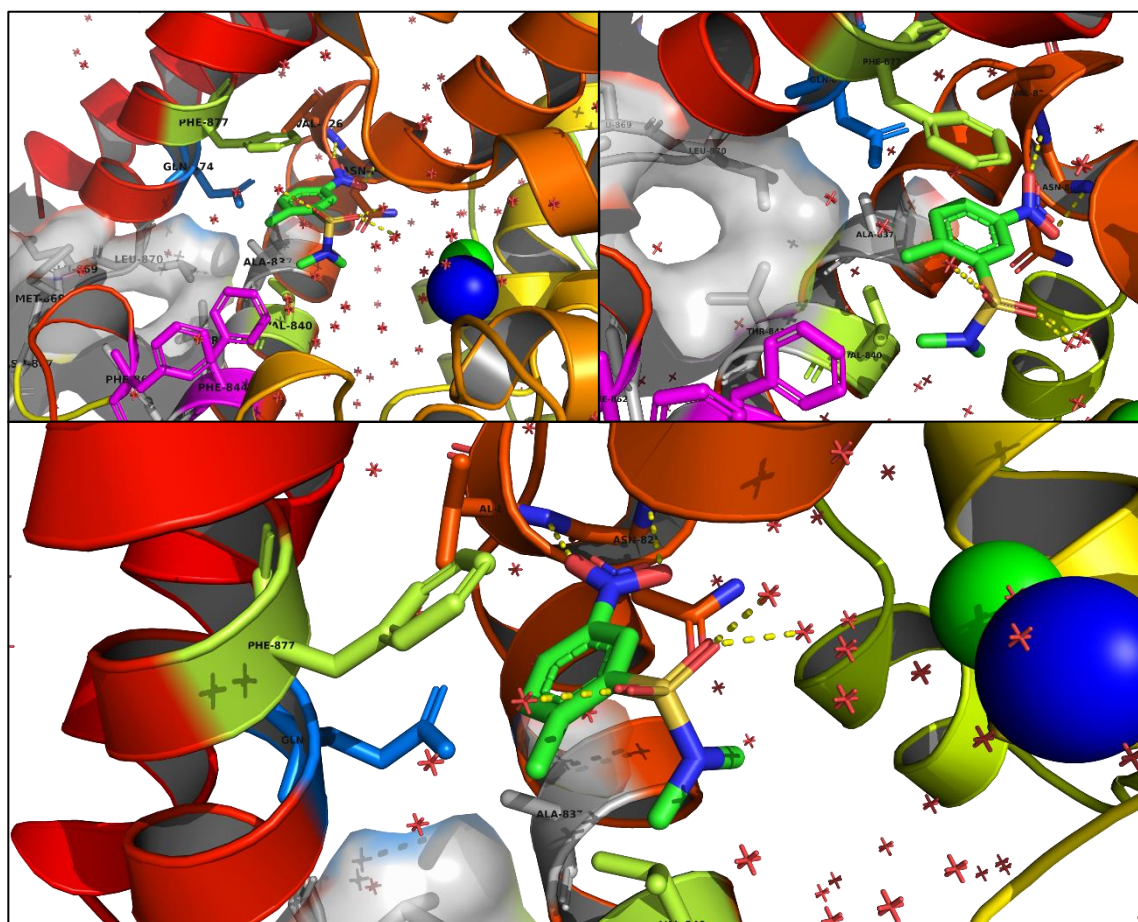


Figure 40:

TbrPDEB1-Lig29 3D structure and binding from 3 different angles, made with PyMOL

Instead of Q-switch hydrogen bond formation, the ligand nitro oxygens form hydrogen bonds with the nitrogen atoms present on the Asparagine 825 and Valine 826 residues, 3.78Å and 3.10Å in length, respectively. This hydrogen bond formation and water networks present maintain the ligand in the preferential binding pocket of the crystal structure. Whilst this is promising despite the lack of Glutamine 874 interaction, further protein-ligand binding determination experiments are required to validate Ligand 29 as a potential TbrPDEB1 binder.

3.4.2.2 Ligand 27 Data Analysis

Similarly, Ligand 27 (ID Z220370518) was a successful hit in a soak concentration of 50mM and a soak duration of 72 hours. The crystal details are seen presented in *Table 9* below:

Table 9:
X-Ray data collection and refinement statistics for TbrPDEB1-Lig27 complex

Data collection -	
Synchrotron and Beamline	Diamond Light Source, io4-1
Space group	C 1 2 1
Molecules	2
Cell dimensions a, b, c (Å) α , β , γ (°)	115.40, 114.77, 68.29 90.00, 108.35, 90.00
Resolution (Å) R_{merge} $I/\sigma I$ CC (1/2) Completeness (%) Multiplicity	1.48 – 57.39 0.044 11.6 0.999 89.50% 3.0
Refinement -	
Resolution (Å)	79.24 – 1.47
No. reflections	6300
$R_{\text{work}} / R_{\text{free}}$	0.244 / 0.270
No. atoms	5130
R.m.s deviations Bond lengths (Å) Bond angles (°)	0.0163 1.658

Additionally, a 3D model interpretation of Ligand 27 binding to TbrPDEB1 is seen:

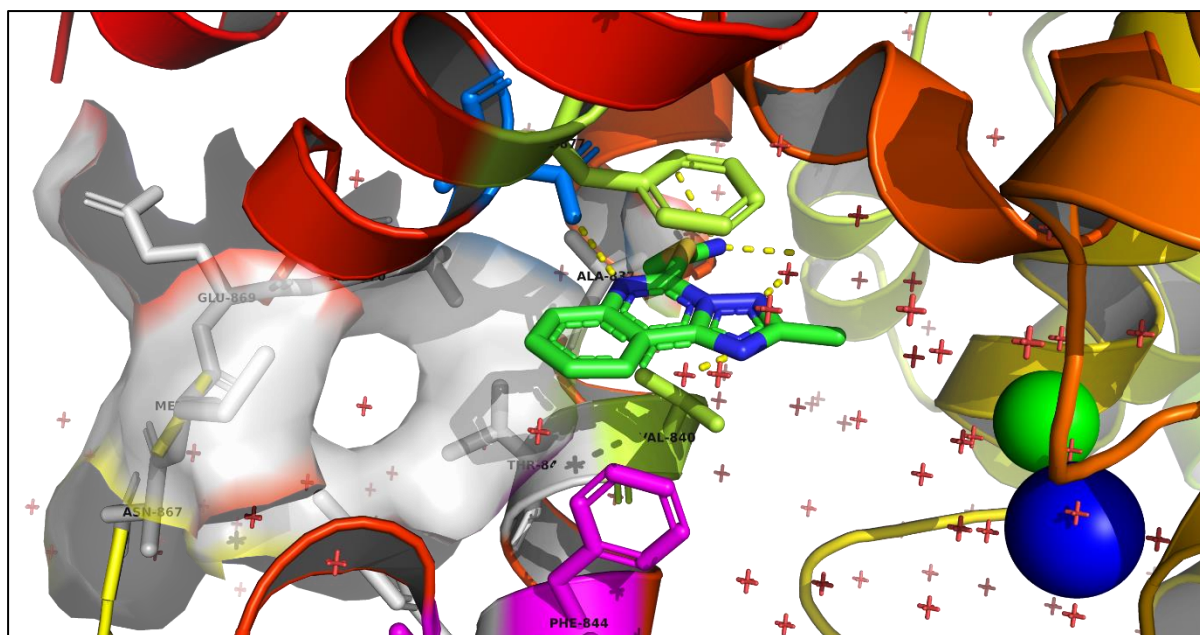


Figure 41:
TbrPDEB1-Lig27 3D structure and binding made with PyMOL

As can be seen, Ligand 27 is well established within the TbrPDEB1 binding pocket. Like Ligand 29 in *Figure 40*, Ligand 27 forms a hydrogen bond water network. However, unlike the latter, Ligand 27 forms the essential hydrogen bond to the invariant Glutamine 874 residue as previously discussed. The hydrogen bond, 3.20Å in length, is likely what gives rise to Ligand 27's potency and ability to bind to the binding pocket. Despite the successful bond formation, the acetonitrile tail of Ligand 27 points away from the parasite specific P-pocket; Therefore, if binding of this ligand to the protein target is to be confirmed by further means, it is highly expected that binding will be non-specific to the TbrPDEB1 target and inhibition to hPDE4D would also be observed. Of course, specific inhibition is not expected in the case of any fragments used in this project, as the purpose of these compounds was merely to determine their binding to TbrPDEB1, as they were originally selected by their tendency to inhibit hPDE enzymes. In any case, binding confirmation by other techniques is required in order to establish an active chemotype optimal for target binding, and to observe for target inhibition.

3.4.2.3 Ligand 6 Data Analysis

Finally, Ligand 6 (ID Z1945707486) is presented as follows:

Table 10:
X-Ray data collection and refinement statistics for TbrPDEB1-Lig6 complex

Data collection -	
Synchrotron and Beamline	Diamond Light Source, io4-1
Space group	C 1 2 1
Molecules	2
Cell dimensions a, b, c (Å) α , β , γ (°)	116.36, 115.17, 68.27 90.00, 108.42, 90.00
Resolution (Å) R_{merge} $I/\sigma I$ CC (1/2) Completeness (%) Multiplicity	1.52 – 79.70 0.076 7.4 0.998 99.40% 3.2
Refinement -	
Resolution (Å)	79.70 – 1.52
No. reflections	6355
$R_{\text{work}} / R_{\text{free}}$	0.188 / 0.212
No. atoms	5130
R.m.s deviations Bond lengths (Å) Bond angles (°)	0.0182 1.789

Additionally, the binding interaction of Ligand 6 to TbrPDEB1 can be seen below:

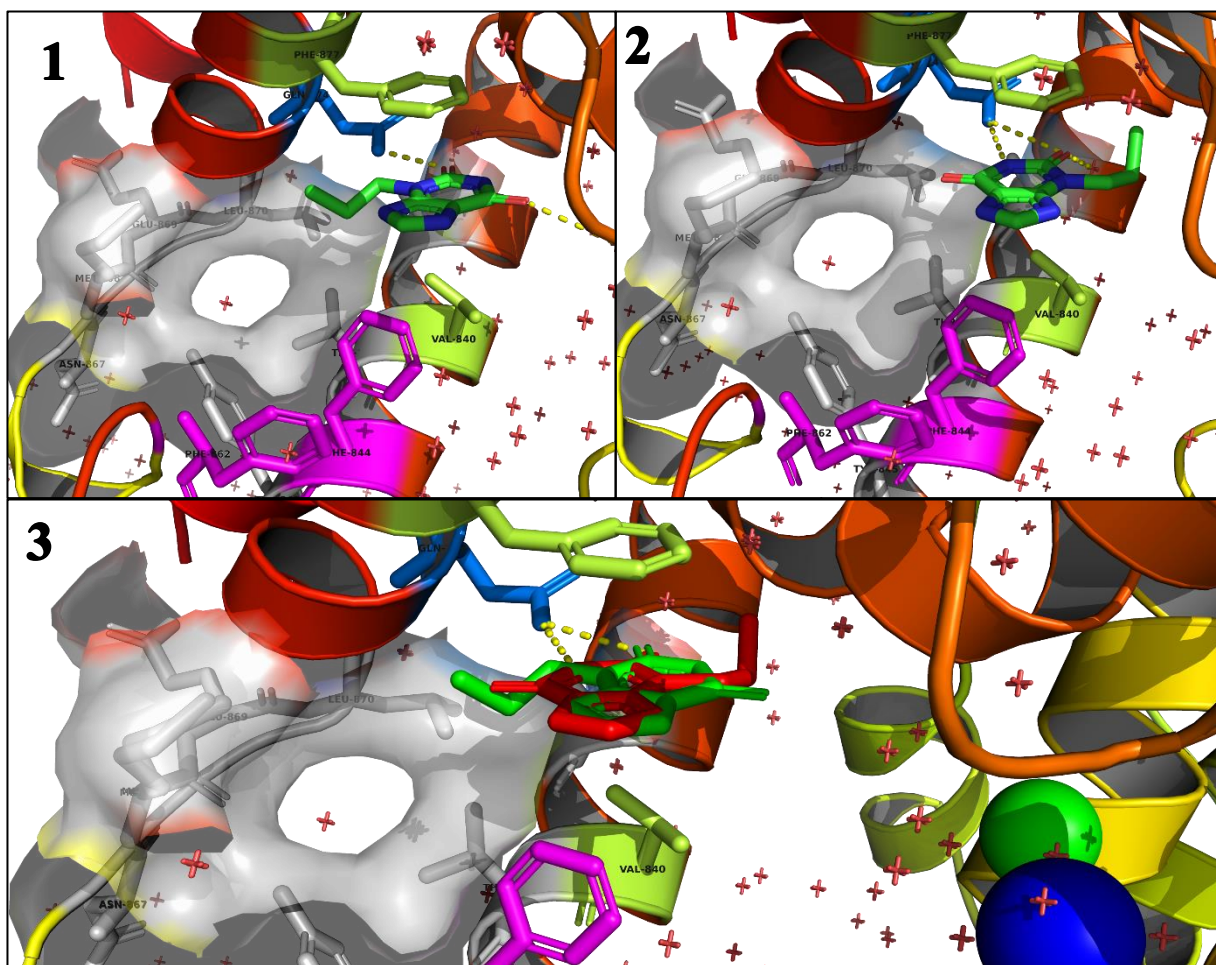


Figure 42:
TbrPDEB1-Lig6 3D structure and different binding modes (1 and 2) and binding mode overlap (3) made with PyMOL

In the case of Ligand 6, upon modelling the ligand within the crystal structure, varying electron densities remained present. Additionally, an average B-factor of atoms of 50.5 and an occupancy of 1.0 also suggests 50% occupancy of the ligand. In other words, it is present in the following position 50% of the time. This is due to the interchangeable binding confirmations of Ligand 6 to the substrate binding pocket, as depicted in *Figure 42*. In the first binding confirmation (1), the oxygen on position 6 of the tetrahydro-1H-purine-2,6-dione rings is seen to form a hydrogen bond 3.19Å in length to the invariant Glutamine Q-switch. In the second confirmation (2) the ligand is flipped 180° horizontally in order to create a new binding mode. In this binding mode, two hydrogen bond formations to the Glutamine Q-switch are suggested; the first by the same oxygen on position 6, forming a bond 3.55Å in length, and the second by the NH on position 1 forming a weaker 3.34Å hydrogen bond (permanent dipole-permanent dipole interaction)

to the nitrogen on the Glutamine residue. This interchangeable bond formation likely causes the latter confirmation to have preferred binding to the substrate binding pocket. The observed tetrahydro-1H-purine scaffold as seen in Ligand 6 is repeated across a number of X-Ray crystallography hits, including Ligand 3 and Ligand 9. Similarly, a close derivative chemotype including a quinazoline/chloroquinazoline scaffold, as seen in Ligands 12, 17, 22 and 29 are also found to be successful X-Ray hits. *Figure 43* below shows the structure of an established competitive hPDE inhibitor, IBMX (3-isobutyl-1-methylxanthine).

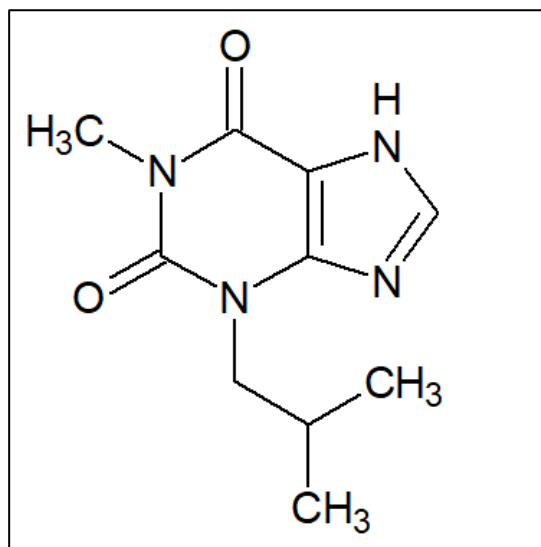


Figure 43:
Molecular structure of IBMX – a known PDE Inhibitor
Drawn on Chemskech

As can be seen, the structure of IBMX is similar to that of the reported fragment hits to TbrPDEB1 (highlighted in red in Appendix 1) by X-Ray Crystallography. The aromatic tetrahydro-1H-purine-2,6-dione scaffold is a key characteristic across most hits, including Ligand 6 previously presented. Substitution of the 5-membered ring to a 6-membered ring, and variation between the position of the Nitrogen atoms give the Phthalazine and Quinazoline derivatives also identified as X-Ray hits in the case of the fragment library used in this project. In other words, it is no surprise that a majority of the set of compounds in *Table 7* were identified by X-Ray crystallography, as the main scaffolds are very close derivatives of the potent non-selective hPDE inhibitor IBMX⁶⁴ and are therefore not novel scaffolds. Nevertheless, the chemotype derivation of some hits to IBMX such as Ligand 29 and understanding the binding interaction to the substrate binding pocket of TbrPDEB1 remains an important task; Being able to confirm and identify compounds in this fragment

library as active binders or potential TbrPDEB1 binders is still of utmost importance and remains the goal of this research project.

Whilst the binding of known scaffold derivatives to the protein target is now understood in greater detail, the binding mode of Ligand 29 needs to be addressed in more detail. As previously mentioned, the importance of the invariant Glutamine Q-switch in parasitic (and human) Phosphodiesterases arises from substrate specificity determination.

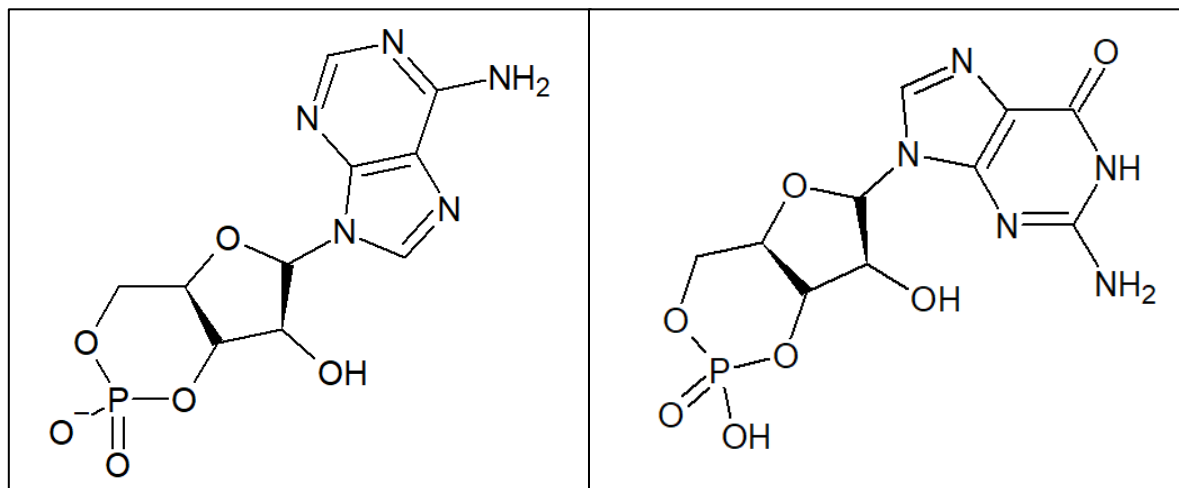


Figure 44:
Molecular structures of cAMP (left) and cGMP (right) drawn on Chems sketch

A 2004 article by K.Y.J. Zhang *et al*⁶⁵ reports this Glutamine switch mechanism for nucleotide selectivity in the case of human Phosphodiesterases. The article discusses the Glutamine-874 residue recognising and forming essential hydrogen bonds to the purine ring present in the cAMP and cGMP substrates and additionally in the fragment hits. This glutamine residue is able to adopt two different orientations⁶⁵ thus enabling substrate dual-specificity (in the case of hPDE1B as an example). In the case of the cAMP-specific TbrPDEB1, the glutamine residue possesses only one confirmation. Upon inhibitor binding to the active site, glutamine hydrogen bond formation is important because of this selectivity mechanism. An already occupied binding site / glutamine residue would prevent cAMP from entering the binding site and in turn the catalytic hydrolysis reaction (assuming a potent enough inhibitor thus to not be displaced by the cAMP substrate). The interesting phenomenon in the case of Ligand 29 is the lack of this glutamine occupation occurring. Instead, the ligand binds to the Asparagine 825 and Valine 826 residues as previously mentioned; However, the high concentration crystal soaks and confirmed ligand binding by X-Ray crystallography is slightly biased and insufficient to determine whether a ligand truly has a binding affinity, or how potent it is. Simply put, the aim of X-Ray crystallography is only to get a general idea of whether a particular ligand chemotype is

capable of binding by extreme methods – a high soak concentration / prolonged soak duration as well as observing the binding interactions. How well a ligand occupies the binding site and competes over the natural cAMP substrate in solution (a more medically realistic view) cannot be determined by X-Ray crystallography. Therefore, the X-Ray hits are further screened and validated by other Biophysical / Biochemical techniques, as will be discussed in the following chapters. This way, the Ligand 29 binding interaction can be examined further by determining the fragment potency.

3.4.3 XCHEM Data Analysis

Looking at *Appendix 30* and *Appendix 31* (and *Table 7*), it is evident that the hit overview comparison between the two techniques does not match. Out of the 31 fragments screened in total, 11 proved to be identified hits by manual soaking, in comparison to 3 successful hits as a result of the XCHEM approach. In the case of poor binding ligands, soak concentration and soak duration were both increased in the case of manual crystallisation in order to test fragments which had not damaged the crystal, as optimal diffraction seemed to occur but appeared to possess a very weak binding affinity as there was no electron density present in the binding pocket. Comparing this to the XCHEM approach, a large number of binders by manual crystallisation such as Ligands 22, 27 and 29 which showed promising $F_o - F_c$ difference maps and even binding affinity confirmation by Biochemical assays and NMR (as seen in the following chapters) did not appear as X-Ray hits in the XCHEM approach. The reason for this is because in most cases, the crystal was either destroyed by the ligand soak or yielded a very poor resolution or an incorrect space group, which is a clear indication of crystal damage occurring. To recap on soak details in both techniques:

As previously mentioned, 2 super stocks of 100mM and 400mM concentration were made up for all 31 compounds. The 400mM super stock was used to make 30mM, 40mM and 50mM manual soak solutions. Note that the ligand super stocks were made up with 100% DMSO, which is capable of heavily damaging the lattice of a crystal at high concentrations. Making a 40mM concentrated solution from the 400mM stock would result in a total of 10% DMSO being present in the soak solution, which appears to be appropriate due to a lack of crystal damage as observed in the manual soaking overview. Note that prior to the XCHEM screening experiment, TbrPDEB1 crystals were initially soaked in a range of DMSO concentrations from 5% to 40% DMSO content within a crystal drop, thus to observe the limits of crystal survival. Crystals appeared to survive up

to a 40% DMSO concentration, however some crystal damage was observed at 40% indicated by slightly higher (and therefore poorer) resolution maps.

The actual XCHEM fragment screen was performed by dispensing appropriate nanolitre drops of 100mM fragment stocks into 400nanolitre crystal drops using an ECHO dispenser to reach the desired soak concentration. Due to using a 100mM stock, the DMSO concentration was 40% instead of 10% as seen in the case of manual soaks. Despite a 40% DMSO concentration being seemingly appropriate following the DMSO crystal survival rate experiment, the question is whether a high DMSO concentration and a high binding affinity ligand in combination has a more damaging effect on a TbrPDEB1 crystal compared to just a high DMSO concentration alone. Going back to the XCHEM hit rate taking this into account, this is a more plausible explanation as the only XCHEM hits (Ligands 1, 3 and 12) were weak binders as indicated by Biochemical assays and NMR in the following chapters. Essentially, a low stock concentration (high DMSO concentration) destroys crystals in the case of higher binding affinity ligands. To put this hypothesis to the test, it would be ideal to repeat the XCHEM fragment screen approach with the same conditions as the manual soaking experiments (30mM and 40mM soaks from a 400mM stock for 24 hours) and observe whether the same hits are determined. Additionally, a separate experiment can be ran in parallel using a high soak concentration (100mM from a 400mM stock) for a much shorter duration, such as 3 hours, in order to best optimise the XCHEM procedure and observe successful ligand hits and the crystal survival rate, as it is concerning that ligands with a higher binding affinity are not identified by the XCHEM approach.

3.5 Luciferase Biochemical Assays

As previously discussed in section 2.6, the PDE4NPD Standard Operating Procedures protocol was followed after slight adaptation for the determination of a suitable working PDE concentration and single concentration assessment experiments of new compounds.

3.5.1 Determining a suitable TbrPDEB1 concentration

Before the identified X-Ray crystallography fragment hits can be screened further, the Biochemical assay must initially be optimised by using a wide range of TbrPDEB1 concentrations thus to give suitable readings according to the positive and negative controls used. Initially, the PDE4NPD protocol followed advised using a range of concentrations starting from a 1:10 PDE dilution and continuing the series in a ten-fold dilution until a 1:100,000 dilution was reached. This protocol was slightly adapted, as the 1:10 TbrPDEB1 dilution seemed to be in excess due to giving much higher RLU (Relative Light Unit) values compared to the AMP positive control and was therefore discarded from the series; This was questionable initially, as it seems invalid that the RLU value of a certain dilution was greater than that of the theoretical maximum RLU value. However, it was suspected that the presence of TbrPDEB1 protein in excess also contributed to detected AMP, and hence resulted in increased RLU values.

Additionally, note that a multi-channel pipette was used to simultaneously add the cAMP reagent to every microplate well in order to initiate the hydrolysis reaction as seen in *Figure 4* at the same time, thus to reduce error in readings and produce more reliable data. Also note that the initial 1:1 concentration was used at 1mg/ml of TbrPDEB1 protein.

Figure 45 on the following page shows a graph depicting the RLU values measured for each TbrPDEB1 dilution:

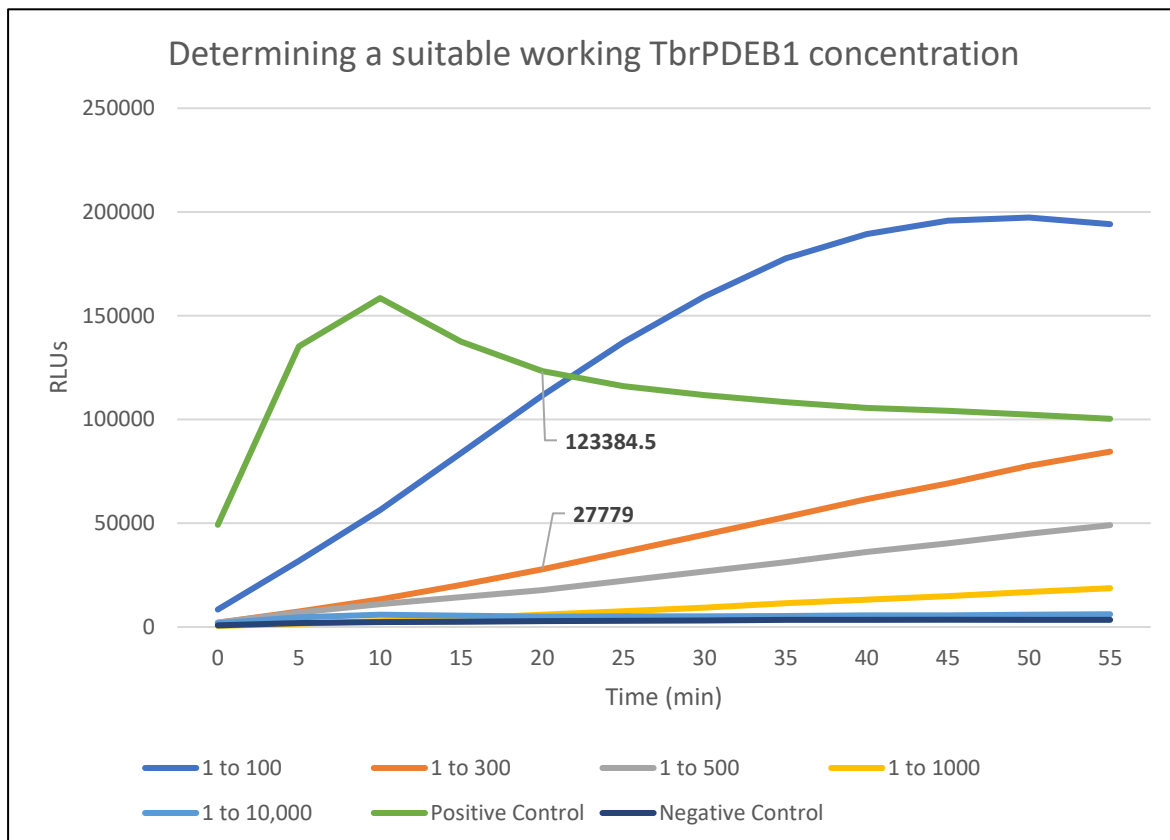


Figure 45:
Determination of a suitable working TbrPDEB1 concentration by use of a Biochemical Assay

As can be seen, the positive control initially rises and slowly begins to decrease following the maximum RLU reading. Initially, the positive control gives the highest light reading, however the 1:100 dilution is seen to surpass this indicating that perhaps this dilution is also in excess. The PDE4NPD protocol instructs to use a dilution that is in the linear range and gives 25% of the maximum RLU value (given by the AMP control) after 20 minutes. The 1:300 dilution is an ideal candidate for this, as it appears linear on the graph; Additionally, the 1:300 dilution gives an RLU value of 27779 after 20 minutes – 22.5% of the measured 123384.5 RLU value of the positive AMP control after 20 minutes. An additional change made to the original protocol can be seen in the dilutions chosen for the series: Instead of following a ten-fold series from 1:10 to 1:100,000, 1:300 and 1:500 were also added to the series. This was done due to the 25% of the maximum RLU value target being observed somewhere between the 1:100 and 1:1000 dilution range, hence an additional 2 dilutions were added in order to increase accuracy. The 1:100,000 dilution was also discarded as it is too dilute, as the 1:10,000 dilution is seen to be very close to the negative control in *Figure 45*. Additionally, TbrPDEB1 enzyme activity is also confirmed. Hence, the determined optimal TbrPDEB1 concentration of 1:300 was then used to run the assay screen against the identified X-Ray crystallography hits from the previous chapter.

3.5.2 Single concentration assessment of identified X-Ray Fragment Hits against TbrPDEB1

100mM DMSO stocks of Ligands 1, 3, 4, 6, 9, 12, 17, 22, 23, 27 and 29 (X-Ray Crystallography hits) were used to prepare the appropriate ligand concentration solution dissolved in Stimulation Buffer + 6% DMSO. The recommended ligand concentration in the PDE4NPD protocol was determined to be 60 μ M; However, this protocol was initially designed to assess officially selected PDE4NPD compounds synthesised to yield as much potency as possible. In the case of this project, mere fragments are being assessed therefore it is expected that the potency yield would be much lower. Hence, an increased ligand concentration of 200 μ M was used instead. NPD-008, a known potent TbrPDEB1 inhibitor previously mentioned in this report was used as a control in this experiment. Fragment inhibition of the TbrPDEB1 target, if any, was compared to the NPD-008 control. *Figure 46* shows a graph depicting the RLU values measured for each Ligand (at a 200 μ M concentration) in TbrPDEB1 protein (diluted to 1:300 from an initial concentration of 1mg/ml) over a 55-minute period. NPD-008 inhibitor, and no inhibitor at all are two controls used in addition to the positive (AMP instead of cAMP) and negative (Stimulation Buffer instead of TbrPDEB1) controls.

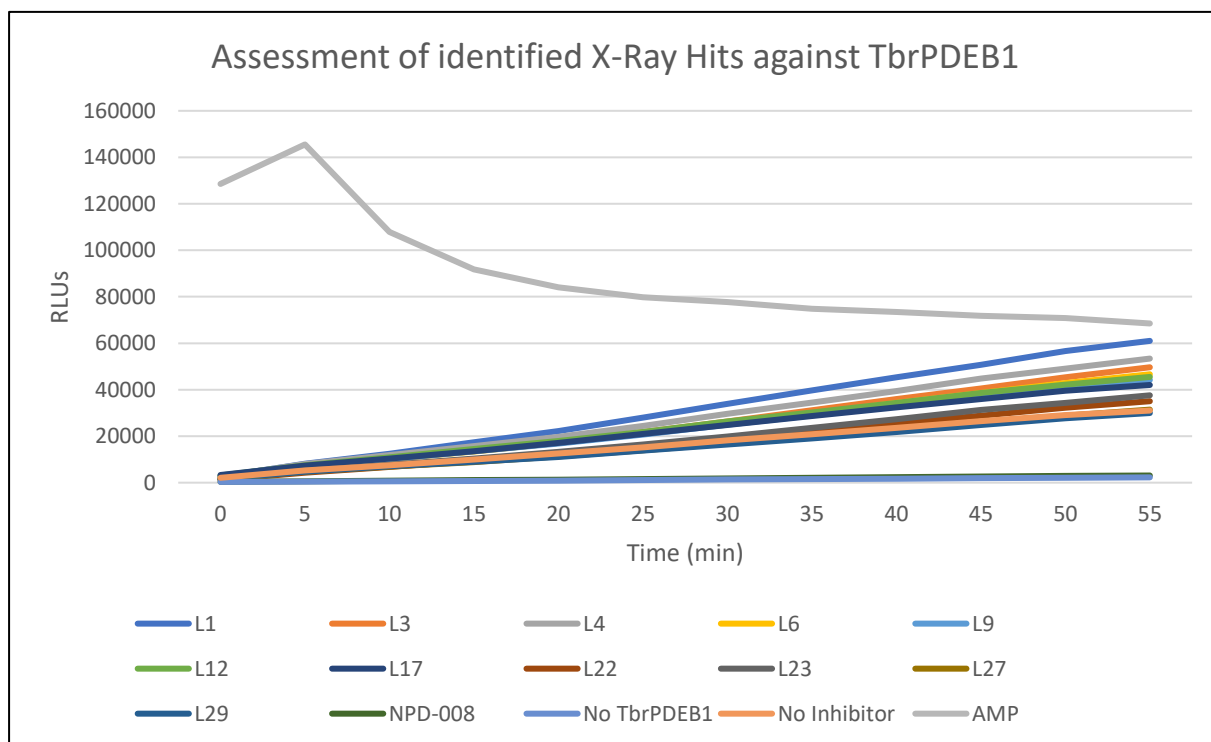


Figure 46:
Assessment screen of X-Ray fragment hits against determined optimal TbrPDEB1 concentration by use of a Biochemical Assay

The remaining activity can be calculated and normalised using the following equation at a selected time point of 20 minutes:

$$\text{Remaining activity} = \frac{(\text{Sample value} - \text{'No TbrPDEB1' control})}{(\text{'No Inhibitor' control} - \text{'No TbrPDEB1' control})} \times 100\%$$

The remaining activity of each ligand was calculated as seen in *Table 11*. Additionally, the percentage inhibition is also seen, calculated using the following equation:

$$\text{Maximum inhibition (\%)} = 100\% - \text{Remaining Activity (\%)}$$

Table 11:
Biochemical Assay determined inhibition of X-Ray hits and controls to TbrPDEB1

Sample	Remaining Activity (%)	Maximum Inhibition (%)
Ligand 1	183.74	-83.74
Ligand 3	143.735	-43.735
Ligand 4	162.44	-62.44
Ligand 6	138.98	-38.98
Ligand 9	136.77	-36.77
Ligand 12	146.95	-46.95
Ligand 17	139.12	-39.12
Ligand 22	97.89	2.11
Ligand 23	105.8	-5.8
Ligand 27	89.5	10.5
Ligand 29	87.14	12.86
NPD-008	3.64	96.36
No TbrPDEB1	0	100
No Inhibitor	100	0

The Remaining Activity for each sample can also be seen in graph form in *Figure 47* on the following page:

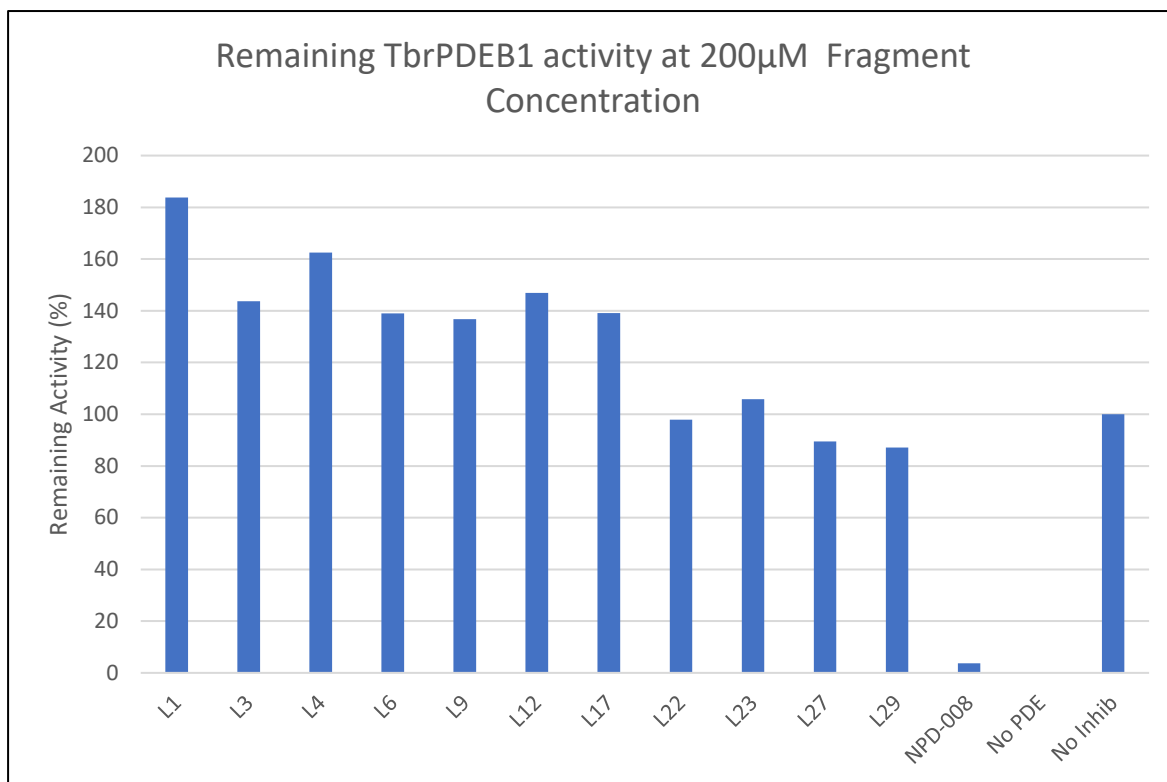


Figure 47:
Remaining TbrPDEB1 activity at 200µM X-Ray Fragment Hit Concentration in graph form

As can be seen, the quality of the screen looks promising: The high concentration of the NPD-008 control yields almost total inhibition of the TbrPDEB1 target enzyme. Additionally, Ligand 29 also shows great promise as it yields the highest Maximum Inhibition percentage out of the other ligands at 12.86% inhibition. This seems incredibly minor compared to the 96.36% inhibition of NPD-008 - however keeping in mind that Ligand 29 is an unrefined fragment it is still promising that a small level of inhibition is observed. Ligand 27 also shows a 10.5% maximum inhibition. The majority of the other ligands show Remaining Activity above 100% due to errors common in Biochemical assays. Despite being successful binders in X-Ray Crystallography, it is apparent that actually in solution most fragments in the set are not potent enough to be able to compete against the cAMP substrate or displace it from the TbrPDEB1 binding site to yield any inhibition.

Additionally, the screen was repeated against the hPDE4D protein target in order to observe and compare fragment binding between the two targets.

3.5.3 Determining a suitable hPDE4D concentration

The protocol used in section 3.4.1 was repeated in the case of hPDE4D for the purpose of optimization and determining an optimal protein concentration to be used for the screening experiment. *Figure 48* depicts this optimization graph as seen below:

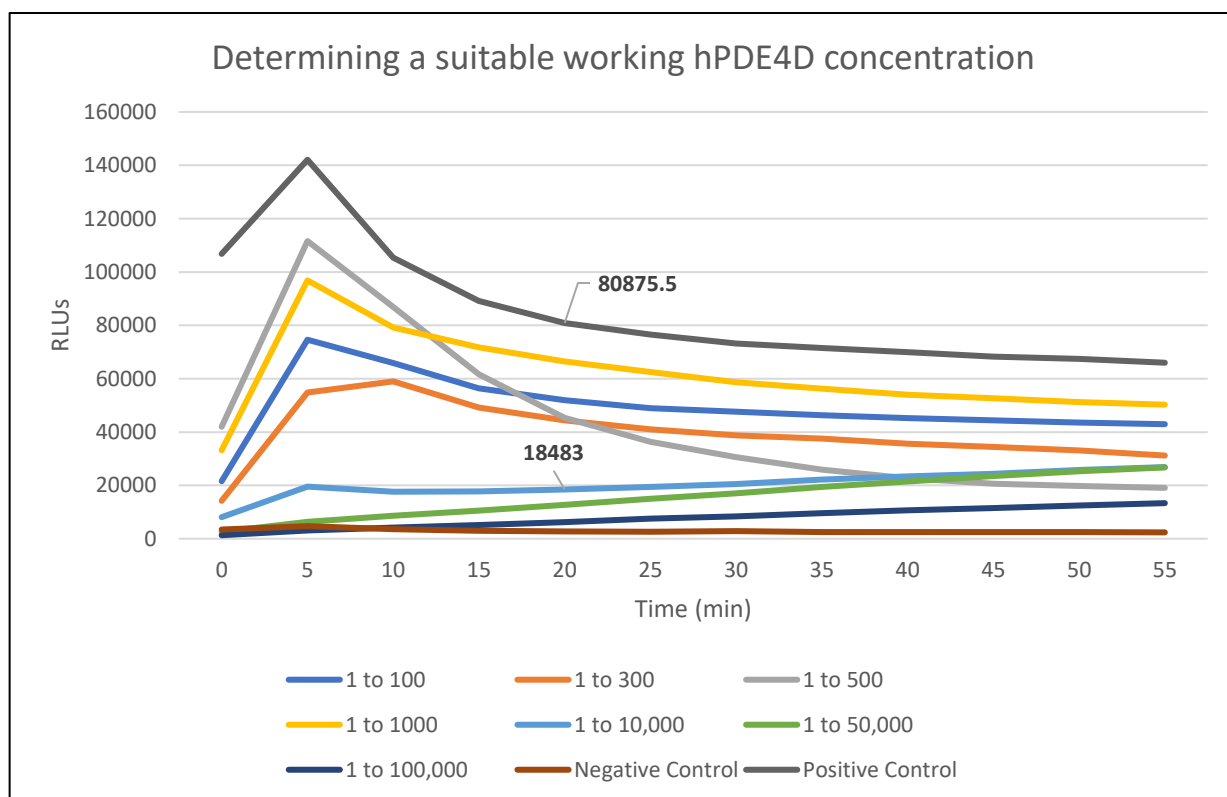


Figure 48:
Determination of a suitable working hPDE4D concentration by use of a Biochemical Assay

As can be seen, a different dilution range is used for this particular assay. Initially, the 1:300 dilution from a starting concentration of 1mg/ml used in the case of the TbrPDEB1 optimization assay was also used for the hPDE4D screen. However, this was incorrect as this assumes both enzymes are of equal activity, which gave an error in the assay as RLU limit detection was being hit very quickly due to the protein being in excess. Hence, the 1:50,000 and 1:100,000 dilutions were also added to this assay after it was made evident that this particular enzyme is much more active and hence requires far less PDE to reach a suitable concentration. It was determined that the 1:10,000 dilution was best suited (from a starting concentration of 1mg/ml) for an assay screen, as it is within a linear range and also gives an RLU value of around 25% compared to the RLU value given by the positive control at 20 minutes, as seen highlighted. Hence, this particular dilution was used to run the assay screen against the identified X-Ray crystallography hits.

3.5.4 Single concentration assessment of identified X-Ray Fragment Hits against hPDE4D

The protocol used in section 3.4.2 was also repeated at 200 μ M fragment concentration against hPDE4D protein target at a 1:10,000 dilution from an initial concentration of 1mg/ml. The purpose of this experiment was to confirm previous speculations on how this particular fragment library would interact with hPDE4D. A similar graph depicting the RLU values measured for each ligand in hPDE4D protein over a 55-minute period can be seen in *Figure 49* below:

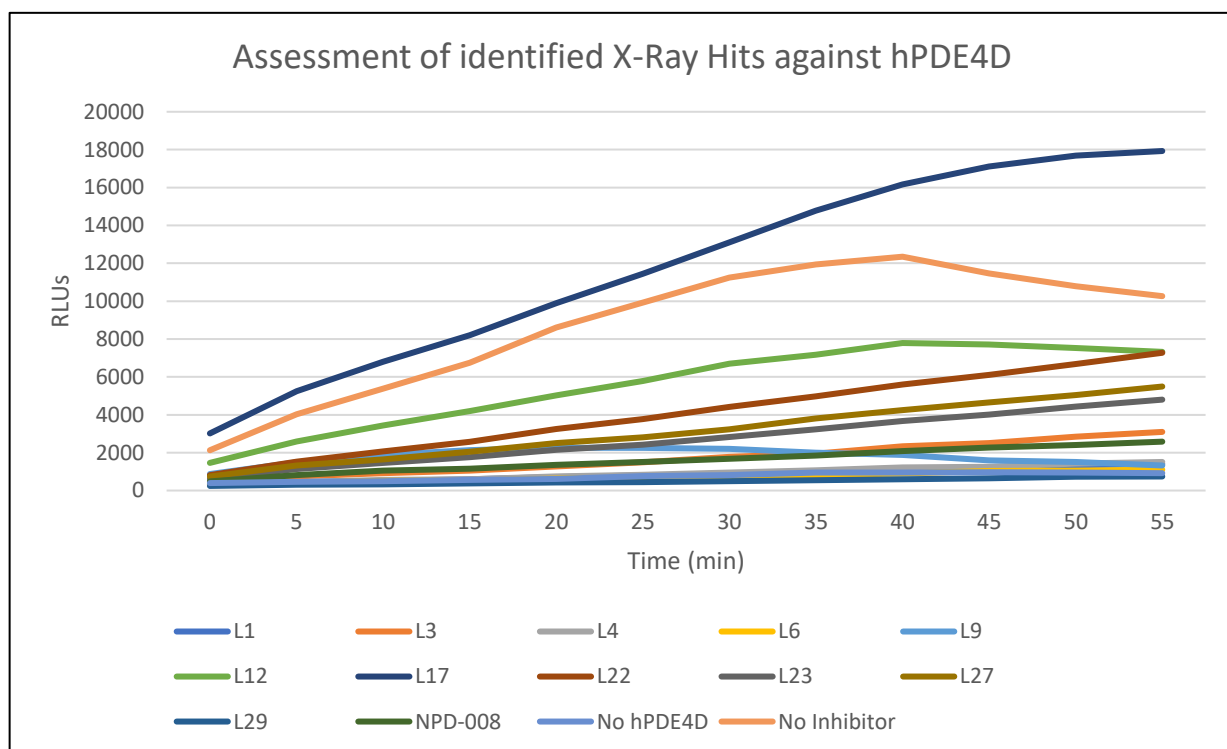


Figure 49:
Assessment screen of X-Ray fragment hits against determined optimal hPDE4D concentration by use of a Biochemical Assay

The remaining activity and maximum inhibition of each ligand were also calculated as shown previously, seen in *Table 12* on the following page:

Table 12:
Biochemical Assay determined inhibition of X-Ray hits and controls to hPDE4D

Sample	Remaining Activity (%)	Maximum Inhibition (%)
Ligand 1	1.23	98.77
Ligand 3	8.27	91.73
Ligand 4	1.58	98.42
Ligand 6	0.13	99.87
Ligand 9	20.56	79.44
Ligand 12	55.07	44.93
Ligand 17	115.99	-15.99
Ligand 22	32.94	67.06
Ligand 23	19.26	80.74
Ligand 27	23.65	76.35
Ligand 29	-2.4	102.4
NPD-008	9.32	90.68
No hPDE4D	0	100
No Inhibitor	100	0

The Remaining Activity for each sample can also be seen in graph form in *Figure 50*:

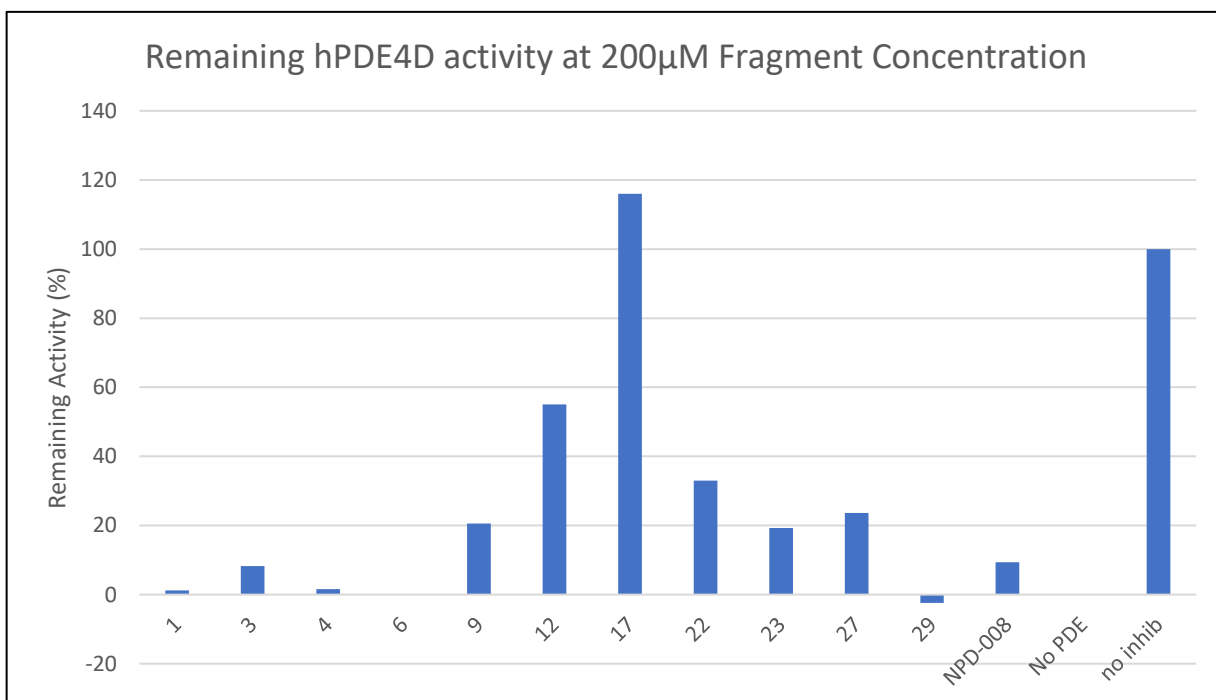


Figure 50:
Remaining hPDE4D activity at 200µM X-Ray Fragment Hit Concentration in graph form

As can be seen, the assay screen results are as expected; At an optimal enzyme concentration, it appears that most ligands used in the screen almost completely inhibit hPDE4D at 200 μ M inhibitor concentration. It is observed that the NPD-008 control also yields a 90.68% inhibition - slightly less than the 96.36% inhibition measured against TbrPDEB1, showing that NPD-008 is indeed slightly selective against the desired target, however not to a degree of optimal selectivity to advance to further trials, as concluded by previously cited literature³³.

As previously mentioned, the fragment library was put together based on their potency against human PDE enzymes. *Table 13* below shows the K_i and IC_{50} values for every fragment used in this particular screen against hPDE4D as reported from ChEMBL⁴⁷. K_i is an inhibitor dissociation constant indicating inhibitor potency and the functional strength of an inhibitor; The lower the K_i value, the greater the binding affinity of a particular inhibitor. The IC_{50} (half maximal inhibitory concentration) of an inhibitor measures the concentration required of a potent compound to give a 50% inhibition response. This can be generated using dose response curves of a particular compound, similarly to the single concentration fragment assessment assays as previously presented, instead using a range of inhibitor concentrations as will be later discussed.

Table 13:
ChEMBL Reported IC_{50} and K_i values of fragments against hPDE4D (μ M)

Lig #	1	3	4	6	9	12	17	22	23	27	29
K_i	/	36500	/	43100	63000	/	/	/	/	/	/
IC₅₀	14,000	/	520	/	/	/	/	231	/	/	1800

These values may be compared to measured IC_{50} data against the TbrPDEB1 target for several fragments following assays generating dose response curves. Ligands 29 and 27 are ideal candidates for this experiment, as both fragments yielded the highest inhibition against TbrPDEB1. Whilst both ligands also yielded high inhibition against hPDE4D, Ligand 27 yielded significantly less (around 30%) compared to Ligand 29, making it an increasingly attractive candidate as it also possesses the essential Q-switch binding as previously discussed. The lack of Q-switch binding of Ligand 29 and large measured inhibition against both targets would also be of interest to compare to Ligand 27.

However, before these assay experiments can be carried out, it is necessary to determine a K_m value for TbrPDEB1.

3.5.5 Km determination of TbrPDEB1

A similar protocol as seen in section 3.4.1 was used for this experiment, with the exception of using a set TbrPDEB1 concentration (at a 1:300 dilution as determined by the previous optimisation assay). Instead, a range of cAMP substrate concentrations were used, as follows: 0 μ M, 0.625 μ M, 1.25 μ M, 2.5 μ M, 5 μ M, 10 μ M, 20 μ M and 40 μ M. A suitable time point was chosen for analysis, determined by observing the point of RLU signal saturation. This time point was used to plot a graph in order to calculate K_m , as seen in *Figure 51* below. A K_m value is derived from and used in Michaelis-Menten enzyme kinetics. K_m is a measurement of the concentration of substrate required to achieve half of the maximum enzymatic response, described by $\frac{1}{2} V_{max}$ (in this case, half of the maximum RLU value). In other words, K_m can be used to represent the overall enzyme efficiency by determining the rate at which the enzyme is capable of binding to the substrate, and hence the overall enzyme-substrate affinity. A low K_m implies the enzyme has a higher affinity towards the substrate, and vice versa. In this case, the V_{max} is observed to be 72,024 RLU as seen in *Figure 51*, resulting in a $\frac{1}{2} V_{max}$ of 36,012 RLU, as indicated on the graph.

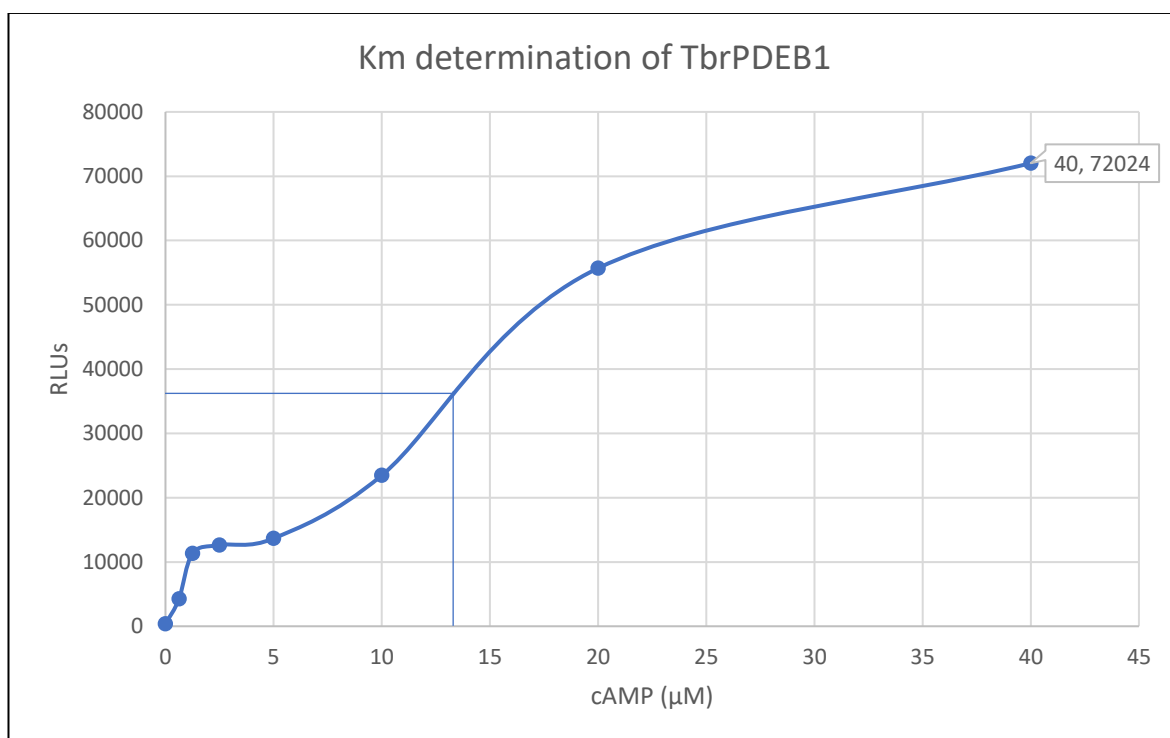


Figure 51:
TbrPDEB1 enzyme K_m determination Biochemical assay

By taking the relevant equivalence point, 36,012 RLU is equivalent to 13.5 μ M cAMP substrate concentration. Compared to the recorded 1.19 K_m value of TbrPDEB1 in the PDE4NPD protocol, 13.5 is a significantly larger K_m value. A reason for this could be due

to enzyme activity variability; As previously shown, a 1:300 dilution was required to give optimal linear results for the Biochemical assay optimisation, which also implies a less active enzyme. Nevertheless, Dose response curve calculations can be appropriately adjusted using the newly recorded K_m value.

3.5.6 Dose response curve Biochemical assays

With the optimised cAMP substrate and target protein concentration, dose response curves were ran following the TbrPDEB1 fragment screen. 10mM stock solutions were prepared for the appropriate compounds being tested. 6% of ligand stock solutions were mixed in a total 100 μ l of stimulation buffer in the first row of a Corning 96-well plate resulting in a 6×10^{-4} M or 600 μ M concentration. The serial dilution continued to a final concentration of 6×10^{-10} M or 0.6nanomolar, with the last row of the plate consisting of 6% DMSO control in a total of 100 μ l of stimulation buffer. 17.5 μ l of the compound solution wells was transferred to a fresh Corning 96-well plate with a multichannel pipette and the procedure continued as described previously. The plate was read for a total of 55 minutes, and a suitable time point was chosen for analysis for each curve. Additionally, a positive and negative control were also ran as seen in previous assays. The cAMP substrate concentration remained at 10 μ M, just slightly below the calculated K_m value of 13.5 μ M. The protein dilution also remained at 1:300. The purpose of these assay experiment was to determine the IC_{50} of the most promising fragments as identified by X-Ray crystallography and previous Biochemical assays by observing the particular concentration range where an inhibition response is seen.

Hence, Ligand 27 and Ligand 29, with the addition of NPD-008 as a control were tested in dose response curves, as seen in the following pages:

3.5.6.1 Ligand 29 Dose response curve

The assay experiment curve can be seen in *Figure 52* below: Note that the logarithm of the inhibitor concentration range was used for the x-axis.

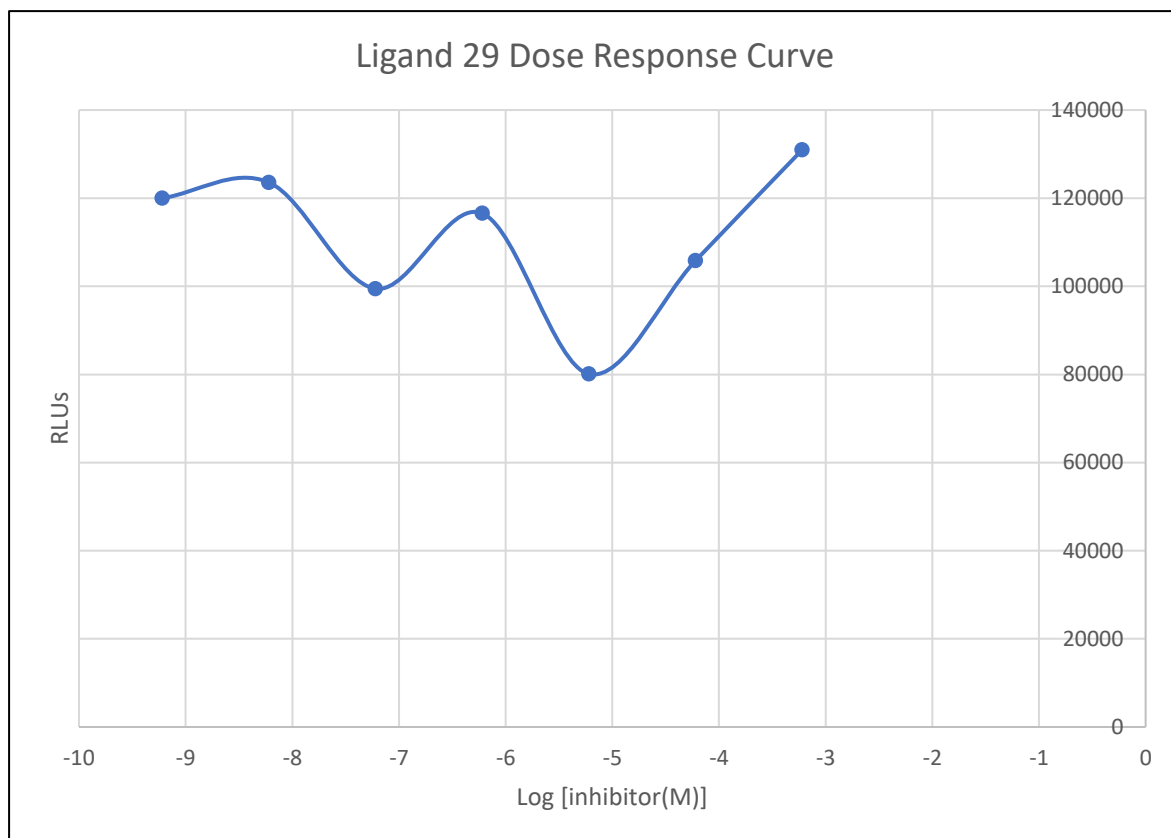


Figure 52:
Ligand 29 dose response curve against TbrPDEB1

As can be seen, no real trend can be deduced from this particular curve. The decreasing RLU value does not seem to be proportional with increasing inhibitor concentration as would be expected with a potent inhibitor. The RLU fluctuation is likely due to assay reading errors, as in this case measured inhibition is practically non-existent. As previously seen in section 3.5.2, 12.86% inhibition of TbrPDEB1 was measured in the case of Ligand 29 at 200 μ M concentration. At the beginning of the serial dilution (600 μ M or log[-3.22M]), the inhibitor is likely in excess, and inhibition past this concentration would be very difficult to observe. In other words, despite yielding the most inhibition against TbrPDEB1, this compound is simply not potent enough to be viable for dose response curves or yield any IC₅₀ value as 50% inhibition cannot be achieved.

3.5.6.2 Ligand 27 Dose response curve

This is similarly seen in the case of Ligand 27, as can be seen in *Figure 53* below:

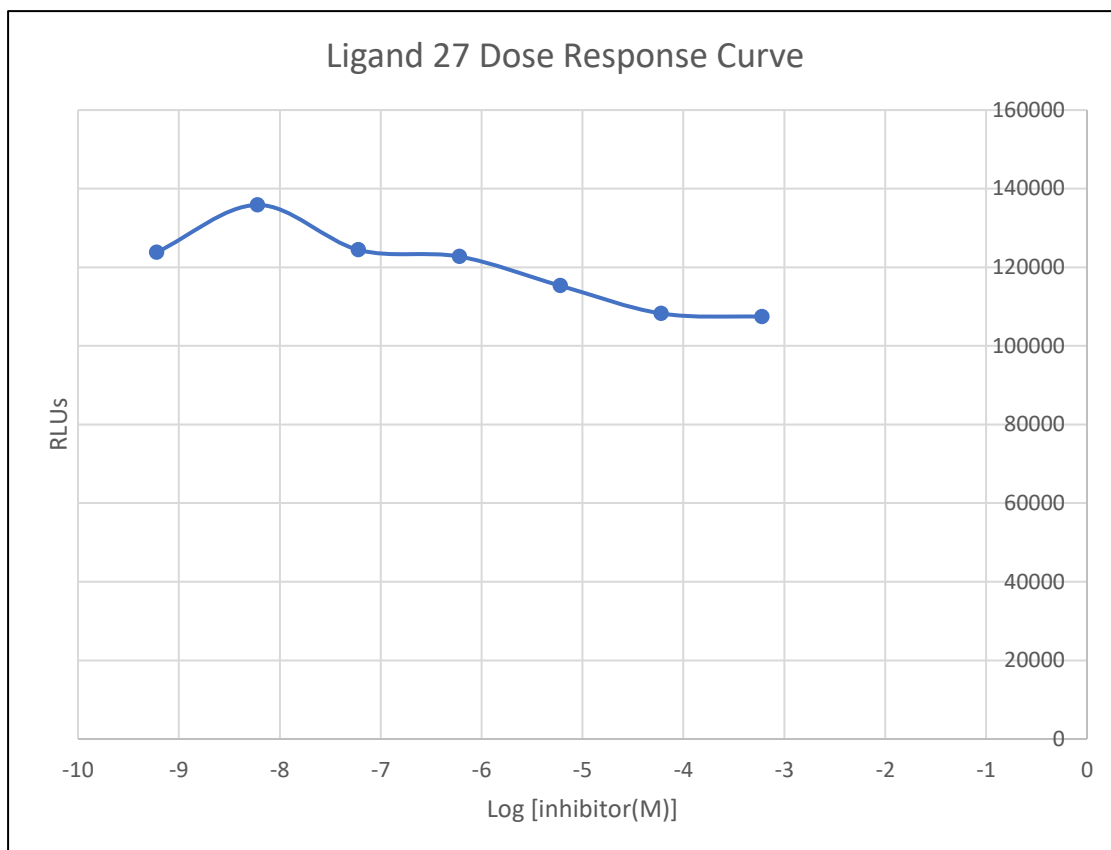


Figure 53:
Ligand 27 dose response curve against TbrPDEB1

As previously mentioned, a Dose response curve was also produced using NPD-008 as a control, as can be seen in *Figure 54* on the following page.

3.5.6.3 NPD-008 Dose response curve

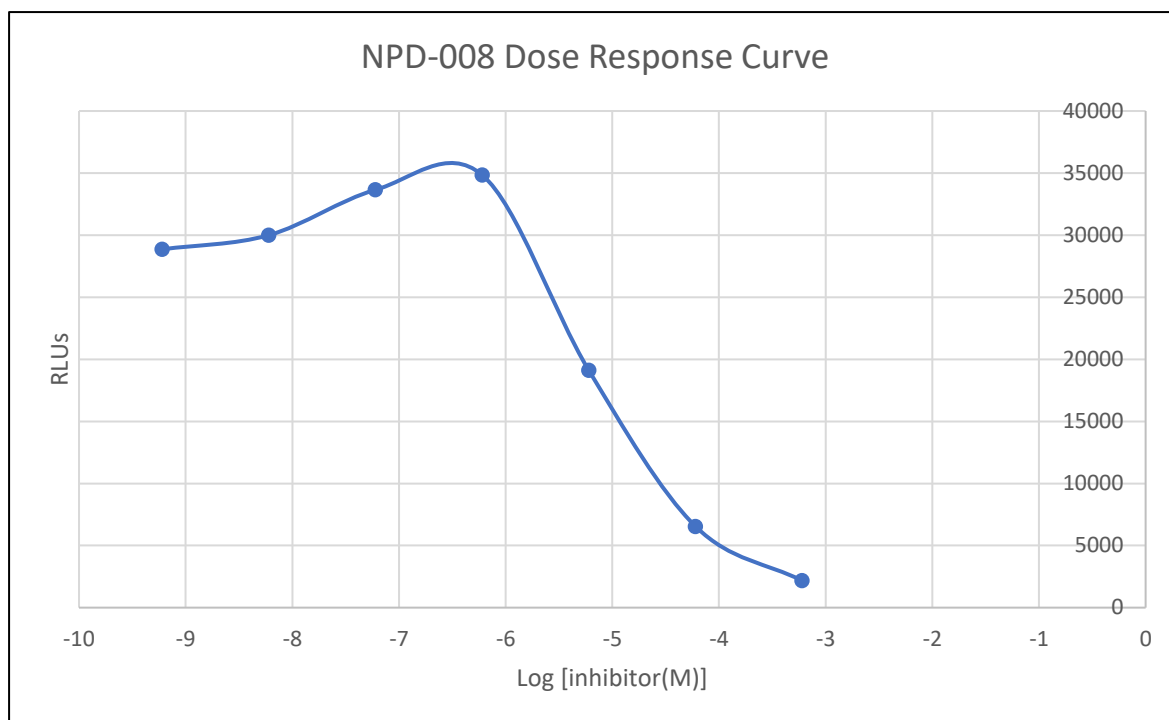


Figure 54:
NPD-008 dose response curve against TbrPDEB1

In the case of NPD-008, the curve is clearly seen, and a concentration point at which an inhibition effect is observed can be seen. The IC_{50} can be seen in *Table 14*, calculated by inverting the logarithm concentration at which 50% inhibition is observed:

Table 14:
NPD-008 IC_{50} table

	NPD-008
Bottom	2168.667
Top	34849
Log IC_{50}	-5.2814
IC_{50}	$5.23 \times 10^{-6} M$

The IC_{50} is calculated to be $5.23 \times 10^{-6} M$, or $5.23 \mu M$. This agrees with the reported IC_{50} value of NPD-008 of $5.5 \mu M$ against TbrPDEB1 as seen in the previously cited article by A.R. Blaazer *et al*³³ on a similar study. This in turn increases confidence in the dose response curves of Ligands 29 and 27.

A lack of compound potency in the case of Ligands 29 and 27 is due to ineffective binding to the TbrPDEB1 substrate binding pocket compared to cAMP substrate. However, this is expected in the case of fragments. Nevertheless, novel fragments capable of binding to TbrPDEB1 have been identified within this research project. Despite not showing optimal target inhibition, this is still beneficial as the ‘cherry-picked’ binders from the starting fragment library can be modified in future work thus to increase potency and in turn inhibition against the target enzyme.

3.6 Ligand-Protein NMR

The final screening method used in this project was through several NMR techniques ideal for observing protein-ligand interactions. The purpose of this additional technique is to support the results yielded from the Biochemical assays by being able to identify fragment binding to the TbrPDEB1 target. Like the Biochemical assay screen, this technique can quantitatively measure ligand binding affinity to the protein target indirectly by exploiting several NMR techniques such as the Nuclear Overhauser Effect (NOE), as will be later discussed. *Table 15* below shows a brief overview of the X-Ray Crystallisation hits screened by each NMR method:

Table 15:
Overview of X-Ray Crystallography hits screened by various protein-ligand NMR techniques

Ligand #	WaterLOGSY Hit	STD Hit	CPMG Hit
1	✓	✓	✓
3	✗	✗	✗
4	✓	✗	✓
6	✗	✗	✗
9	✗	✓	✗
12	✗	✗	✓
17	✗	✗	✗
22	✗	✗	✓
23	✗	✗	✗
27	✗	✗	✗
29	✓	✓	✓

This will be expanded upon further in the following pages.

3.6.1 TbrPDEB1 1D Proton NMR Spectrum

A simple 1D Proton NMR experiment was ran on the TbrPDEB1 target protein as previously mentioned in section 2.8.2. The respective spectrum yielded from this experiment can be seen in *Figure 55*:

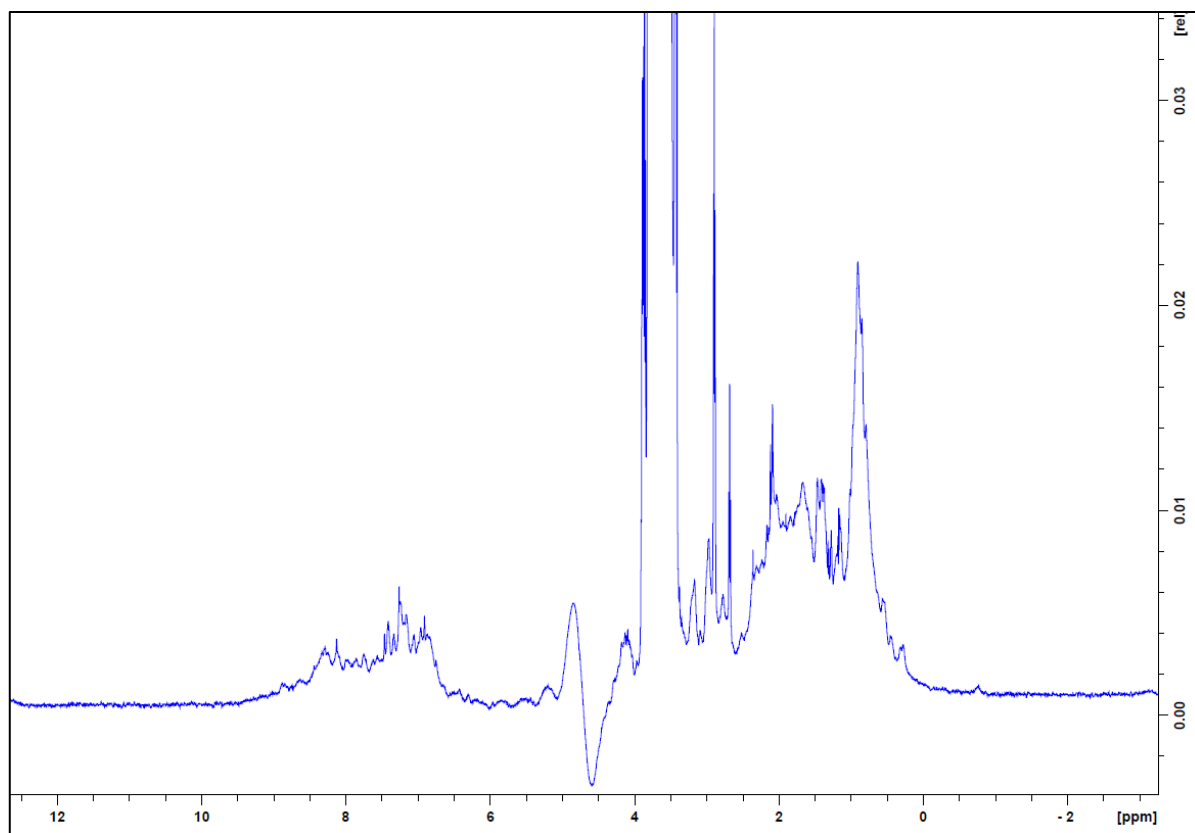


Figure 55:
TbrPDEB1 protein Proton 1D NMR Spectrum

As can be seen, a ^1H NMR spectrum of a folded protein is far more complex in comparison to simple molecules or fragments and contains many regions of interest. The jagged peak regions between 6.5-8.5ppm correspond to the amide NH hydrogens present in the peptide backbone and amino acid side chain residues of the protein. The region between 6.5-7.5ppm also corresponds to the aromatic rings present in amino acid residues. The inverted peak as seen at 4.5ppm is an artefact as a result of water suppression on the spectrum. Another key region common in folded proteins is seen between 0.5-2ppm region corresponding to hydrogen atoms present on CH_3 methyl side chains. Additionally, the peaks seen between 2.2-3ppm also correspond to the open, aliphatic hydrocarbon chains. A large impurity peak is seen 3.7ppm; This impurity peak was identified as remaining glycerol present in the protein sample from the previous storage buffer. Despite excessive buffer exchanging, the glycerol remained abundant, yielding a large impurity peak present

in the protein sample. A 1D Proton NMR spectrum of glycerol was determined as can be seen in *Figure 56* (right) and compared to the impurity present in the TbrPDEB1 spectrum (left).

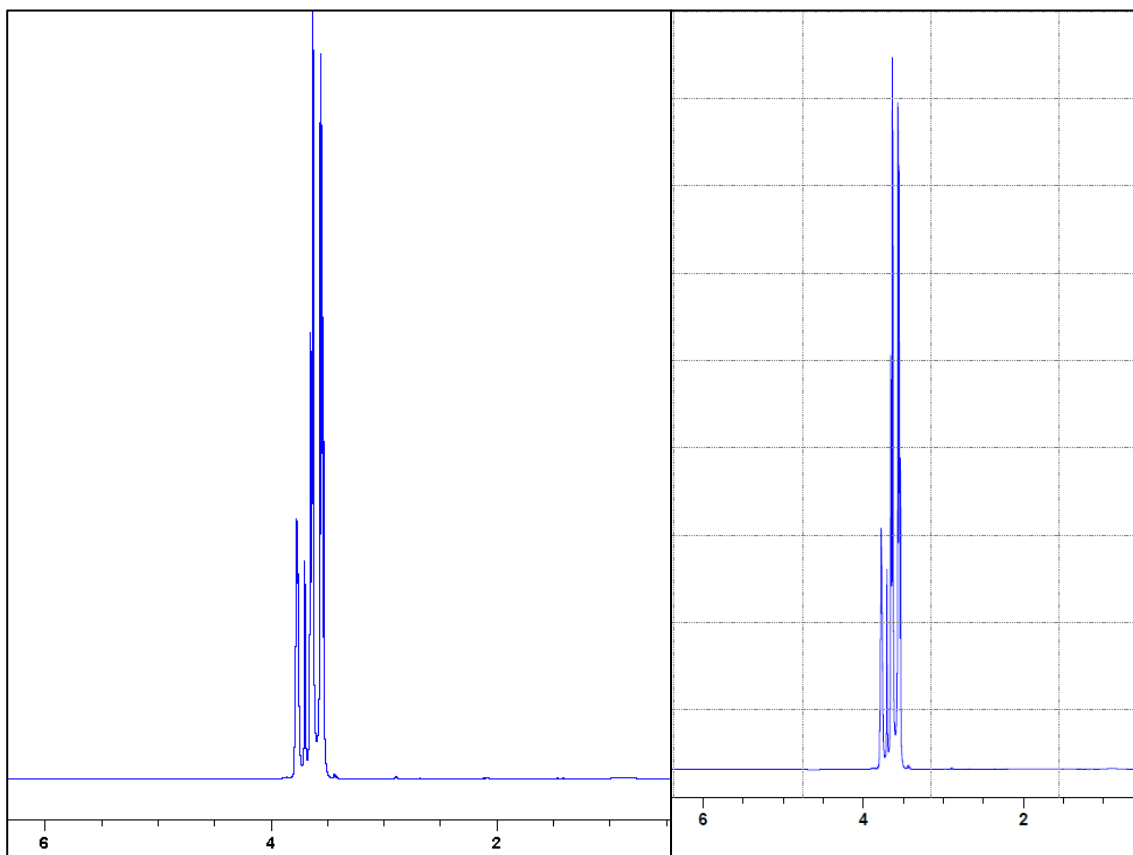


Figure 56:
*TbrPDEB1 Proton 1D NMR spectrum impurity (left) and glycerol
Proton 1D NMR spectrum (right)*

3.6.2 Protein-Ligand Interaction NMR

Whilst it is promising to obtain validation of the target protein integrity by ^1H NMR, it is not beneficial in terms of protein-ligand interaction experiments. As previously mentioned in section 2.8.3, the samples tested for the protein-ligand experiments consisted of 0.5mM ligand and 0.01mM protein. At such a low concentration, protein peaks would be incredibly faint and difficult to distinguish. Instead, binding affinity can be observed by comparing ligand behaviour in presence of protein against a ligand solution with the same concentration. These ligand-based screening approaches rely on fast dissociation of bound ligands and measuring the Nuclear Overhauser Effect (NOE) relaxation time in order to observe differences between protein-ligand and ligand spectra⁷³. These experiments are ideal for fragments, as a low affinity to the protein target is expected and also required in order to achieve a rapid exchange between the bound and unbound forms. In other words, high affinity / tightly bound ligands would have a very short or no dissociation from the protein structure at all and would therefore not be ideal for these experiments.

These particular techniques are based on observing the intensity of magnetization transferred to the free, unbound ligand state by the protein-bound ligand protons⁷⁴. Again, note that both of these states are in constant exchange due to the rapid dissociation from the protein-bound to the unbound ligand state. It is expected that a small molecular weight molecule such as a fragment would have a small positive NOE value, whilst a larger molecule such as a 40.57kDA protein would have a large negative NOE value⁷⁴. Upon ligand binding to the protein, the magnetization transfer to the ligand in bound state following excitation on the protein would also cause it to acquire a large negative NOE state. Upon dissociation, the new NOE state of unbound ligand differs from its original small positive NOE value. This change in NOE values as a result of magnetization transferred is indirectly proportional to ligand-protein binding. The observed response is dependent on the quantity of unbound ligand experiencing a magnetization transfer, which in turn is dependent on the quantity of unbound ligand bound to the protein to begin with⁷⁴. Hence, binding affinity can be indirectly and quantitatively observed.

The techniques used are discussed in greater detail and presented in the case of Ligand 29 and Ligand 27 in the following pages. The NMR spectra of the other ligands can be seen presented from *Appendix 41* to *Appendix 67* in the Appendices.

3.6.2.1 Ligand 29 waterLOGSY NMR

This particular experiment uses a phenomenon similar to that previously described above, however the transfer arises from bulk water magnetization. In other words, present water molecules in the protein-ligand sample are excited and transfer a bulk magnetization moment to the protein. This in turn is transferred to the ligand⁷⁴; The difference in magnetization can be quantified upon dissociation as previously discussed. In this case, the observed response would be a resonance with opposing peaks in the case of bound compounds⁷⁴. This can be seen in *Figure 57*, depicting the waterLOGSY NMR spectra for Ligand 29.

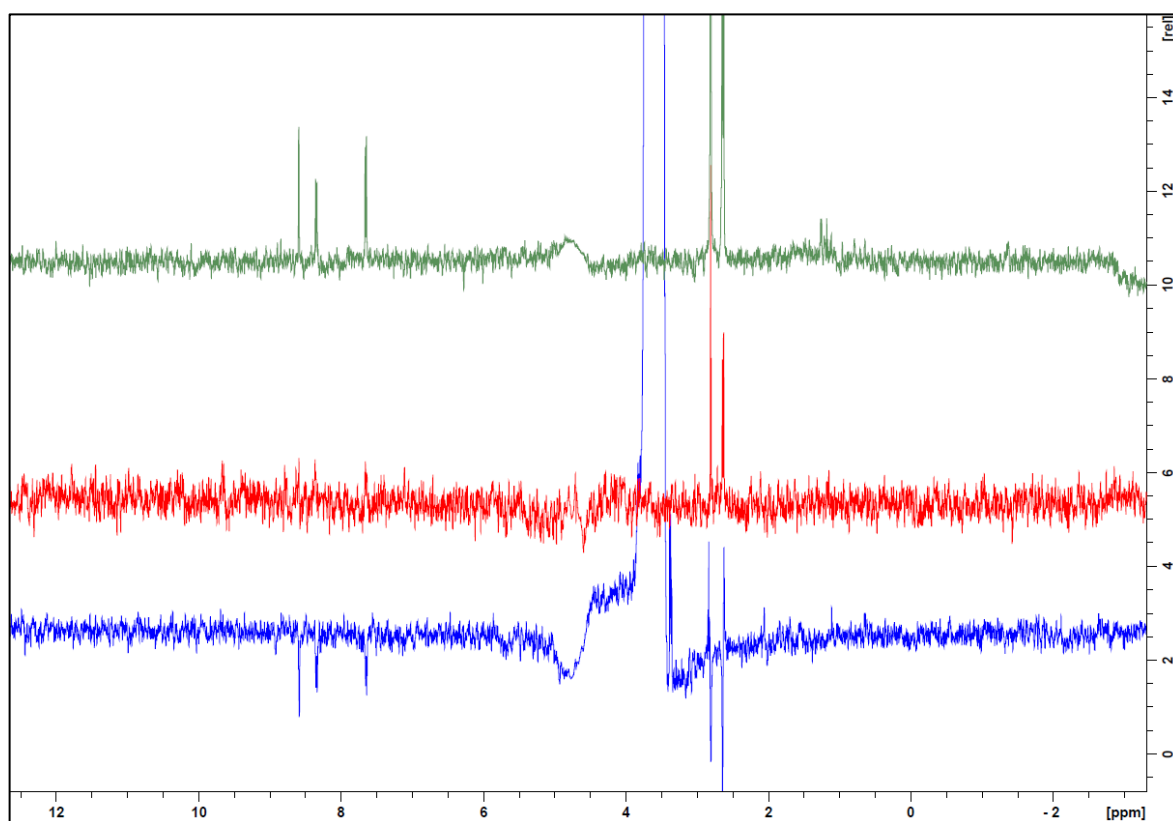


Figure 57:
Ligand 29-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 29 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)

As can be seen, Ligand 29 peaks (seen present in the reference spectrum in green) are seen inverted in the Ligand 29-TbrPDEB1 waterLOGSY spectrum in blue. The spectrum in red is another waterLOGSY experiment performed on a ligand sample outside of protein solution, in order to normalise the signal intensity as the ligand can directly acquire magnetisation from the bulk solvent present in the sample which can lead to false positives⁷⁴. In order to correctly compare each spectrum, the phases and baselines of all spectra were corrected and normalized. The line broadening was also corrected to a value

of 3hz. In the case of Ligand 29, a successful hit can be seen for the waterLOGSY experiment. Note however, that the 1D Proton spectrum in green may vary from the 1D Proton spectrum determined from the Compound QC as a result of a different solvent being present, and the possibility of peak shifting to occur. The 1D NMR Proton spectrum of the phosphate solvent used to make up each solution can be seen in *Appendix 40*; however, no impurities appear to be present. Despite this, the purpose of the previous QC experiments was simply to verify the purity of each compound prior to any further experimentation. Since no compounds were flagged as a result of a questionable purity due to excessive contamination, peak identification in the case of Protein-Ligand experiments are not as vital.

3.6.2.2 Ligand 29 STD NMR

Similarly, Saturation Transfer Difference NMR works by initially saturating a protein by exciting a local region, which in turn affects the spin states of the protons on the entire protein, including the binding site hydrogens. Bound ligand protons would also experience this change in polarization, and in turn the unbound state ligand hydrogens, causing the response to be quantitatively measured as previously discussed⁷⁴ and observed by a reduced signal on the spectrum. In addition to protein saturation, another experiment is run simultaneously where an off-resonance region is excited, thus not saturating the protein. The purpose of this is to cancel out any background noise once a difference spectrum is taken, leaving only peaks of any molecule that has at some point experienced a change in magnetization and hence an association to the protein. In other words, in the case of a bound ligand, low intensity peaks would be observed in the difference spectrum. An empty difference spectrum is a result of no magnetization transfer occurring due to unsuccessful ligand binding⁷⁵ and cancelling out of the on-resonance and off-resonance spectra. This is depicted in the case of Ligand 29 as shown in *Figure 58* on the following page:

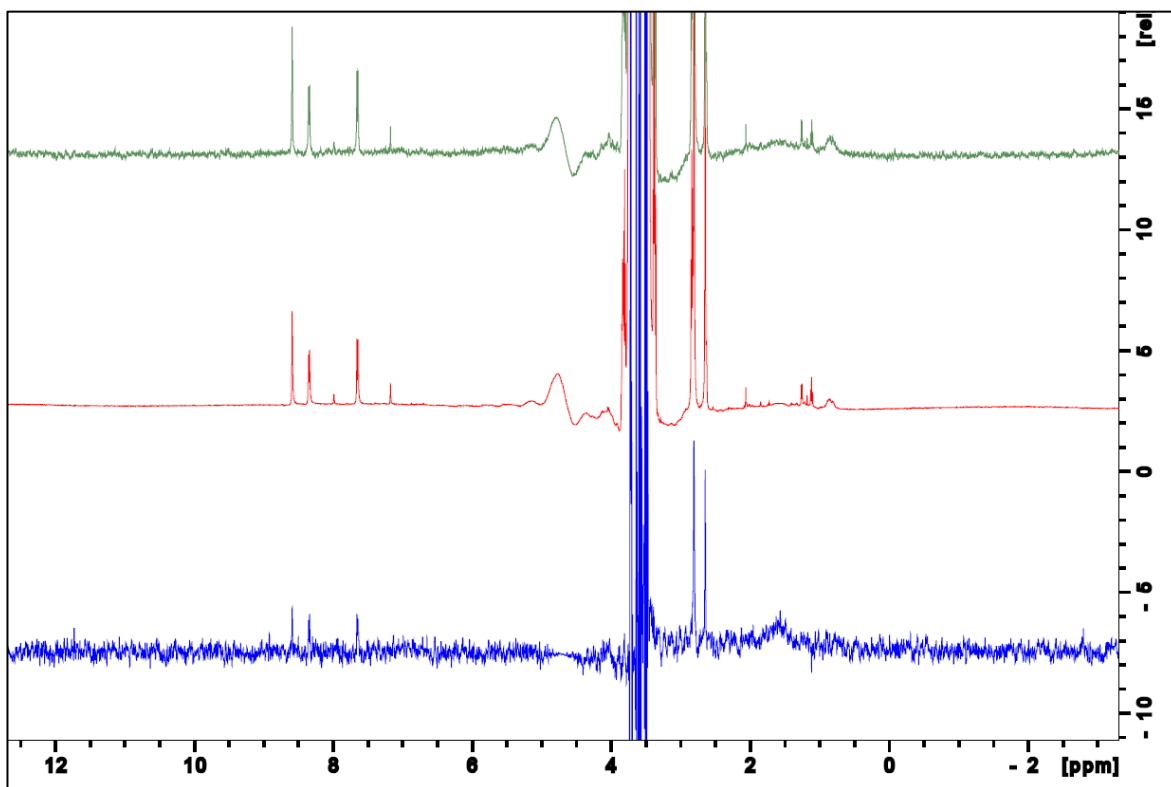


Figure 58:
*Ligand 29-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top),
 Ligand 29 on-resonance control spectrum (Middle), difference spectrum (Bottom)*

As can be seen, ligand-protein binding is indicated by STD NMR due to reduced intensity ligand peaks being present in the difference spectrum (in blue), implying a successful magnetization transfer on the ligand occurring. Like in the waterLOGSY experiment, the phases were adjusted so that they are equal, and the line broadening was increased to 3hz.

An STD experiment was also performed on the ligand solution, as seen in *Figure 59* on the following page. The purpose of this simultaneous STD experiment is to ensure no false positives occur as a result of directly exciting the ligand. A lack of signals in the difference spectrum (in blue) indicates the ligand is not directly excited, and the presence of peaks in the difference spectrum in *Figure 58* is indeed due to successful ligand-protein binding, rather than a false positive.

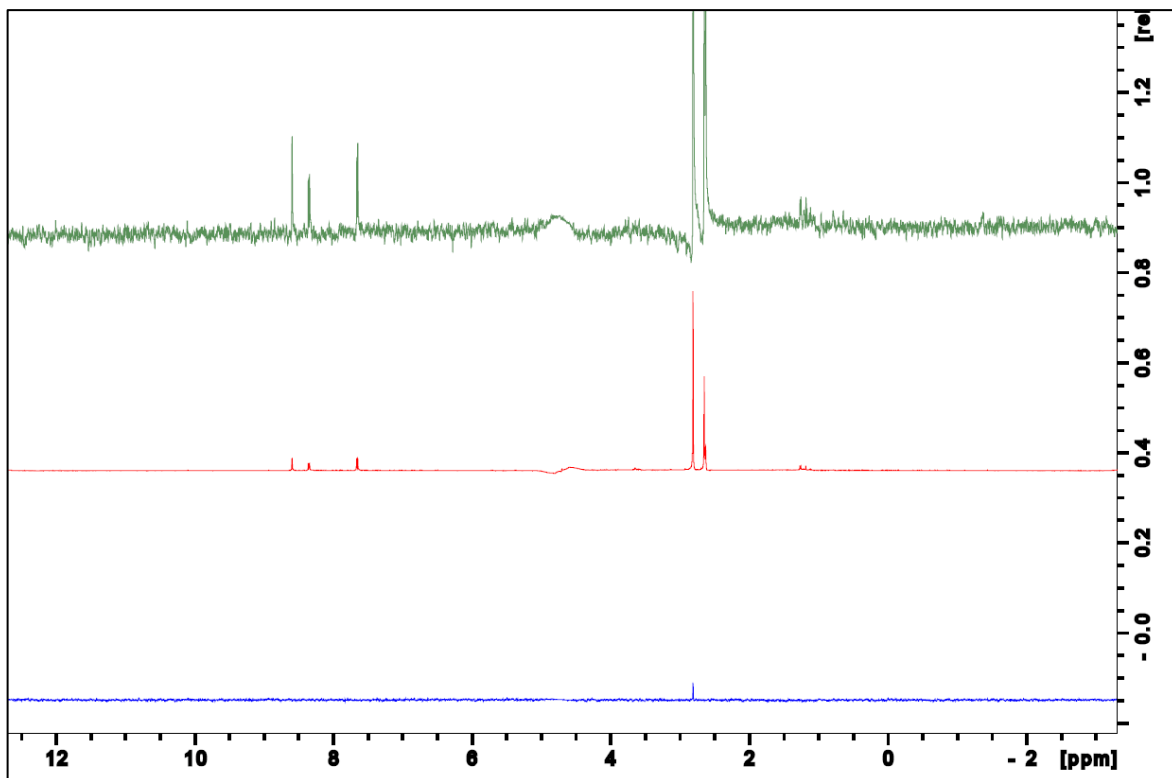


Figure 59:
Ligand 29 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 29 on-resonance control spectrum (Middle), difference spectrum (Bottom)

3.6.2.3 Ligand 29 CPMG NMR

CPMG NMR is the final protein-ligand affinity determining experiment. CPMG NMR works by measuring spin-spin T_2 relaxation times and applying a series of spin-echo pulse elements to transverse magnetization⁷⁶ in order to detect ligand exchange processes. In practice, a CPMG spectrum of the ligand sample is taken (which has a fixed relaxation time) and compared to the CPMG spectrum of the ligand + protein containing sample. Variability in T_2 relaxation times as a result of successful magnetisation corresponds to successful ligand binding, which is indicated by reduced peak signals on the protein-ligand NMR spectrum⁷⁷.

This can be seen in the case of Ligand 29 in *Figure 60* on the next page:

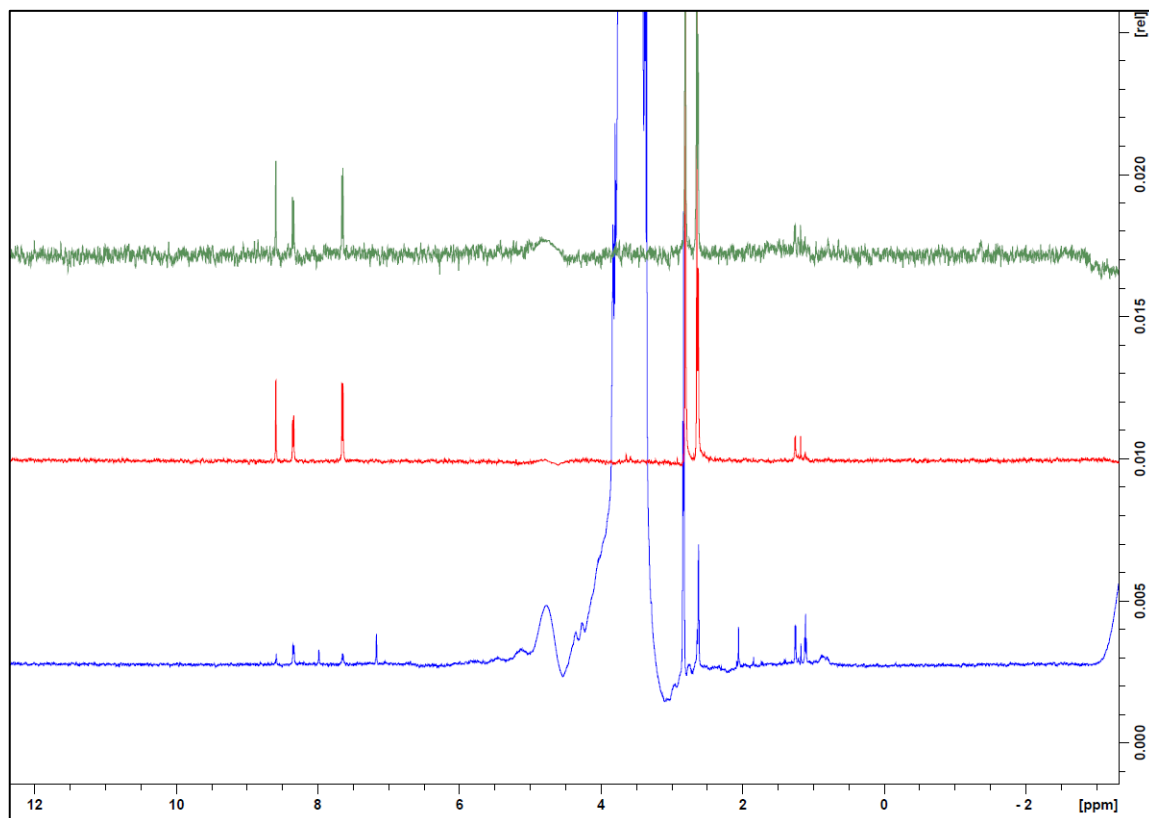


Figure 60:
*Ligand 29-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top),
Ligand 29 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum
(Bottom)*

As can be seen, the reduced peak intensities as seen in the protein-ligand CPMG spectrum indicated in blue suggest successful binding determined by CPMG NMR. In addition to determination by waterLOGSY NMR and STD NMR, Ligand 29 is confirmed as a successful hit to the TbrPDEB1 target by NMR protein-ligand interactions as well as Biochemical assays and X-Ray crystallography.

3.6.2.4 Ligand 27 waterLOGSY NMR

In the case of Ligand 27, binding to TbrPDEB1 does not occur, as can be seen in *Figure 61* below:

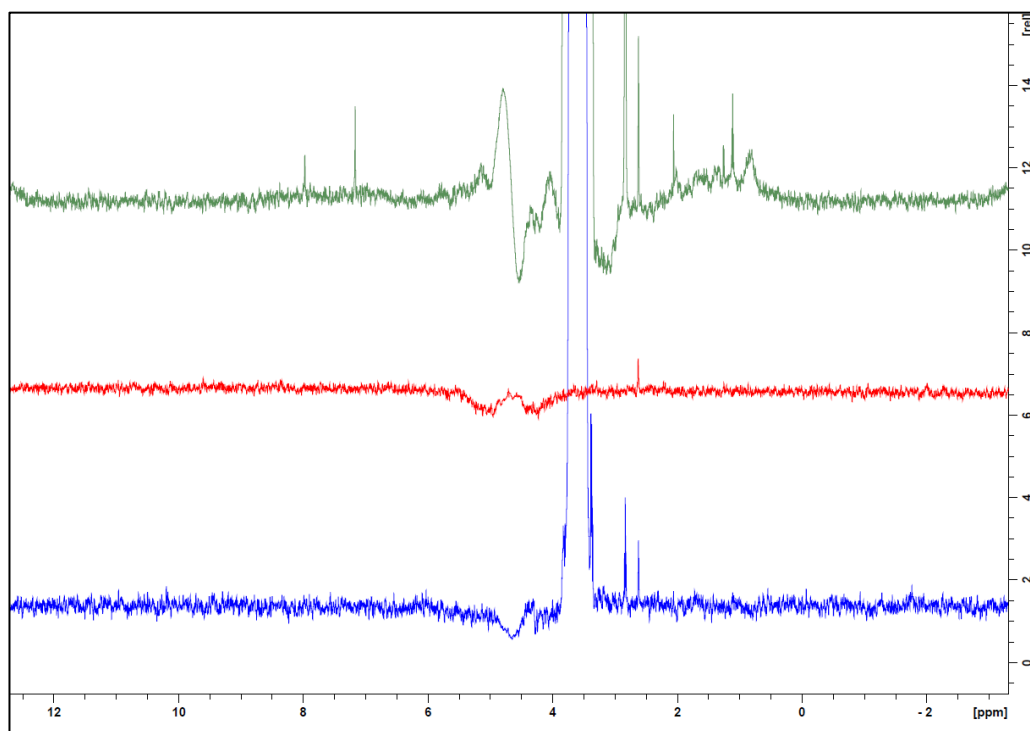


Figure 61:
Ligand 27-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 27 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)

Additionally, the waterLOGSY control spectrum as seen in red shows no signal peaks.

3.6.2.5 Ligand 27 STD NMR

Furthermore, Ligand 27 also appears as a non-binder in STD NMR:

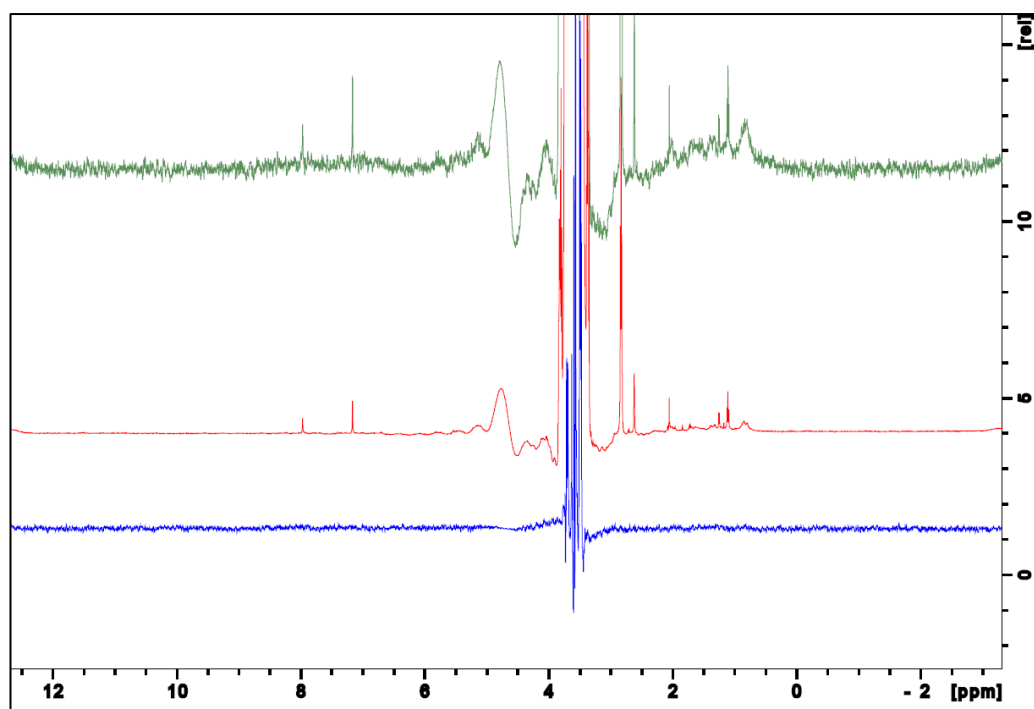


Figure 62:
Ligand 27-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 27 on-resonance control spectrum (Middle), difference spectrum (Bottom)

As can be seen, the on-resonance and off-resonance STD spectra cancel out, giving a flat difference spectrum as seen in blue, indicating unsuccessful ligand-protein binding.

3.6.2.6 Ligand 27 CPMG NMR

Finally, CPMG NMR remains in agreement with previous NMR experiments in the case of Ligand 27:

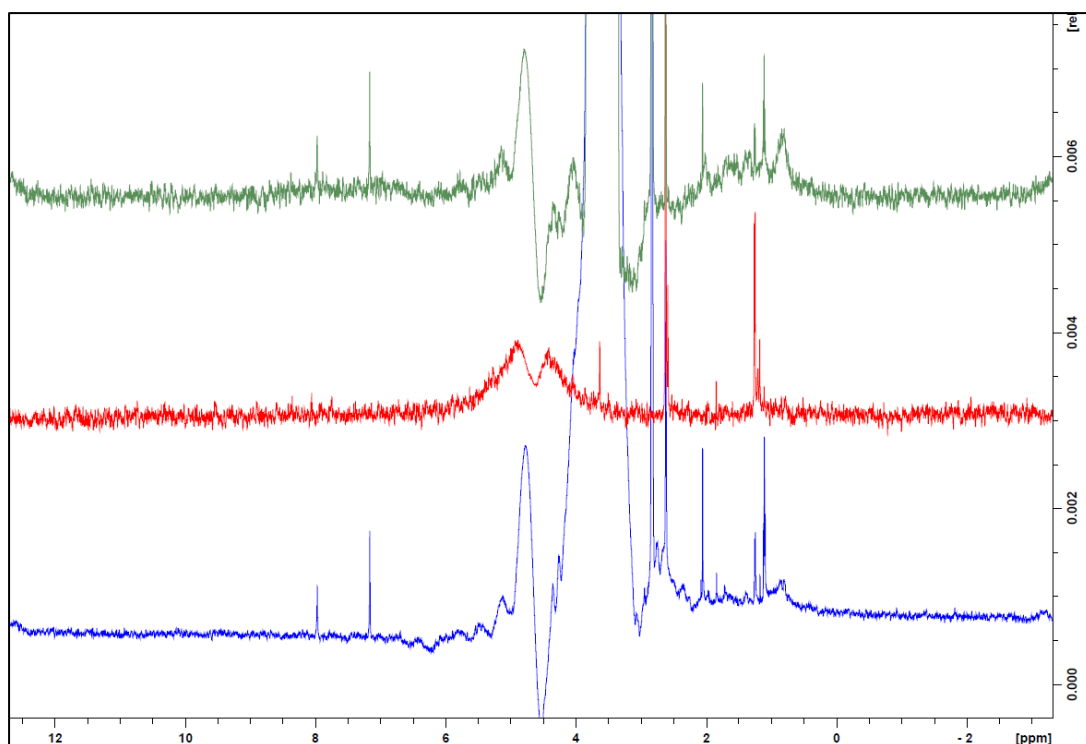


Figure 63:
Ligand 27-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 27 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)

This is determined due to a lack of signal reduction in the ligand-protein CPMG spectrum seen in blue compared to the reference spectrum in green.

To summarise: alongside Ligand 29, Ligand 1 is the only other ligand with triple NMR protein-ligand interaction confirmations. Ligand 29 was an expected hit outcome as it was the strongest hit observed in the Biochemical assays as previously shown. Ligand 1 has also shown a binding affinity to the TbrPDEB1 target protein following the NMR screen, despite showing no inhibition in the Biochemical assays. However, binding of a particular ligand occurring does not necessarily mean that ligand is also capable of target inhibition. In the case of Biochemical assays, it is likely that Ligand 1 is not a competitive enough inhibitor against the cAMP substrate and therefore cannot displace it from the binding site, which is a possible explanation for the lack of measured inhibition.

Surprisingly, Ligand 27 did not appear as a hit following the NMR techniques used, despite being identified as a hit in the Biochemical assay experiments. As previously presented, Ligand 27 proved to be an optimal ligand hit, as it forms the vital Q-Switch binding as shown by X-Ray crystallography and also has a small measured inhibition towards the TbrPDEB1 target as shown by the Biochemical assays. However, as discussed, these particular NMR experiments are strictly for weakly-binding ligands with a low binding affinity towards the target, as rapid dissociation from the bound and unbound state is required in order to transfer magnetisation. Hence, Ligand 27 not showing up as a hit in these particular experiments does not necessarily mean it is a weak binder by NMR. On the contrary, it is possible that it binds too tightly to the TbrPDEB1 target in order to satisfy the required fast dissociation, resulting in a false negative occurring. Further experimentation on Ligand 27 is required to put this hypothesis to the test.

3.6.3 Ligand 27 competitive STD NMR

An additional protein-ligand STD experiment was repeated to further determine Ligand 27 as a binder by NMR. The experiment was repeated as described in section 2.8.3 with a weak binding ligand. Hence, the appropriate samples were prepared in the case of Ligand 1. STD spectra were taken for the respective samples, as seen in *Figure 64* below:

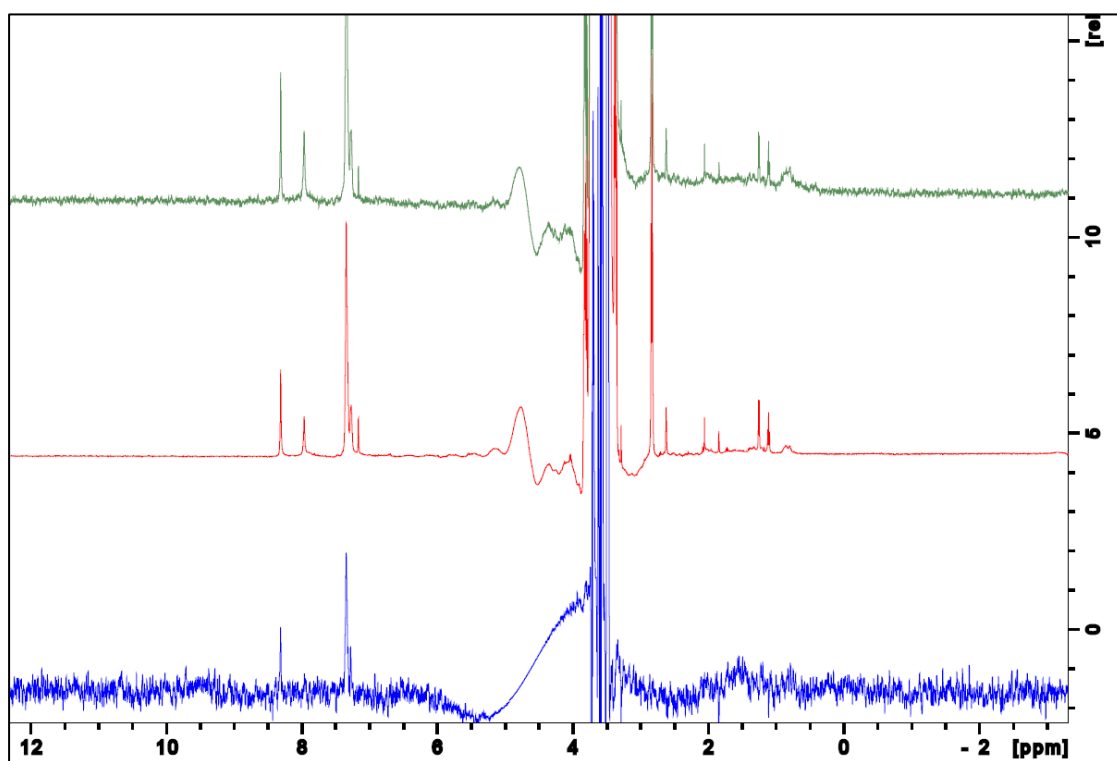


Figure 64:
Ligand 1-TbrPDEB1 STD Competitive NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 1 on-resonance control spectrum (Middle), difference spectrum (Bottom)

As can be seen, the weakly binding Ligand 1 shows binding affinity to TbrPDEB1 by STD NMR. Adding a further 0.5mM of Ligand 27 to the samples would assist in determining whether Ligand 27 is indeed a strong-binding ligand as predicted, indicated by a decrease in peak intensity as a result of Ligand 27 displacing Ligand 1 from the substrate binding pocket. This additional STD experiment is seen in *Figure 65* below:

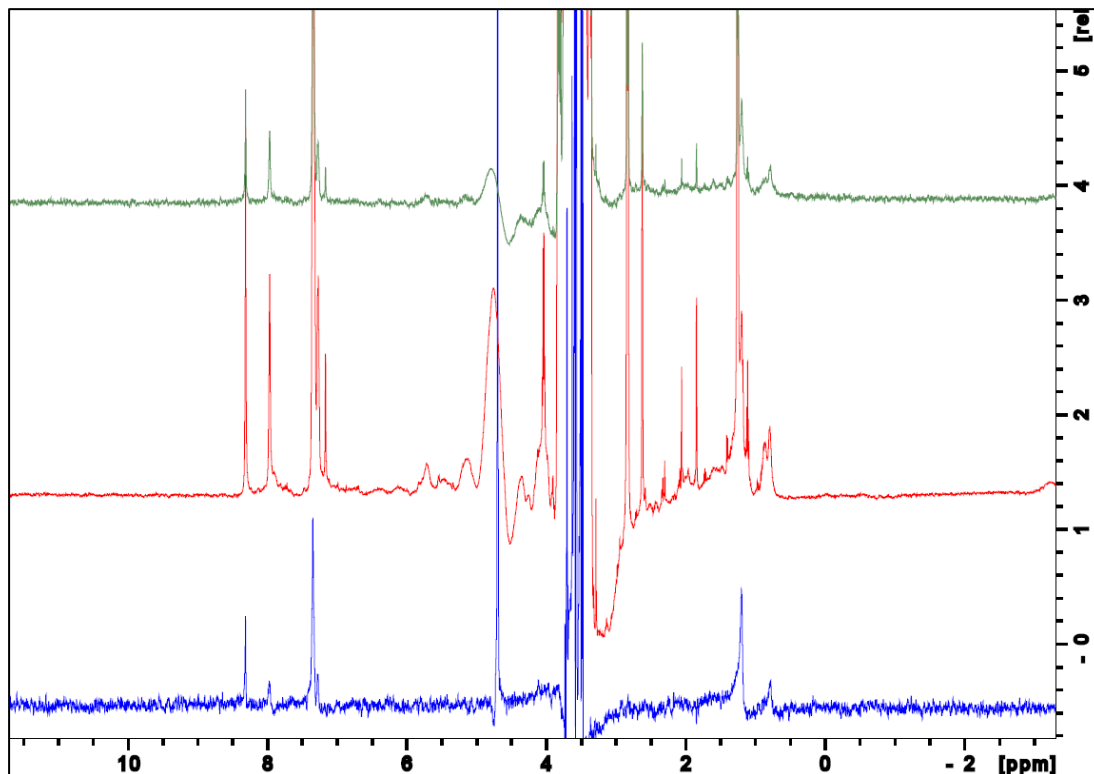


Figure 65:
Ligand 1+27-TbrPDEB1 Competitive STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 1+27 on-resonance control spectrum (Middle), difference spectrum (Bottom)

It can be determined from this figure that Ligand 27 is actually not a strong-binding ligand, as the Ligand 1 peaks remain present in the difference spectrum as seen in blue.

After further investigation, it was determined that Ligand 27 was unsuccessfully validated as a binder by NMR due to poor solubility. As previously mentioned, a lack of peaks on the control spectra in the case of Ligand 27 waterLOGSY, STD and CPMG experiments causes the difference spectra measured to be false negatives as they cannot be compared to the control. Proton 1D spectra of Ligand 27 were once again taken in DMSO and in phosphate buffer, as seen presented on the following page:

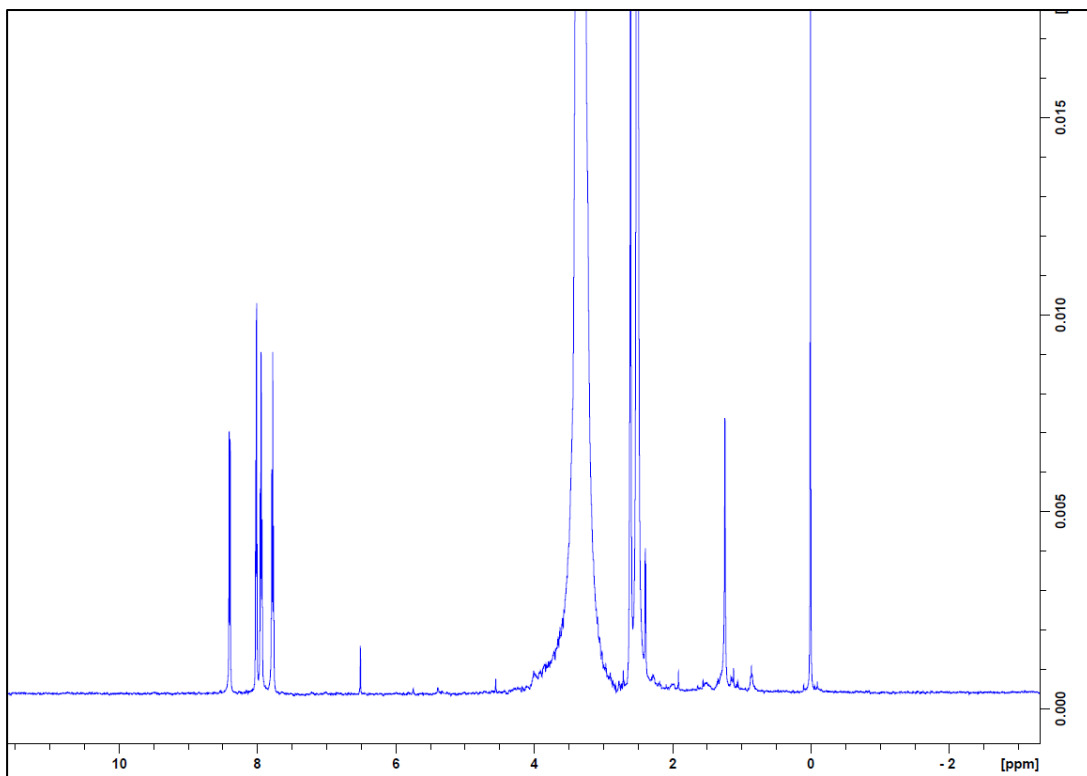


Figure 66:
Ligand 27 in DMSO Proton 1D

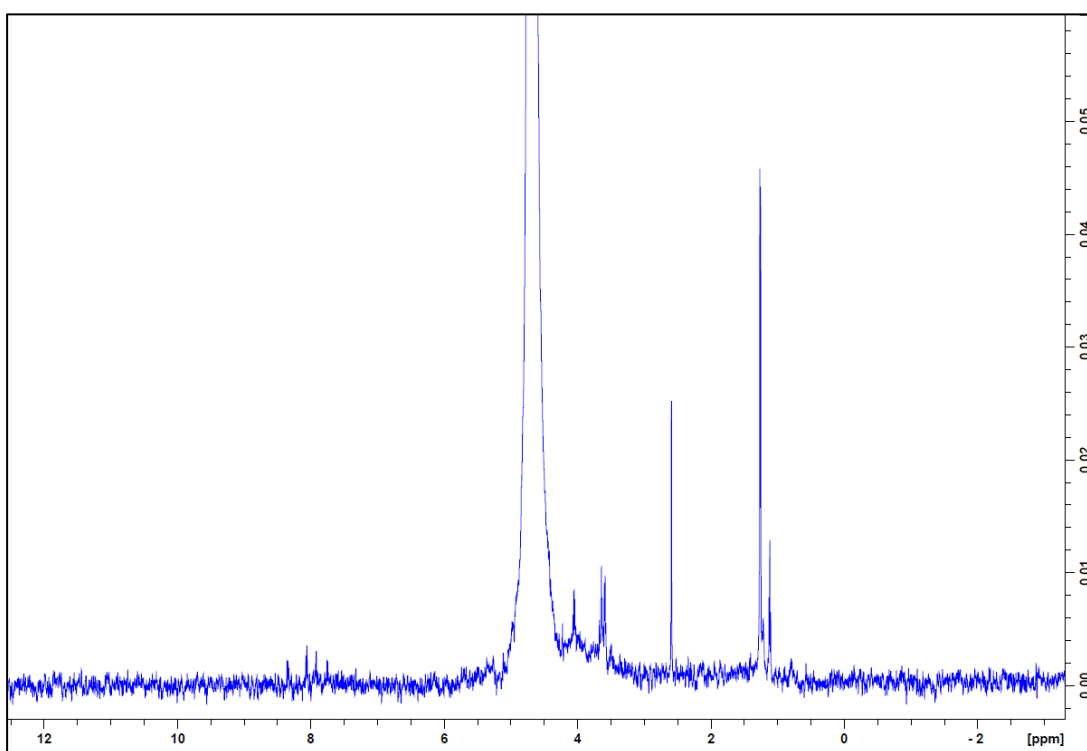


Figure 67:
Ligand 27 in phosphate buffer Proton 1D

It can be deduced from the seen 1D spectra that Ligand 27 is in fact insoluble in phosphate buffer, explaining the lack of peaks in the control spectra on the previously shown ligand-protein affinity NMR spectra. However, peaks appear to be present in the reference spectra, meaning that Ligand 27 is soluble in the presence of TbrPDEB1 protein. A lack of oxygen atoms or lone pairs on the Ligand 27 methyl-triazolo-quinazoline structure make hydrogen bonding extremely difficult, resulting in very poor solubility in water, making it a poor lead candidate compared to Ligand 29, which is evidently far more soluble due to the sulphonamide oxygen atoms and the nitrobenzene negative charge.

3.7 Substructure Similarity Search

In collaboration with Lorena Zara, a PhD student from VU University in Amsterdam, a ROCS⁷⁹ shape similarity search follow-up was performed on Ligand 29. ROCS is a powerful virtual screening method used to identify potentially active compounds by shape comparison⁷⁹. Hence, an SAR by catalogue (Structure-activity relationship) search of Ligand 29 was performed cross referenced with commercially available ligands from ZINC⁸⁰ – a database consisting of millions of purchasable compounds available for virtual screening.

10 compounds with a ROCS ComboScore of above 1.0 were identified. A ROCS ComboScore is a rating system used to rank compounds by their shape similarity to the original compound using a score value between 0 and 2. The top 3 structures (compared to Ligand 29) are presented in *Figure 68* as seen on the following page. The remaining structures are seen presented in *Appendix 68 – Appendix 74* in the Appendices.

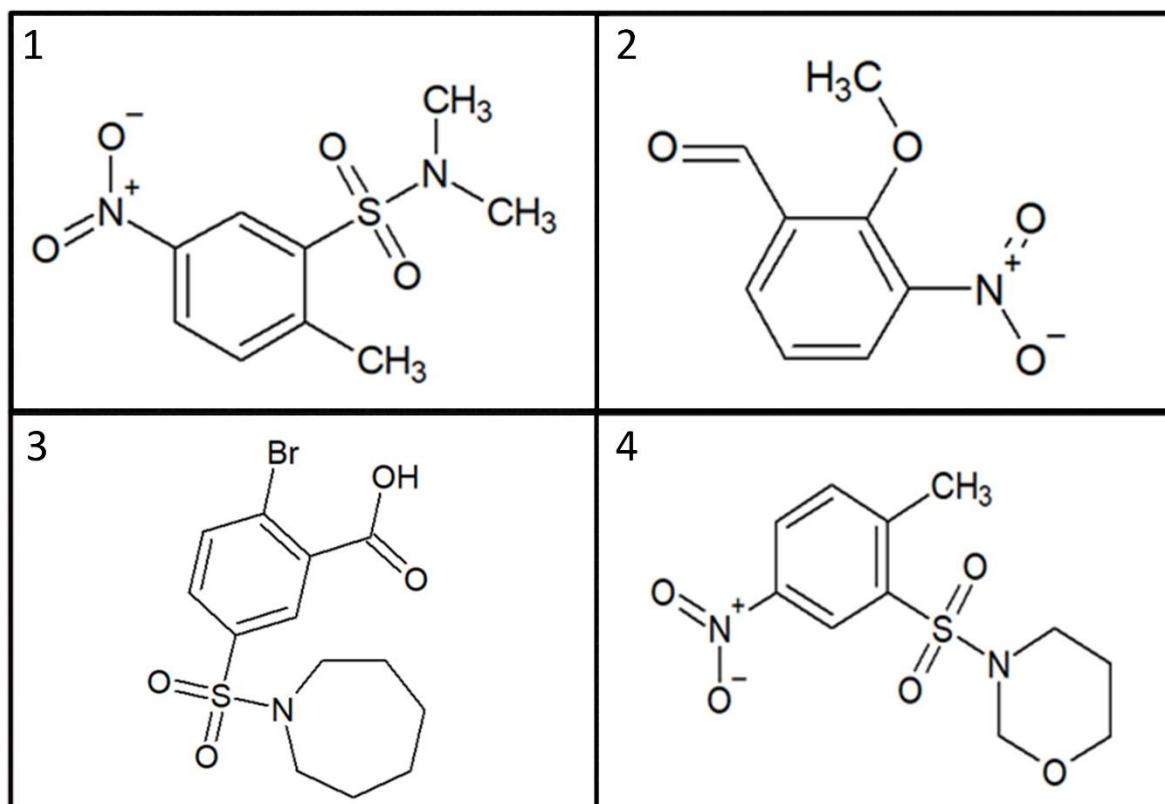


Figure 68:
 Ligand 29 (1) ROCS similarity search, VM_STL367219 (2), CB-5346663 (3),
 EN_Z2157984556 (4)

VM_STL367219 (Figure 68-2) is the closest derivative to Ligand 29 (Figure 68-1) with a ComboScore of 1.4299. CB_5346663 (Figure 68-3) and EN_Z2157984556 (Figure 68-4) also possess a ComboScore of 1.397 and 1.395, respectively. A follow-up project of interest may involve performing similar screening experiments on a set of compounds identified by this substructure search in order to observe how potency is affected by differentiation of certain side chains and functional groups. In turn this would lead to the expansion of an active TbrPDEB1 fragment library, which would meet the core aims of this particular research project.

4 Conclusion

In this research project, Ligand 29 has proven to be the most suitable ‘stand-out’ fragment ‘cherry-picked’ from a library of 31 compounds following various experiments and screening methods in order to validate it for a similarity-search follow up. Ligand 29 has been identified as a binder to TbrPDEB1 by X-Ray Crystallography and protein-ligand NMR screening, and a low potency inhibitor to TbrPDEB1 through the use of Biochemical assay experiments. Despite also possessing a small inhibition towards the drug target at a high concentration, this fragment is not deemed suitable for dose response curves or further experimentation as previously discussed. However, derivatives of the active scaffold have been identified which would be of interest in a follow-up project. Another follow-up experiment of interest could include the screening of the fragment library against hPDE4d by X-Ray Crystallography in order to compare and contrast the key differences between the binding modes and interactions of the compounds of interest against the two different targets. Additionally, repeating the X-Ray Crystallography screening of the fragment set used in this project using the XCHEM approach as previously discussed would also be of interest and useful in the optimisation of the XCHEM protocol to aid in identifying high affinity ligands using this platform.

Fragment screening has proven to be an effective means of hit identification for the discovery of inhibitors against a validated drug target. Being able to quickly assess the interactions and potency of low molecular weight fragments against a particular drug target is undoubtedly beneficial, allowing for quick identification of ideal compound scaffolds by being able to categorise and differentiate fragments from non-binders to potential leads. A clear example of this is seen in the case of a previously cited article by AR. Blazeer *et al*³³ as well as other similar scientific research papers referenced within this report discussing the research of the PDE4NPD project. Fragment-based and high-throughput screening as well as hit-to-lead follow up experiments resulted in the synthesis of a number of potent TbrPDEB1 inhibitors such as NPD-001 and NPD-008, the latter also consisting of partial selectivity towards the TbrPDEB1 target as was previously discussed throughout this report. Despite ideal potency being met, selectivity remains the biggest issue in identifying an ideal compound for the treatment of HAT.

Table 16 on the following page shows a brief comparison overview of the screening techniques used in this research project.

Table 16:
Brief overview comparison of screening techniques used

Fragment screening technique comparison	X-Ray Crystallography	NMR	Biochemical Assays
Sensitivity range	millimolar	micromolar- millimolar	nanomolar- micromolar
Binding interaction information	Yes	Yes/No	No
Binding affinity	No	Yes/No	Yes
Cost	High	Medium	Low
Number of compounds per screen	10-100	100-1000	1000-10,000

When attempting to identify novel fragments, it is of utmost importance to use as many resources and techniques as possible, as each screening method possesses certain advantages and disadvantages and indeed yield different results; Much more can be learnt about the interaction of a compound when using a wide variety of techniques. Additionally, being able to compare results from different screening methods and observing overlaps and agreements, as can be seen in the example of Ligand 29 being identified as a binder by Biochemical assays and protein-ligand NMR, is greatly beneficial and can increase confidence and validation for a particular scaffold.

The identified binder to TbrPDEB1 within this research project through the use and understanding of various screening approaches essential for drug discovery is without a doubt a success, and has led to all of the project aims to have been met. To conclude, *Figure 69* on the following page depicts a graphic overview of the findings within this research project.

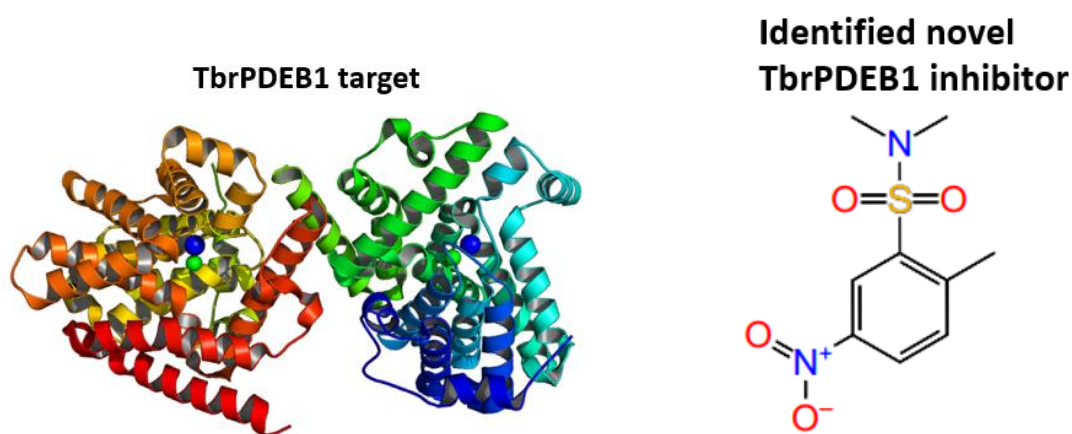
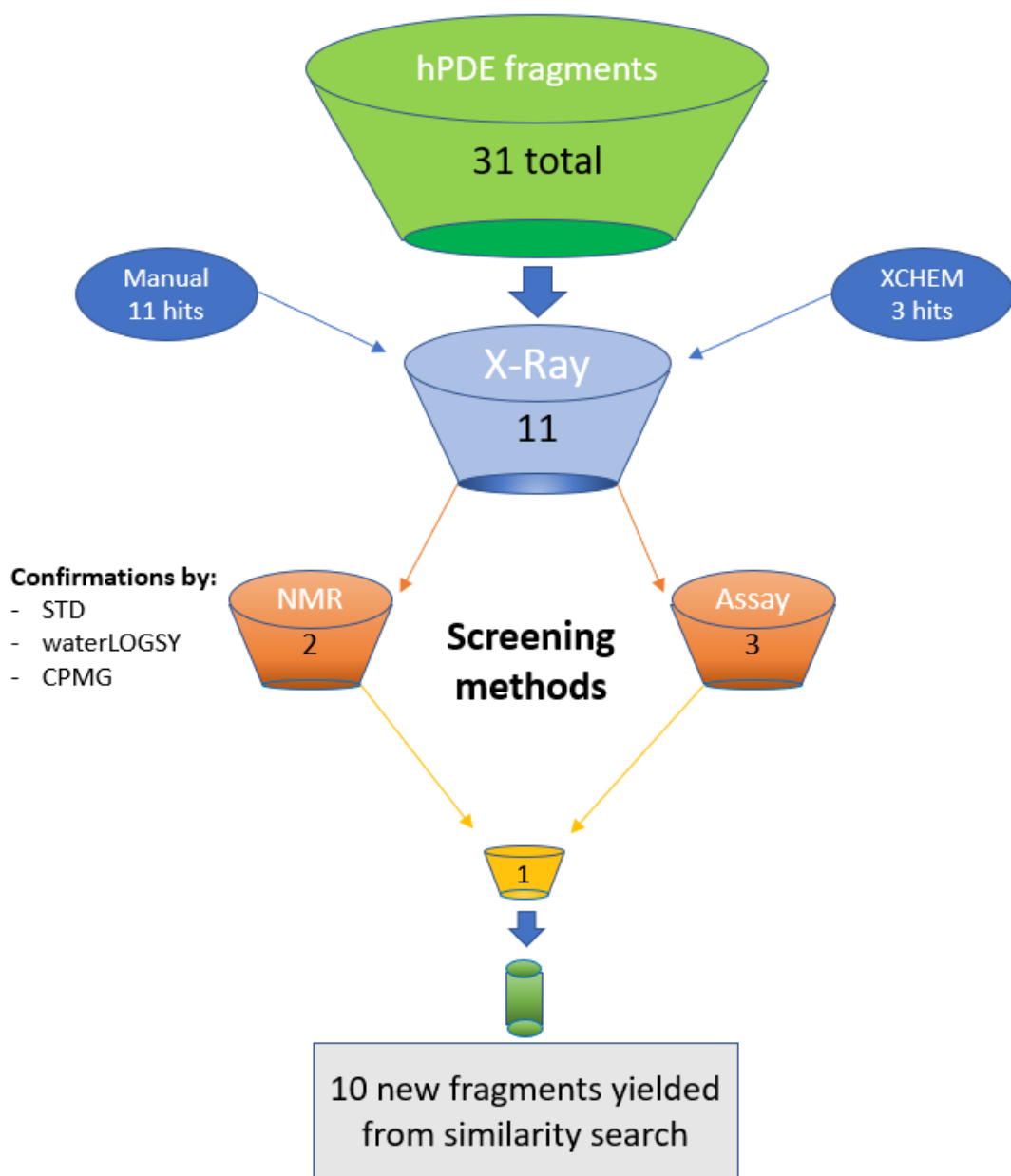


Figure 69:
Project overview

References

1. *Causes of death over 100 years.*
<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/causesofdeathover100years/2017-09-18>
2. WHO. *Neglected tropical diseases.* http://www.who.int/neglected_diseases/diseases/en/
3. WHO. *Chagas disease (American trypanosomiasis).* [http://www.who.int/en/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](http://www.who.int/en/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis))
4. D.-A. Álvarez-Hernández et al. Chagas disease: Current perspectives on a forgotten disease. *Enfermedad de Chagas: Perspectivas actuales sobre una enfermedad olvidada. Revista Medica del Hospital General de Mexico* 2018; 81(3): .
<https://www.sciencedirect.com/science/article/pii/S0185106316301123>
5. CDC. *Parasites - American Trypanosomiasis (also known as Chagas Disease).*
https://www.cdc.gov/parasites/chagas/gen_info/detailed.html
6. Parise ME, Hotez PJ, Slutsker L. Neglected parasitic infections in the United States: needs and opportunities. 2014; (): <https://www.ncbi.nlm.nih.gov/pubmed/24808243>
7. PB. Hamilton, JR. Stevens, MW. Gaunt, J. Gidley, WC. Gibson. Trypanosomes are monophyletic: evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *International Journal for Parasitology* 2004; 34(12): .
<https://www.sciencedirect.com/science/article/pii/S0020751904001894?via%3Dihub>
8. Stich A, Abel PM, Krishna S. Human African trypanosomiasis. 2002; (): .
<https://www.ncbi.nlm.nih.gov/pubmed/12142311>
9. B. Philippe, C. Giuliano, J. Vincent, P. Gerardo. Human African trypanosomiasis. 2017; (): <https://www.ncbi.nlm.nih.gov/pubmed/28673422>
- 10 - Simarro PP, Cecchi G, Paone M, Franco JR, Diarra A, Ruiz JA et al. The atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. *Int J Health Geogr* 2010; 9:57

11. WHO. *Human African trypanosomiasis - The parasite*.
http://www.who.int/trypanosomiasis_african/parasite/en/
12. WHO. *Trypanosomiasis, human African (sleeping sickness)*.
[http://www.who.int/en/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](http://www.who.int/en/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness))
13. S. Aksoy, P. Buscher, M. Lehane, P. Solano, JVD. Abbeele. Human African trypanosomiasis control: Achievements and challenges. 2017; (): .
<http://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0005454>
14. Nigeria Galleria. *Sleeping sickness*. <https://www.nigeriagalleria.com/Community-Health/Sleeping-Sickness.html>
15. D. Steverding. The history of African trypanosomiasis. 2008; (): .
<https://www.ncbi.nlm.nih.gov/pubmed/18275594>
16. WHO. *WHO report on global surveillance of epidemic-prone infectious diseases*.
http://www.who.int/csr/resources/publications/surveillance/WHO_CDS_CSR_ISR_2000_1/en/
17. S. Aksoy, P. Buscher, M. Lehane, P. Solano, VD. Abbeele. Human African trypanosomiasis control: Achievements and challenges. 2017; (): .
<https://www.ncbi.nlm.nih.gov/pubmed/28426685>
18. PG. Bray, MP. Barrett. SA. Ward, HP. de Koning. Pentamidine uptake and resistance in pathogenic protozoa: past, present and future. 2003; (): .
<https://www.ncbi.nlm.nih.gov/pubmed/12763430?dopt=Abstract>
19. V. Delespaux, HP. de Koning. Drugs and drug resistance in African trypanosomiasis. 2007; (): .
<https://www.ncbi.nlm.nih.gov/pubmed/17409013?dopt=Abstract>

20. C. Burri, R. Brun. Eflornithine for the treatment of human African trypanosomiasis. *Parasitology Research* 2003; 90(1): .
<https://link.springer.com/article/10.1007%2Fs00436-002-0766-5>
21. GL. Card et al. Structural Basis for the Activity of Drugs that Inhibit Phosphodiesterases. *Cell Press* 2004; 12(12): .
<https://www.sciencedirect.com/science/article/pii/S0969212604003727>
22. F. Svensson, A. Bender, D. Bailey. Fragment-Based Drug Discovery of Phosphodiesterase Inhibitors. *Journal of Medicinal Chemistry* 2018; 61(4): .
<https://pubs.acs.org/doi/abs/10.1021/acs.jmedchem.7b00404>
23. JC. Leffingwell. *G-Protein Coupled Receptors*. <http://www.leffingwell.com/olfact3.htm>
24. S. Laxman, JA. Beavo. Cyclic Nucleotide Signaling Mechanisms in Trypanosomes: Possible Targets for Therapeutic Agents. *Molecular Interventions* 2007; 7(4): .
<https://triggered.clockss.org/ServeContent?url=http%3A%2F%2Fmolinterv.aspetjournals.org%2Fcontent%2F7%2F4%2F203.full>
25. YSJ. Ho, LM. Burden, JH. Hurley. Structure of the GAF domain, a ubiquitous signalling motif and a new class of cyclic GMP receptor. *EMBO J* 2000; 19(20): .
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC314001/>
26. RL. Williams. FYVE Domains. *Handbook of Cell Signalling (Second Edition)* 2010; ();
<https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/fyve-domain>
27. R. Diaz-Benjumea, S. Laxman, TR. Hinds, JA. Beavo, A. Rascon. Characterization of a novel cAMP-binding, cAMP-specific cyclic nucleotide phosphodiesterase (TcrPDEB1) from *Trypanosoma cruzi*. *Biochem J* 2006; 15(2): .
<https://www.ncbi.nlm.nih.gov/pubmed/16776650/>
28. *Handbook of Experimental Pharmacology - Phosphodiesterases as Drug Targets*. : ; 2011.

29. KW. Gong et al. cAMP-specific phosphodiesterase TbpDE1 is not essential in *Trypanosoma brucei* in culture or during midgut infection of tsetse flies. *Molecular and Biochemical Parasitology* 2001; 116(2): .
<https://www.sciencedirect.com/science/article/pii/S0166685101003152?via%3Dihub>
30. R. Zoraghi, T. Seebeck. The cAMP-specific phosphodiesterase TbpDE2C is an essential enzyme in bloodstream form *Trypanosoma brucei*. *PNAS* 2002; 99(7): .
<http://www.pnas.org/content/99/7/4343>
31. HP. de Koning, MK. Gould, GJ. Sterk, H. Tenor, S. Kunz, E. Luginbuehl, T. Seebeck. Pharmacological validation of *Trypanosoma brucei* phosphodiesterases as novel drug targets. *J Infect Dis* 2012; 15;206((2)): . <https://www.ncbi.nlm.nih.gov/pubmed/22291195>
32. T. Seebeck, GJ. Sterk, H. Ke. Phosphodiesterase inhibitors as a new generation of antiprotozoan drugs: exploiting the benefit of enzymes that are highly conserved between host and parasite. *Future Med Chem* 2011; 3(10): .
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3164761/>
33. AR. Blaazer, AK Singh et al. Targeting a Subpocket in *Trypanosoma brucei* Phosphodiesterase B1 (TbrPDEB1) Enables the Structure-Based Discovery of Selective Inhibitors with Trypanocidal Activity.. *J Med Chem* 2018; 10;61(9): .
<https://www.ncbi.nlm.nih.gov/pubmed/29672041>
34. M. Oberholzer, G. Marti, M. Baresic, S. Kunz, A. Hemphill, T. Seebeck. The *Trypanosoma brucei* cAMP phosphodiesterases TbrPDEB1 and TbrPDEB2: flagellar enzymes that are essential for parasite virulence. *FASEB* 2006; 21(3): .
<https://www.fasebj.org/doi/pdf/10.1096/fj.06-6818com>
35. S. Kunz, E. Luginbuehl, T. Seebeck. Gene conversion transfers the GAF-A domain of phosphodiesterase TbrPDEB1 to one allele of TbrPDEB2 of *Trypanosoma brucei*.. *PLoS Negl Trop Dis*. 2009; 9;3(6): . <https://www.ncbi.nlm.nih.gov/m/pubmed/19513125/>

36. H. Wang et al. Biological and Structural Characterization of Trypanosoma cruzi Phosphodiesterase C and Implications for Design of Parasite Selective Inhibitors. *J Biol Chem* 2012; 6;287(15): . <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3320927/>
37. PG. Mckean. Coordination of cell cycle and cytokinesis in Trypanosoma brucei.. *Curr Opin Microbiol* 2003; 6(6): . <https://www.ncbi.nlm.nih.gov/pubmed/14662356>
38. Cambridge MedChem Consulting. *Fragment-Based Screening*. https://www.cambridgemedchemconsulting.com/resources/hit_identification/fragment_based_screening.html
39. GS. Sittampalam, SD. Kahl, WP. Janzen. High-throughput screening: advances in assay technologies. *Current Opinion in Chemical Biology* 1997; 1(3): . <https://www.sciencedirect.com/science/article/pii/S1367593197800786>
40. AD. Baxter, PM. Lockey. *Hit to lead and lead to candidate optimisation using multi-parametric principles*. <https://www.ddw-online.com/winter-2001/p148585-%EF%BF%BDhit%EF%BF%BD-to-%EF%BF%BDlead%EF%BF%BD-and-%EF%BF%BDlead%EF%BF%BD-to-%EF%BF%BDcandidate%EF%BF%BD-optimisation-using-multi-parametric-principles.html>
41. JP. Hughes, S. Rees, SB. Kalindjian, KL. Philpott. Principles of early drug discovery. *Br J Pharmacol* 2011; 162(6): . <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3058157/>
42. A. Ciulli, C. Abell. Fragment-based approaches to enzyme inhibition. *Curr Opin Biotechnol* 2007; 18(6): . <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4441723/>
43. GM. Keseru, GM. Makara. Hit discovery and hit-to-lead approaches. *Drug Discovery Today* 2006; 11(15-16): . <https://www.sciencedirect.com/science/article/pii/S1359644606002285>
44. A. Elumalai, MC. Eswariah. A review on combinatorial chemistry in drug discovery. *International Journal of Pharmaceutical Research and Analysis* 2012; 2(2): . <http://ijpsr.com/bft-article/combinatorial-chemistry-a-review/?view=fulltext>

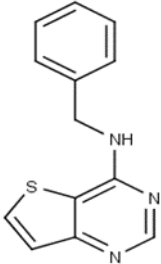
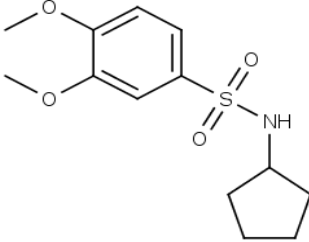
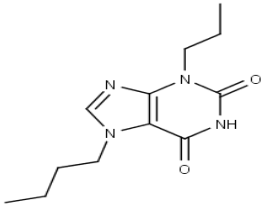
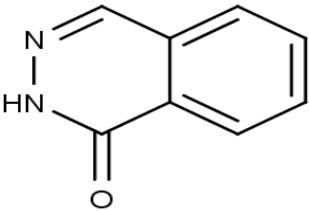
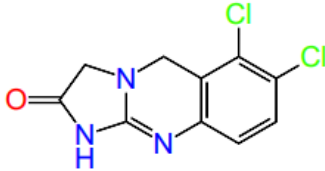
56. JF. Riordan. The Role of Metals in Enzyme Activity. *Annals of Clinical and Laboratory Science* 1977; 7(2): . <http://www.annclinlabsci.org/content/7/2/119.full.pdf>
57. NJ. Pace, E. Weerapana. Zinc-Binding Cysteines: Diverse Functions and Structural Motifs. *Biomolecules* 2014; 4(2): .
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4101490/>
58. KA. McCall, CC. Huang, CA. Fierke. Function and Mechanism of Zinc Metalloenzymes. *The Journal of Nutrition* 2000; 130(5): .
<https://academic.oup.com/jn/article/130/5/1437S/4686409#111924025>
59. S. Maric, SM. Donnelly, MW. Robinson et al. The M17 Leucine Aminopeptidase of the Malaria Parasite *Plasmodium falciparum*: Importance of Active Site Metal Ions in the Binding of Substrates and Inhibitors. *Biochemistry* 2009; 48(23): .
<https://pubs.acs.org/doi/pdf/10.1021/bi9003638>
60. V. WB. The role of magnesium in nucleic-acid and protein metabolism. *Magnesium* 1988; 7(5-6): . <https://www.ncbi.nlm.nih.gov/pubmed/2472534>
61. F. Ehrenmann, Z. Ouaray, MP. Lefranc. *IMGT classes of the 20 common amino acids*.
http://www.imgt.org/IMGTeducation/Aide-memoire/_UK/aminoacids/IMGTclasses.html
62. KM. Orrling, C. Jansen, XL. Vu et al. Catechol Pyrazolinones as Trypanocidals: Fragment-Based Design, Synthesis, and Pharmacological Evaluation of Nanomolar Inhibitors of Trypanosomal Phosphodiesterase B1. *Journal of Medicinal Chemistry* 2012; 55(20): . <https://pubs.acs.org/doi/pdf/10.1021/jm301059b>
63. GL. Card, BP. England, Y. Suzuki et al. Structural Basis for the Activity of Drugs that Inhibit Phosphodiesterases. *Structure* 2004; 12(12): .
<https://www.sciencedirect.com/science/article/pii/S0969212604003727>
64. MM. Flocco et al. Computer-Assisted Drug Design. *Comprehensive Medicinal Chemistry* 2007; (); . <https://www.sciencedirect.com/topics/neuroscience/ibmx>


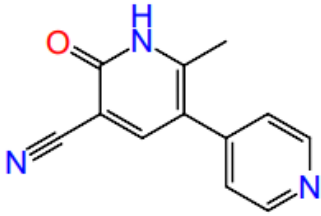
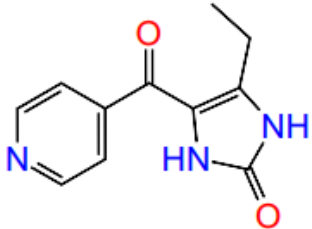

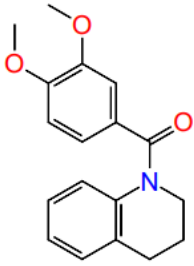
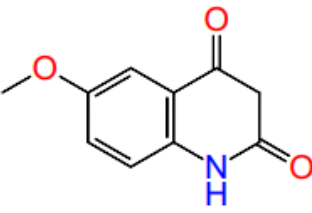
65. KYJ. Zhang, GL. Card, Y. Suzuki et al. A Glutamine Switch Mechanism for Nucleotide Selectivity by Phosphodiesterases. *Molecular Cell* 2004; 15(2): .
<https://www.sciencedirect.com/science/article/pii/S1097276504004083>
66. Collaborative Computational Project No. 4 Software for Macromolecular X-Ray Crystallography. <http://www.ccp4.ac.uk/ccp4i2/>
67. Coot. <https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/>
- 68.. ACD/ChemSketch for Academic and Personal Use.
<https://www.acdlabs.com/resources/freeware/chemsketch/>
69. SnapGene. <http://www.snapgene.com/>
70. FPLC Servicing. https://holmesanalytical.co.uk/fplc-servicing?gclid=CjwKCAjw4uXaBRACeIwAuAUz8EjoUecO_smxVZMYQBjC6wCtPdAIQbj-VhP0FIJtEE9SPLGjApvv0hoCnkQQAyD_BwE
71. Topspin. https://www.bruker.com/products/mr/nmr/nmr-software/nmr-software/topspin/overview.html?gclid=CjwKCAjw4uXaBRACeIwAuAUz8JoqZ86mShaU2K1qRl3LJ4Kbtq8sTuvvrdEurkFKCGNrZPcmUGt6tRoC5xwQAvD_BwE
72. ACD/NMR Predictors. https://www.acdlabs.com/products/adh/nmr/nmr_pred/
73. R. Huang, A. Bonnichon, TDW. Claridge, IKH. Leung. Protein-ligand binding affinity determination by the waterLOGSY method: An optimised approach considering ligand rebinding. *Scientific Reports* 2017; (): . <https://www.nature.com/articles/srep43727>
74. L. Fielding. NMR methods for the determination of protein-ligand dissociation constants. *Curr Top Med Chem* 2003; 3(1): .
<https://www.ncbi.nlm.nih.gov/pubmed/12577990>

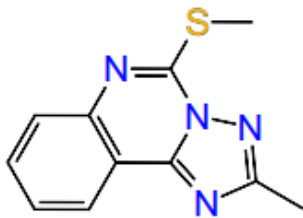
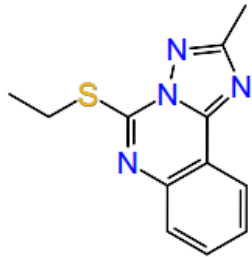
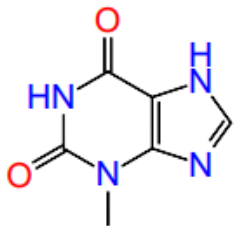
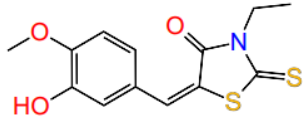
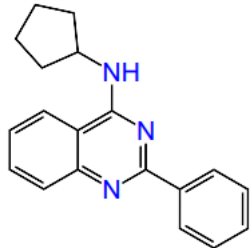
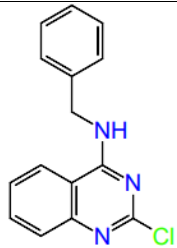
75. A. Viagas, J. Manso, FL. Nobrega, EJ. Cabrita. Saturation-Transfer Difference (STD) NMR: A Simple and Fast Method for Ligand Screening and Characterization of Protein Binding. *Journal of Chemical Education* 2011; 88(7):
<https://pubs.acs.org/doi/pdf/10.1021/ed101169t>
76. IR. Kleckner, MP. Foster. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim Biophys Acta* 2012; 1814(8): .
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3061256/>
77. LP. McIntosh. CPMG. *Encyclopedia of Biophysics* 2013; (): .
https://link.springer.com/referenceworkentry/10.1007%2F978-3-642-16712-6_320
78. Diamond Light Source. <https://www.diamond.ac.uk/Home.html>
79. ROCS. <https://www.eyesopen.com/rocs>
80. ZINC15. <http://zinc15.docking.org/>
81. C. Roumestand, D. Canet. Extending the Excitation Sculpting Concept for Selective Excitation. *Journal of Magnetic Resonance* 2000; 147(2): .
<https://www.sciencedirect.com/science/article/pii/S1090780700922067>
82. MR. Gryk, J. Vyas, MW. Maciejewski. Biomolecular NMR Data Analysis. *Prog Nucl Magn Reson Spectrosc* 2010; 56(4): .
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2899495/>
83. Teodor Parella. *NMR Building Blocks WATERGATE: 3-9-19 Element*.
<http://triton.iqfr.csic.es/guide/eNMR/eNMRblock/watergate.html>
84. Teodor Parella. *Selective Excitation BURP Pulses*.
<http://triton.iqfr.csic.es/guide/eNMR/eNMRselex/burp.html>
85. Isoelectric point calculator. <http://isoelectric.org>

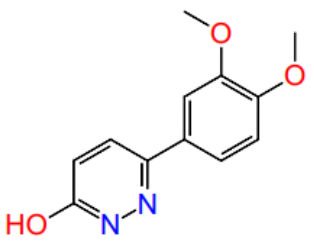
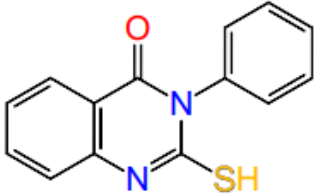
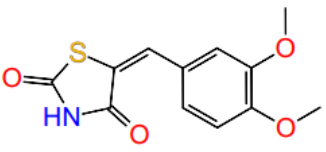
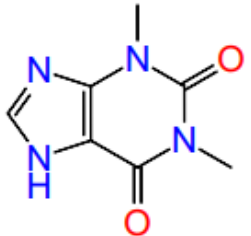
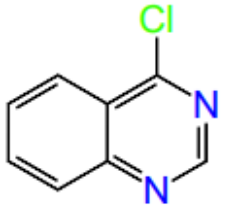
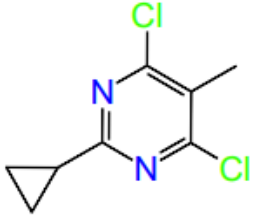
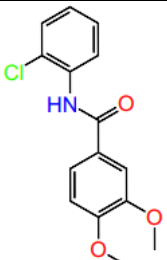
Appendices

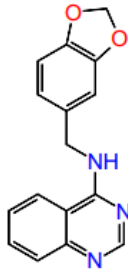
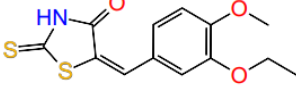
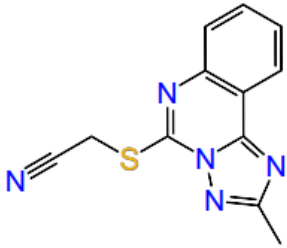
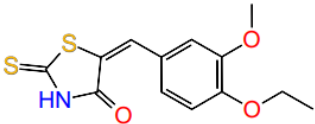
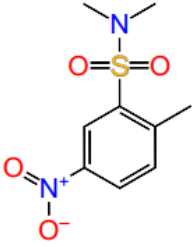
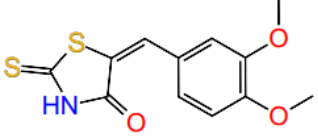
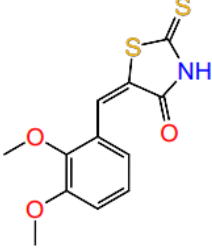
Appendix 1 – List of fragment library (tabulated)

Ligand #	Compound (IOTA ID and name)	Molecular weight	Molecular formula	Chemical structure
1	Z196118054 N-benzylthieno[3,2-d]pyrimidin-4-amine	241.31154	C ₁₃ H ₁₁ N ₃ S	
2	Z45537472 N-cyclopentyl-3,4-dimethoxybenzene-1-sulfonamide	285.35926	C ₁₃ H ₁₉ NO ₄ S	
3	Z56947553 7-butyl-3-propyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione	250.29692	C ₁₂ H ₁₈ N ₄ O ₂	
4	Z203045280 1,2-dihydrophthalazin-1-one	146.14604	C ₈ H ₆ N ₂ O	
5	Z2239077885 6,7-dichloro-1H,2H,3H,5H-imidazolidino[2,1-b]quinazolin-2-one hydrochloride	292.54902	C ₁₀ H ₈ Cl ₂ N ₃ O	

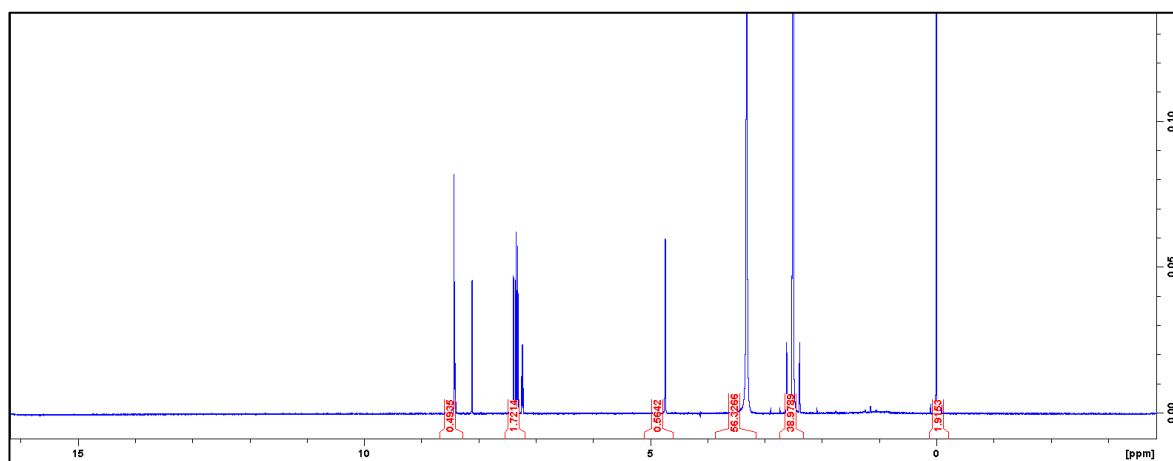
6	Z1945707486 3-propyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione	194.1906	C8H10N4O2	
7	Z1522568219 6-methyl-2-oxo-5-(pyridin-4-yl)-1,2-dihydropyridine-3-carbonitrile	211.21936	C12H9N3O	
8	Z2306534716 4-ethyl-5-(pyridine-4-carbonyl)-2,3-dihydro-1H-imidazol-2-one	217.22394	C11H11N3O2	
9	Z2936508439 3-ethyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione	180.16402	C7H8N4O2	
10	Z27693264 1-(3,4-dimethoxybenzoyl)-1,2,3,4-tetrahydroquinoline	297.34836	C18H19NO3	
11	Z1416282875 6-methoxy-1,2,3,4-tetrahydroquinoline-2,4-dione	191.18336	C10H9NO3	

12	Z226974116 2-methyl-5-(methylsulfanyl)-[1,2,4]triazolo[1,5-c]quinazoline	230.2889	C ₁₁ H ₁₀ N ₄ S	
13	Z226974122 5-(ethylsulfanyl)-2-methyl-[1,2,4]triazolo[1,5-c]quinazoline	244.31548	C ₁₂ H ₁₂ N ₄ S	
14	Z1741977122 3-methyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione	166.13744	C ₆ H ₆ N ₄ O ₂	
15	Z44297214 3-ethyl-5-[(3-hydroxy-4-methoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one	295.37722	C ₁₃ H ₁₃ N ₃ O ₃ S ₂	
16	Z31239904 N-cyclopentyl-2-phenylquinazolin-4-amine	289.37426	C ₁₉ H ₁₉ N ₃	
17	Z57389766 N-benzyl-2-chloroquinazolin-4-amine	269.72888	C ₁₅ H ₁₂ ClN ₃	

18	Z1954801165 6-(3,4-dimethoxyphenyl)pyridazin-3-ol	232.23528	C ₁₂ H ₁₂ N ₂ O ₃	
19	Z55663905 3-phenyl-2-sulfanyl-3,4-dihydroquinazolin-4-one	254.307	C ₁₄ H ₁₀ N ₂ O ₂ S	
20	Z44299539 5-[(3,4-dimethoxyphenyl)methylidene]-1,3-thiazolidine-2,4-dione	265.28504	C ₁₂ H ₁₁ N ₁ O ₄ S	
21	Z271004650 1,3-dimethyl-2,3,6,7-tetrahydro-1H-purine-2,6-dione	180.16402	C ₇ H ₈ N ₄ O ₂	
22	Z235340051 4-chloroquinazoline	164.5917	C ₈ H ₅ ClN ₂	
23	Z1696844447 4,6-dichloro-2-cyclopropyl-5-methylpyrimidine	203.06852	C ₈ H ₈ Cl ₂ N ₂	
24	Z27812338 N-(2-chlorophenyl)-3,4-dimethoxybenzamide	291.72956	C ₁₅ H ₁₄ ClNO ₃	

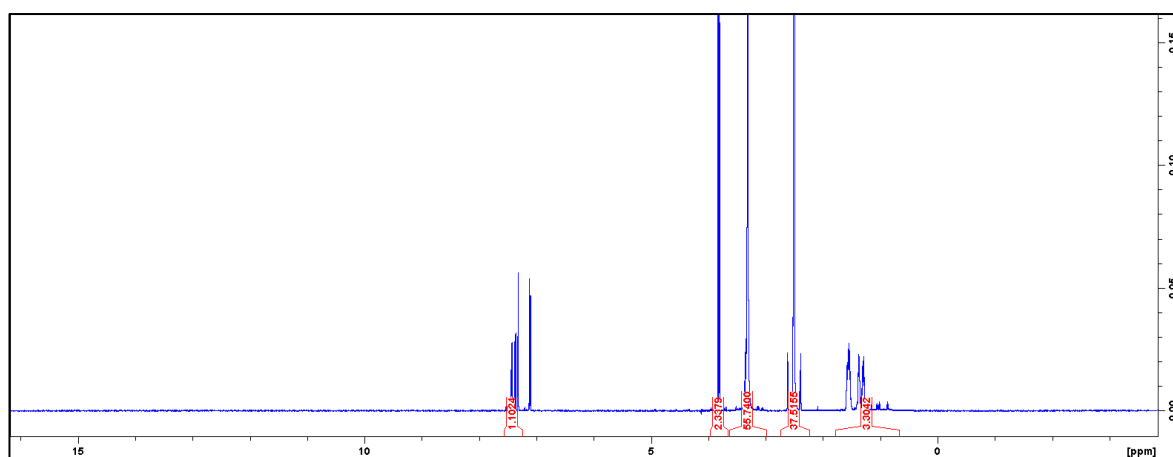
25	Z126964124 N-[(2H-1,3-benzodioxol-5-yl)methyl]quinazolin-4-amine	279.29332	C16H13N3O2	
26	Z44301662 5-[(3-ethoxy-4-methoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one	295.37722	C13H13NO3S2	
27	Z220370518 2-({2-methyl-[1,2,4]triazolo[1,5-c]quinazolin-5-yl}sulfanyl)acetonitrile	255.29836	C12H9N5S	
28	Z44301505 5-[(4-ethoxy-3-methoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one	295.37722	C13H13NO3S2	
29	Z45527541 N,N,2-trimethyl-5-nitrobenzene-1-sulfonamide	244.26758	C9H12N2O4S	
30	Z44301478 5-[(3,4-dimethoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one	281.35064	C12H11NO3S2	
31	Z44301473 5-[(2,3-dimethoxyphenyl)methylidene]-2-sulfanylidene-1,3-thiazolidin-4-one	281.35064	C12H11NO3S2	

Appendix 2 – Ligand 1 Proton QC



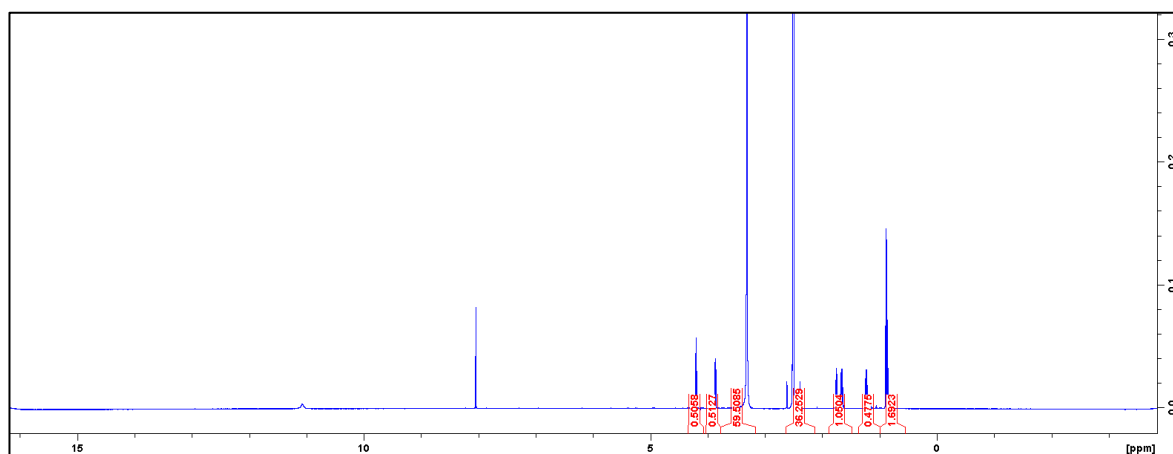
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Triplet	DMSOD6
3.4	Singlet	Impurity
7.3	Multiplet	Aromatic CH
4.9	Doublet	CH ₂
8.5	Doublet	Aromatic CH
Overlapped 8.5	Singlet	NH

Appendix 3 – Ligand 2 Proton QC



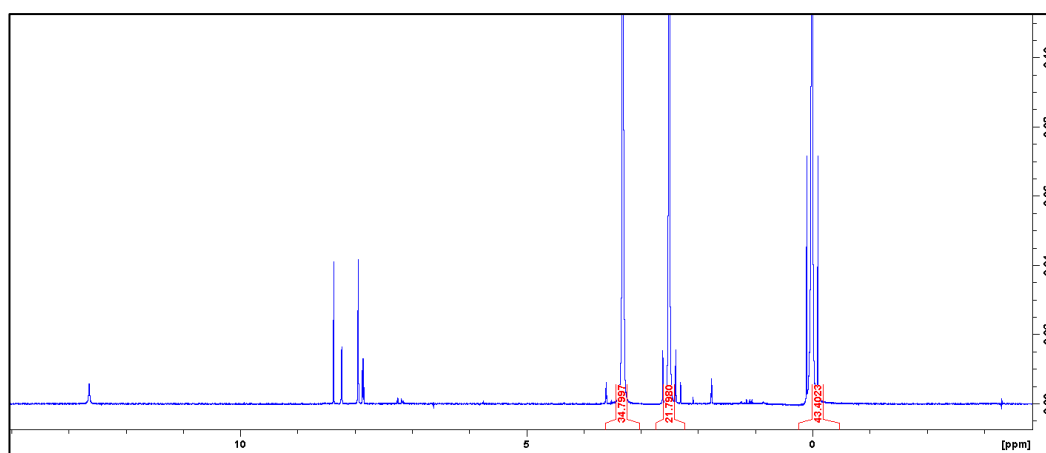
Shift	Splitting	Corresponds to
1.5	Multiplet	CH ₂ R ₂
7.5	Multiplet	Aromatic CH
2.5	Triplet	DMSOD6
3.4	Singlet	Impurity
3.5	Singlet	CH
3.9	Singlet	CH ₃
4.9	Singlet (small)	NH

Appendix 4 – Ligand 3 Proton QC



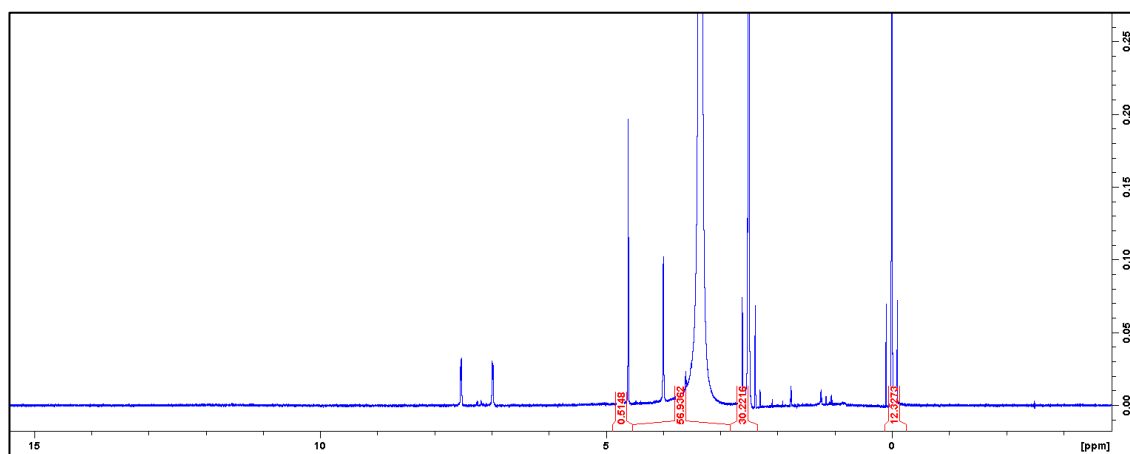
Shift	Splitting	Corresponds to
0.8	Triplet	CH3
1.23	Pentet	CH2
1.75	Multiplet	CH2
2.4	Triplet	DMSOD6
3.4	Singlet	Impurity
3.9	Doublet	CH2
4.2	Doublet	CH2
8.0	Singlet	CH
11.0	Singlet (small)	NH

Appendix 5 – Ligand 4 Proton QC



Shift	Splitting	Corresponds to
0	Triplet	TMS / impurity
2.4	Triplet	DMSOD6
3.4	Singlet	Impurity
7.8	Doublet	CH
8.1	Doublet	CH
12.5	Singlet	NH

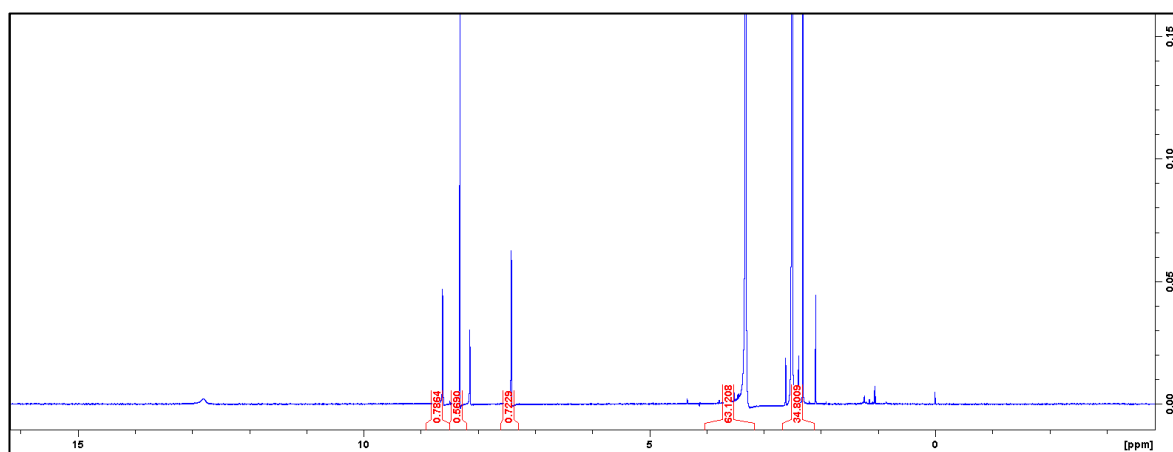
Appendix 6 – Ligand 5 Proton QC



Shift	Splitting	Corresponds to
0	Triplet	TMS / impurity
2.4	Triplet	DMSOD6
3.4	Singlet	Impurity
4.6	Singlet	CH ₂
4.0	Singlet	CH ₂
7.1	Doublet (roofing)	CH
7.3	Doublet (roofing)	CH

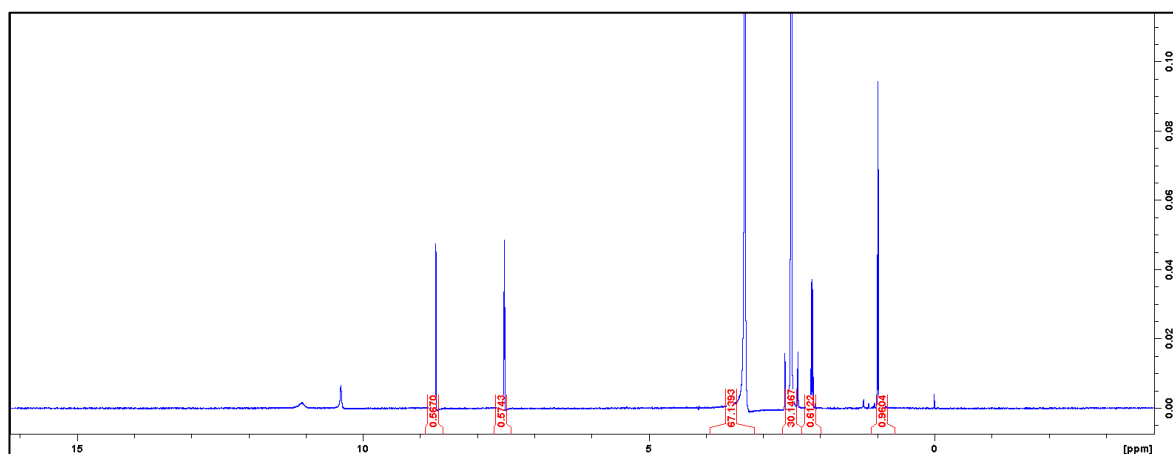
NH Peak not visible. Signal too weak?

Appendix 7 – Ligand 7 Proton QC



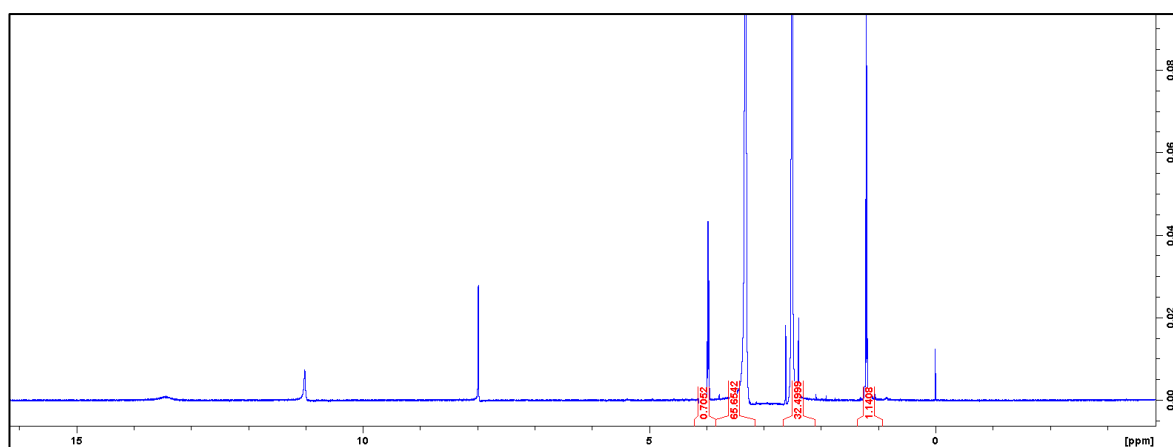
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Triplet	DMSOD6
2.3	Singlet	CH ₃
3.4	Singlet	Impurity
7.5	Doublet	CH
8.6	Singlet	CH
8.4	Doublet	CH
12.8	Singlet	NH

Appendix 8 – Ligand 8 Proton QC



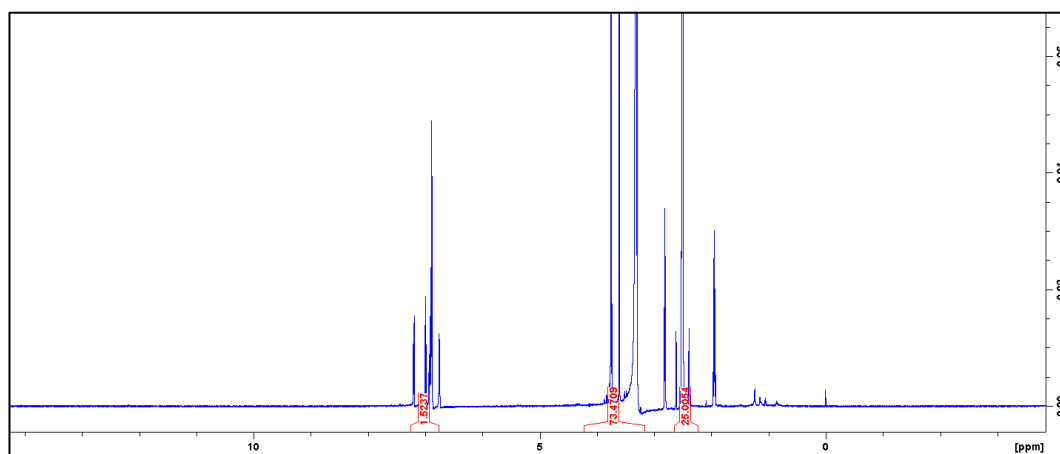
Shift	Splitting	Corresponds to
0	Singlet	TMS
1.0	Triplet	CH3
2.2-2.4	Quartet	CH2
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
7.5	Doublet	CH
8.7	Doublet	CH
10.2	Singlet	NH
11.1	Singlet	NH

Appendix 9 – Ligand 9 Proton QC



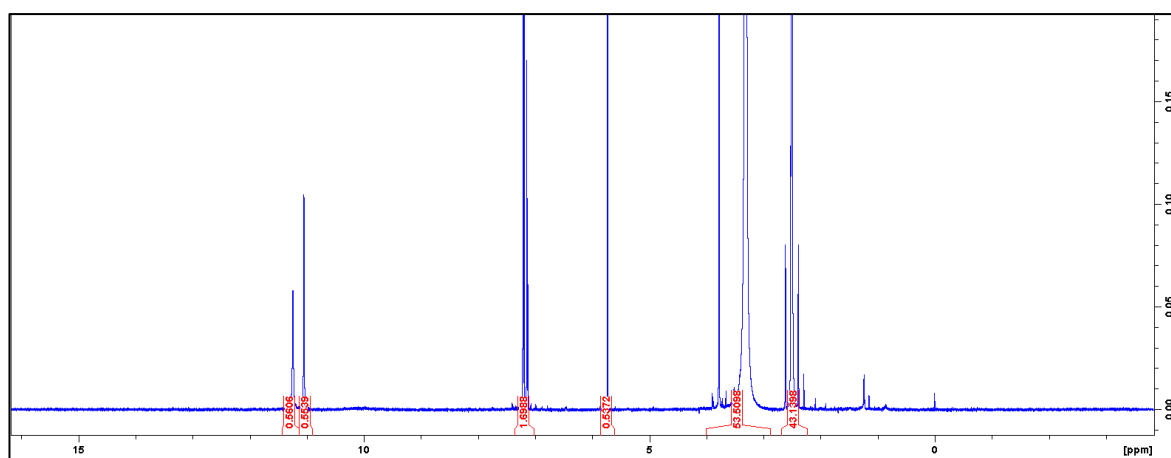
Shift	Splitting	Corresponds to
0	Singlet	TMS
1.2	Triplet	CH3
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
4.0	Quartet	CH2
8.0	Singlet	CH
11.0	Singlet	NH
13.5	Singlet	NH

Appendix 10 – Ligand 10 Proton QC



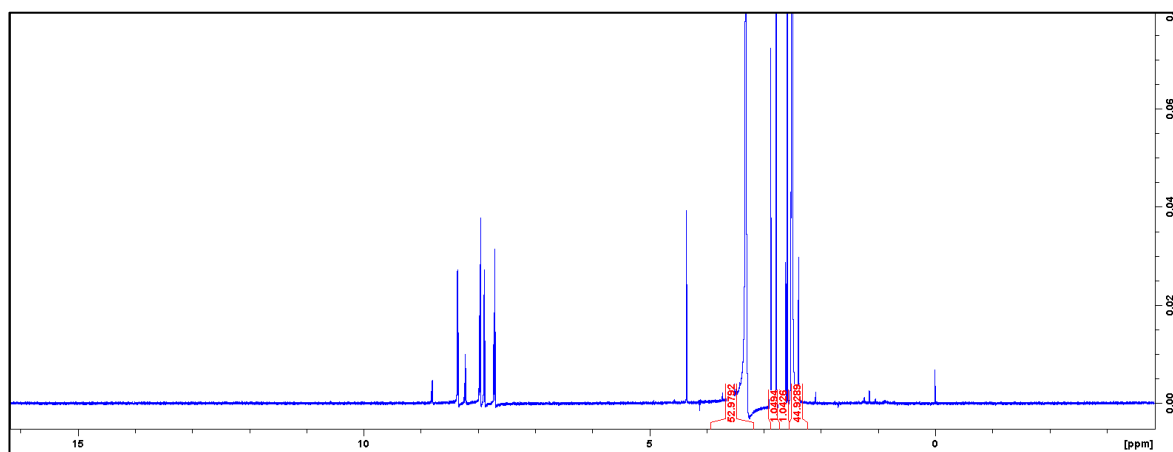
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.0	Doublet	CH2
2.4	Singlet	DMSOD6
2.8	Doublet	CH2
3.4	Singlet	Impurity
3.9	Singlet	CH3
3.8	Singlet	CH
6.9	Doublet	CH
7.0	Triplet	CH
7.2-7.3	Doublet	CH
7.3	Doublet	CH

Appendix 11 – Ligand 11 Proton QC



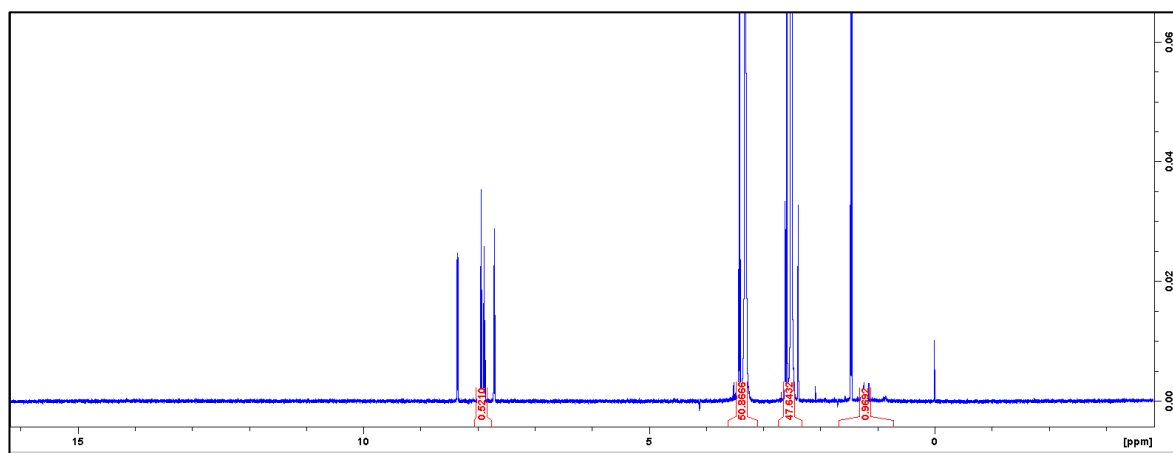
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.7	Singlet	CH3
7.1	Singlet	CH Aromatic
11.0	Singlet	NH
5.9	Singlet	Impurity?

Appendix 12 – Ligand 12 Proton QC



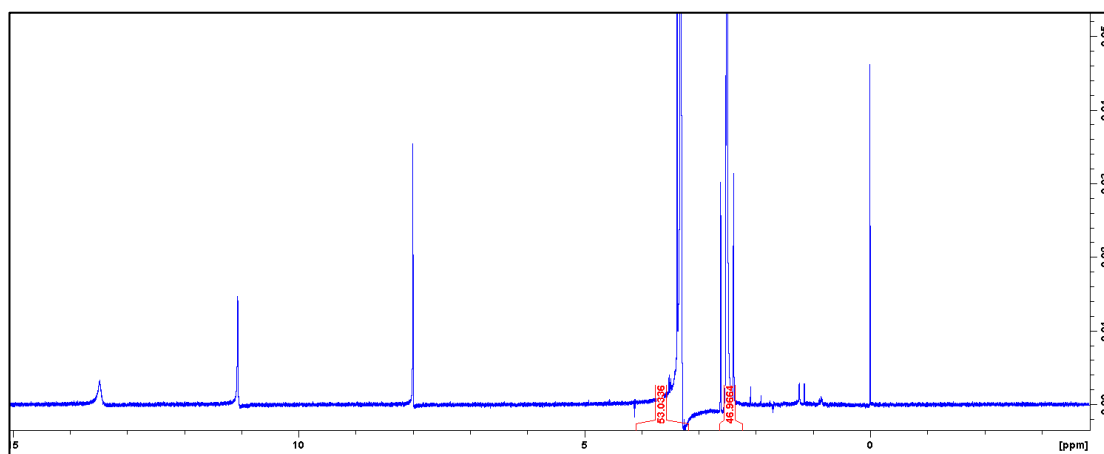
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
2.5	Singlet	CH3
2.9	Singlet	CH3
3.4	Singlet	Impurity
7.8	Doublet	CH
8.0	Doublet	CH
8.1	Singlet	CH
8.8	Singlet	CH
4.2	Singlet	Impurity?

Appendix 13 – Ligand 13 Proton QC



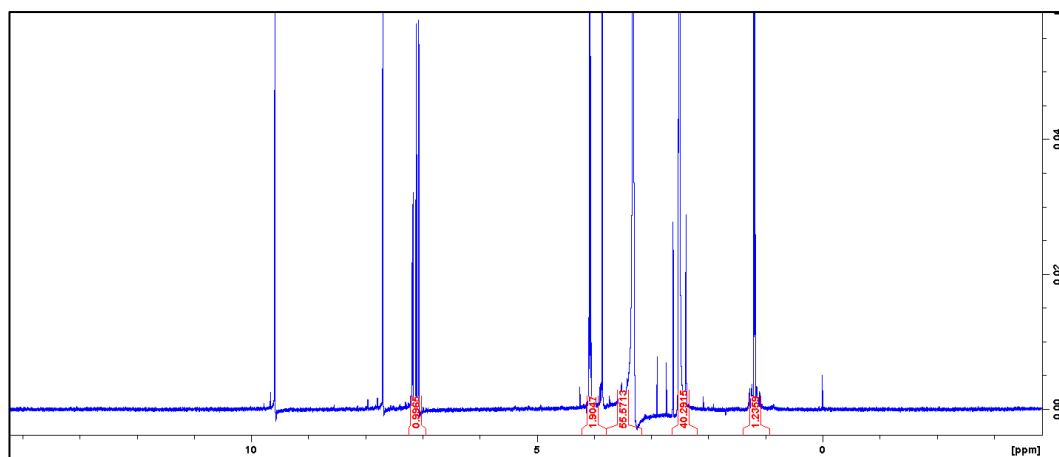
Shift	Splitting	Corresponds to
0	Singlet	TMS
1.4	Triplet	CH3
2.4	Singlet	DMSOD6
2.6	Singlet	CH3
3.4	Singlet	Impurity
3.4 (overlap)	Quartet	CH2
7.8	Doublet	CH
8.0	Doublet	CH
8.1	Singlet	CH
8.8	Singlet	CH

Appendix 14 – Ligand 14 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.4 (overlap)	Singlet	Ch3
8.0	Doublet	CH
11.0	Doublet	NH
13.5	Singlet	NH

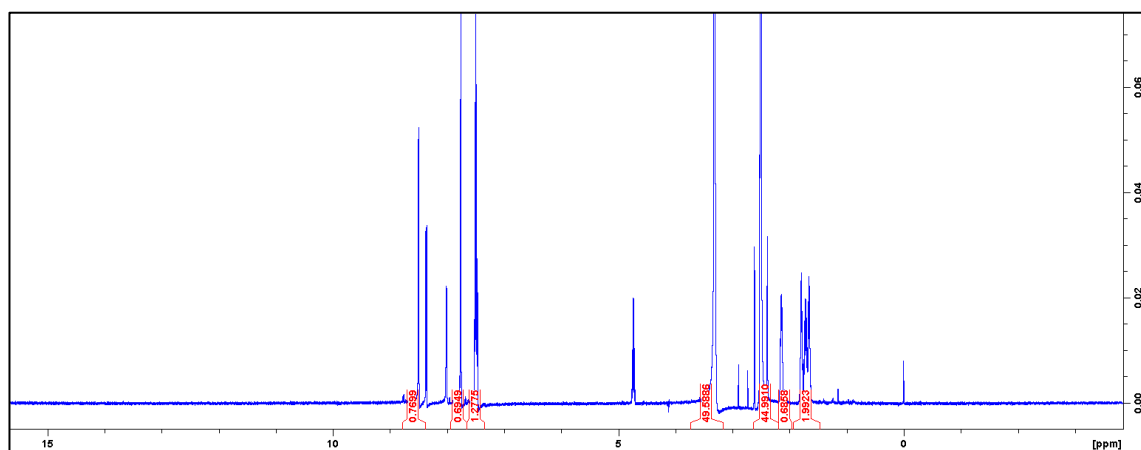
Appendix 15 – Ligand 15 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
1.2	Triplet	CH3
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.9	Singlet	CH3
4.1	Quartet	CH2
7.0	Multiplet	CH
7.1	Doublet	CH
7.7	Singlet	CH
9.8	Singlet	CH

No O-H Peak? Signal too weak?

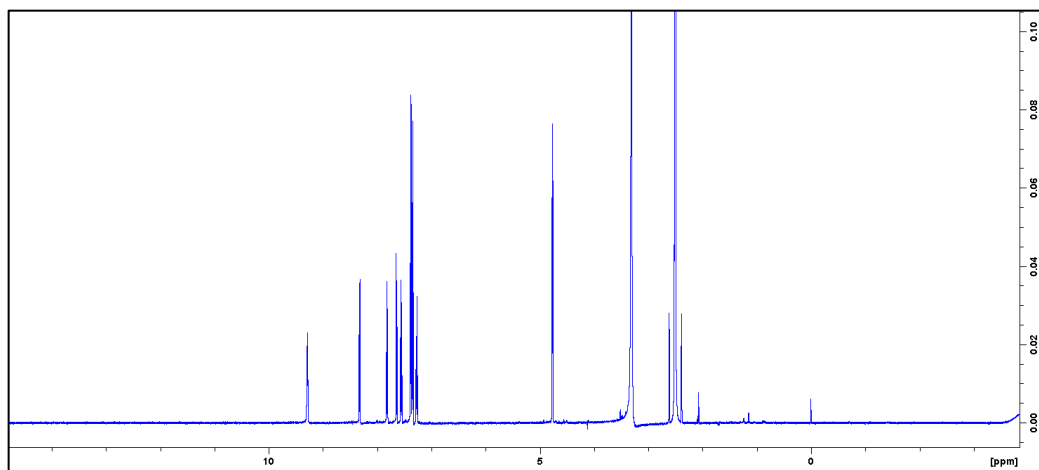
Appendix 16 – Ligand 16 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
1.6	Multiplet	CH ₂
2.1	Doublet	CH ₂
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
4.8	Pentet	CH
7.4	Doublet	CH
7.8	Doublet	CH
8.1	Doublet	CH
8.3	Doublet	CH
8.4	Doublet	CH

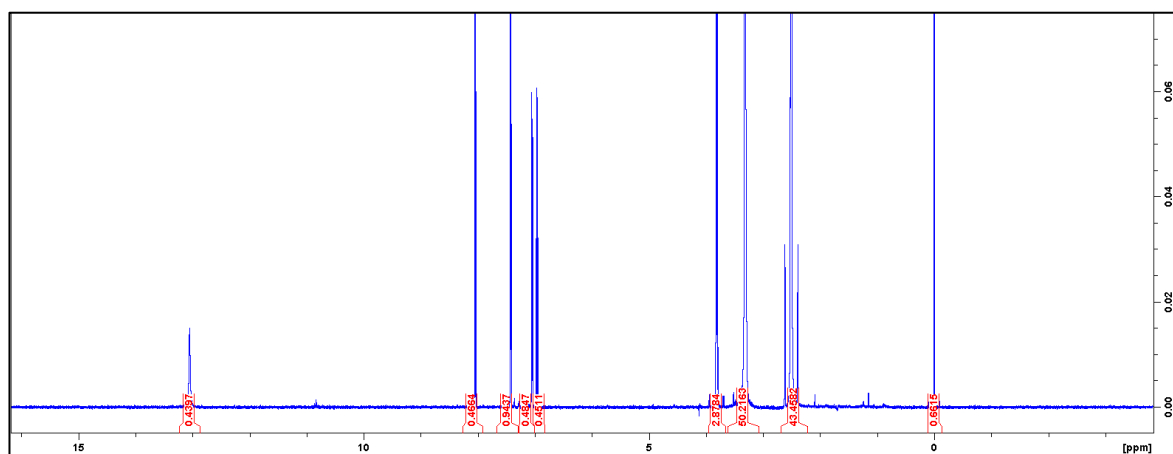
No N-H peak? Signal too weak?

Appendix 17 – Ligand 17 Proton QC



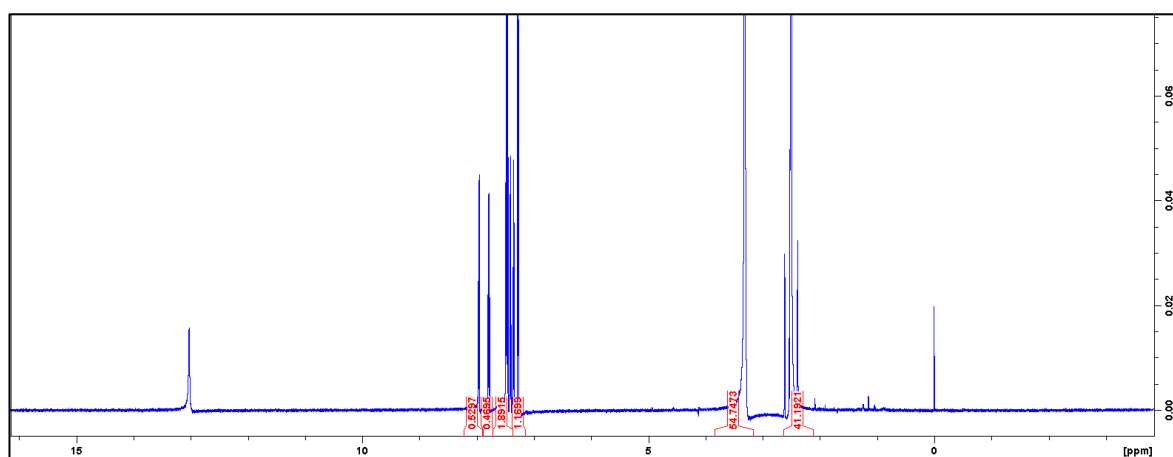
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
4.8	Doublet	CH ₂
7.25	Triplet	CH
7.3	Multiplet	CH
7.4	Doublet	CH
8.3	Doublet	CH
9.4	Triplet	CH

Appendix 18 – Ligand 18 Proton QC



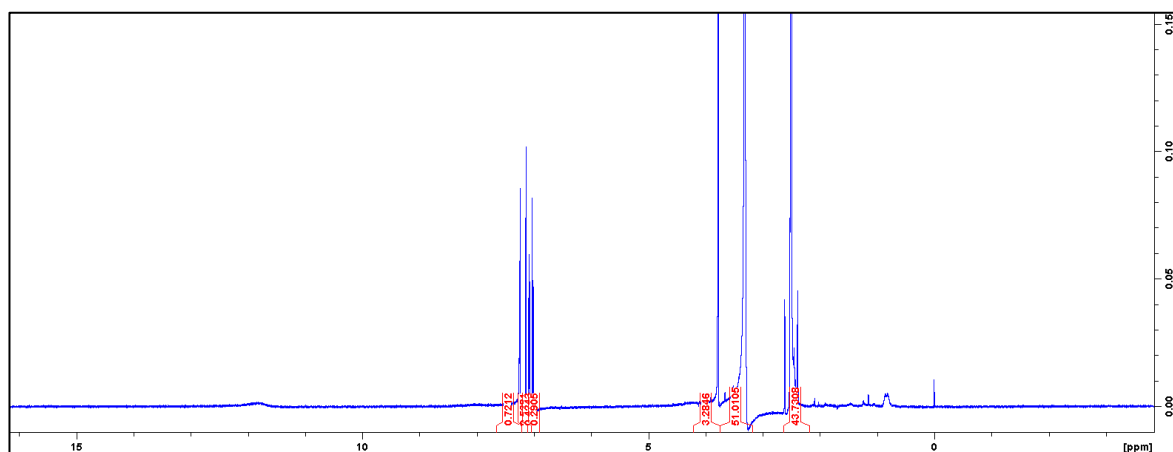
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.8	Singlet	CH3
7.7	Doublet	CH
7.5	Doublet	CH
7.1	Doublet	CH
7.0	Triplet	CH
13.0	Singlet	OH? Or impurity

Appendix 19 – Ligand 19 Proton QC



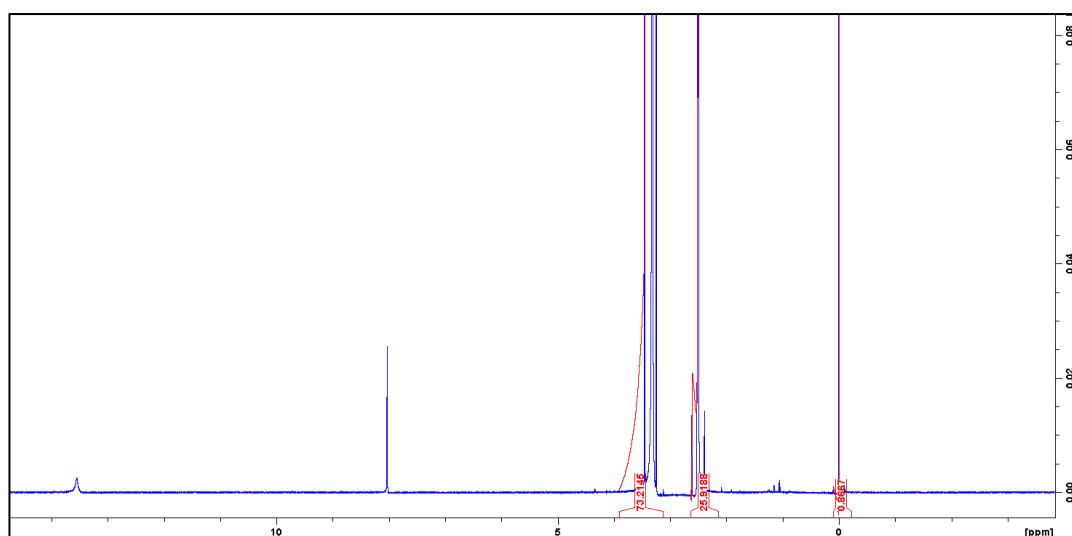
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
7.3-7.5	Multiplet	CH
7.9	Triplet	CH
8.1	Doublet	CH
13.0	Singlet	SH? Or impurity

Appendix 20 – Ligand 20 Proton QC



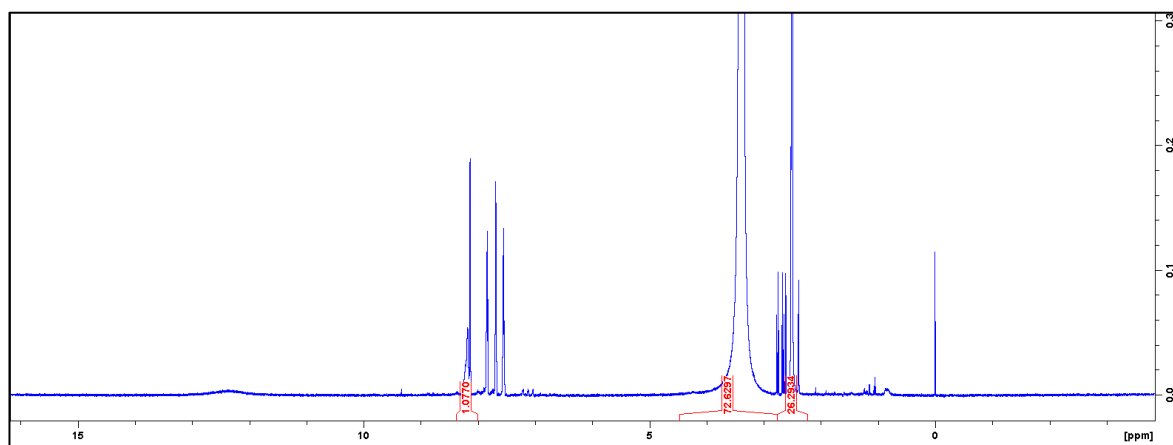
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.9	Singlet	CH3
7.0	Doublet	CH
7.25	Doublet	CH
7.3	Singlet	CH
7.2	Singlet	CH
12	Singlet	NH? Or impurity

Appendix 21 – Ligand 21 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.3	Singlet	CH3
3.4 (overlap)	Singlet	CH3
8.0	Singlet	CH
13.5	Singlet	NH? Or impurity

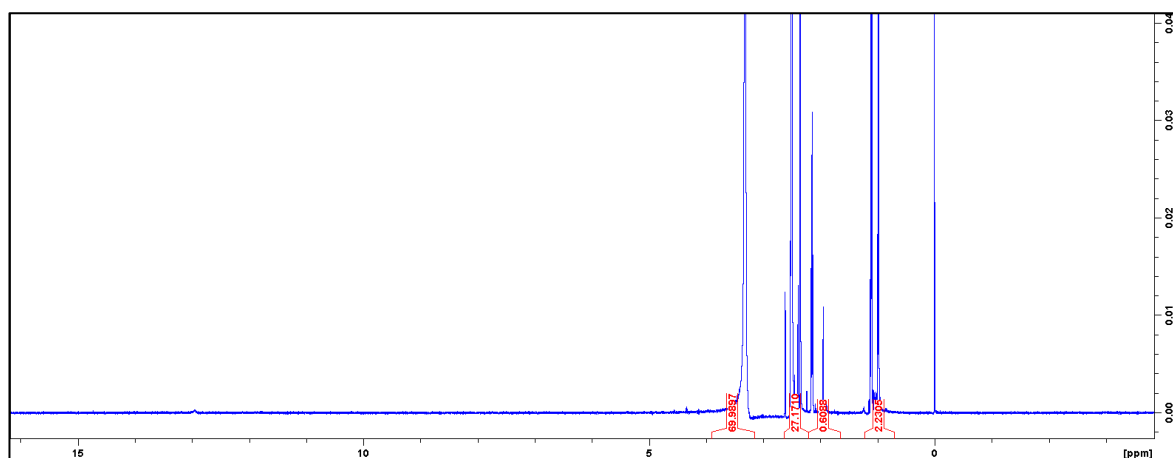
Appendix 22 – Ligand 22 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
2.2-3	Multiplet	???
3.4	Singlet	Impurity
7.6-8.4	Doublets	Aromatic CHs

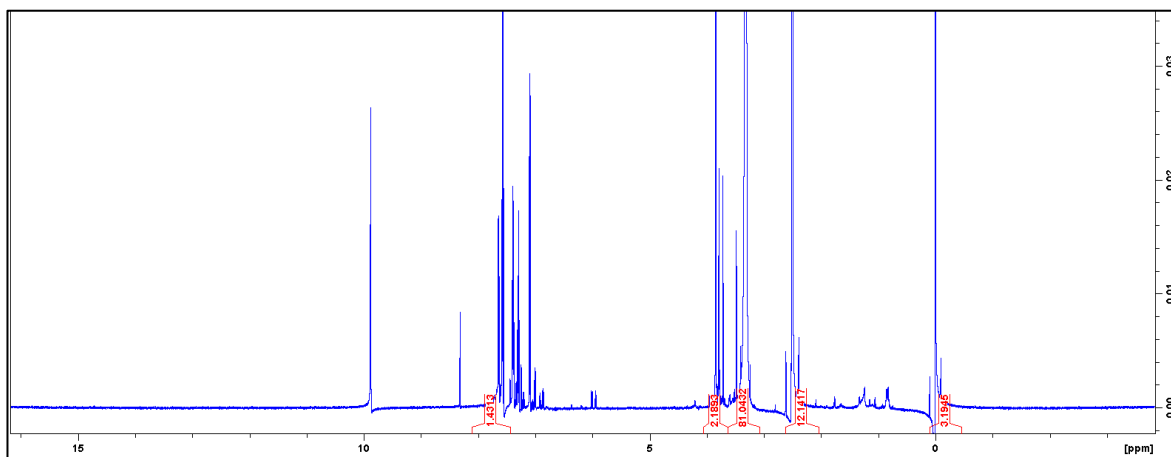
Large peak at 9.0 missing of CH adjacent to nitrogen. Spectrum does not look promising

Appendix 223 – Ligand 23 Proton QC



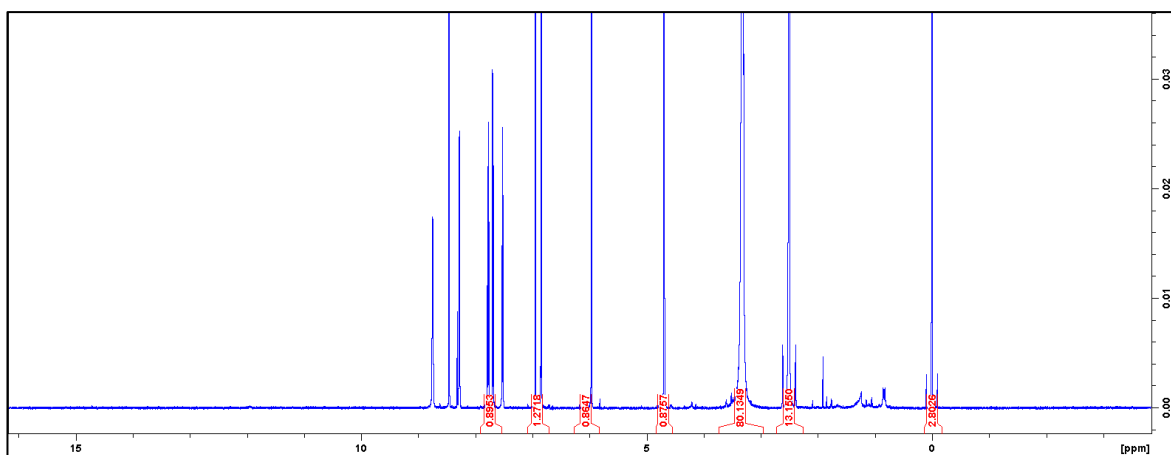
Shift	Splitting	Corresponds to
0	Singlet	TMS
1.0-1.1	Doublet (roofing)	CH ₂
1.8	Multiplet	CH
2.4	Singlet	DMSOD6
2.4 (overlap)	Singlet	CH ₃
3.4	Singlet	Impurity

Appendix 24 – Ligand 24 Proton QC



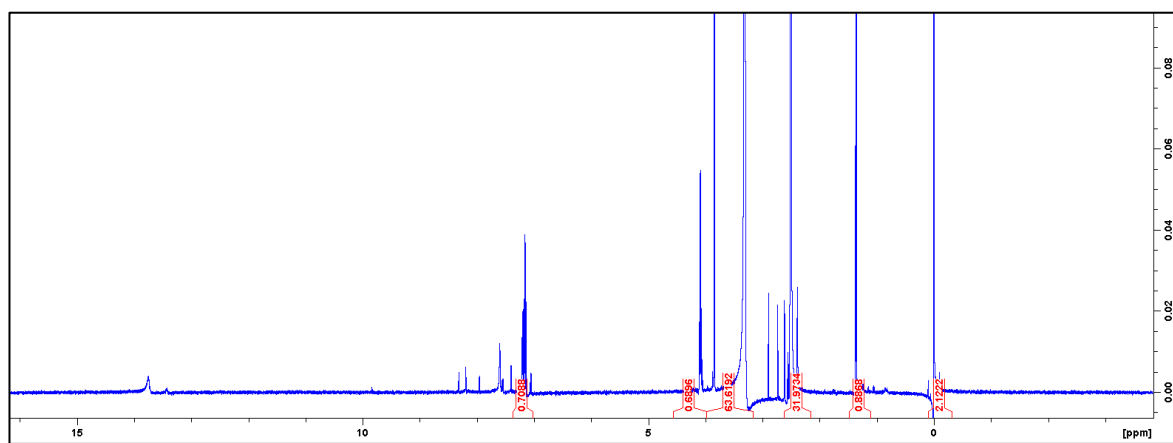
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.9	Singlet	CH3
6.9-8.2	Doublets	Aromatic CH
10.0	Singlet	NH

Appendix 25 – Ligand 25 Proton QC



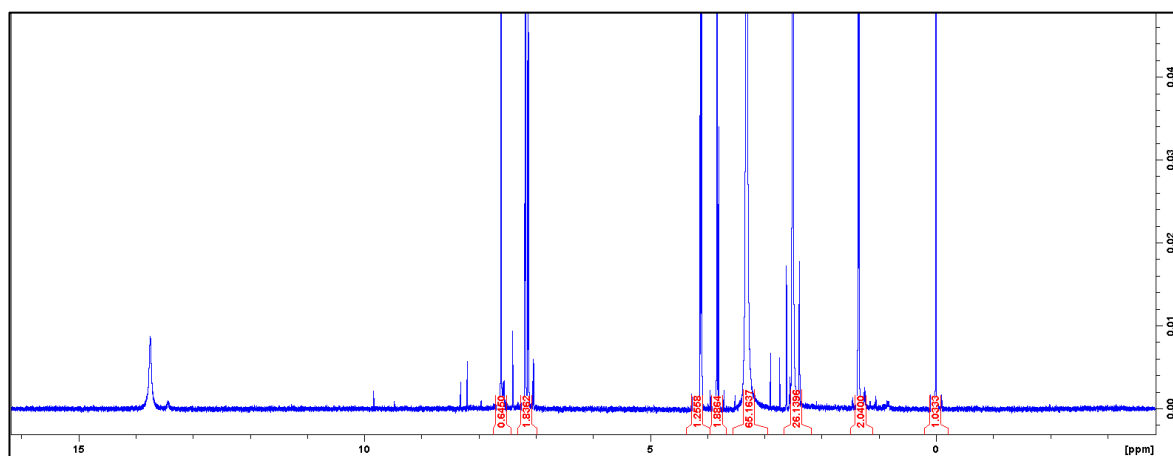
Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
4.6	Doublet	CH2
5.9	Singlet	CH2
6.9	Doublet	CH
6.95	Doublet	CH
7.6	Triplet	CH
7.9	Doublet (roofing)	CH
8.1	Doublet	CH
8.3	Singlet	CH
8.9	Singlet	NH?

Appendix 26 – Ligand 26 Proton QC



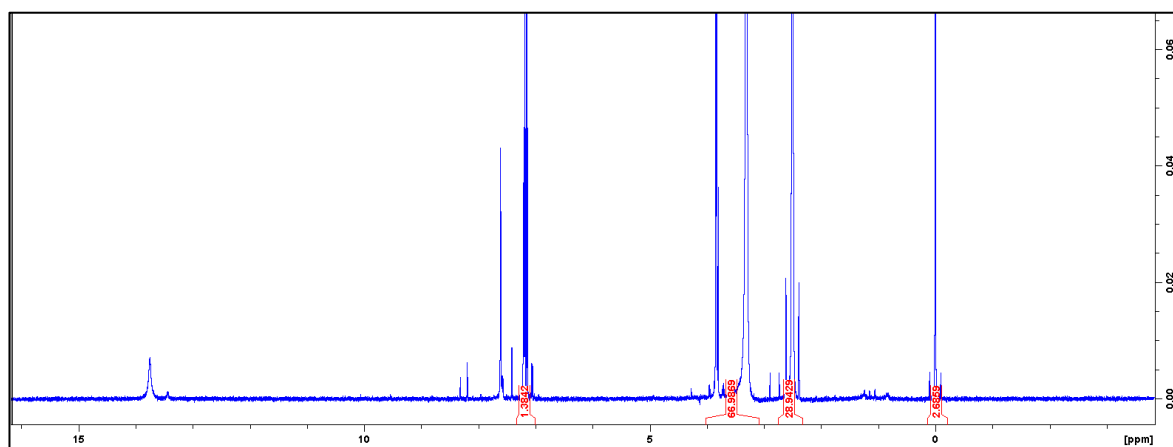
Shift	Splitting	Corresponds to
0	Singlet	TMS
1.4	Doublet	CH3
2.4	Singlet	DMSOD6
2.5	Singlets	Impurity?
3.4	Singlet	Impurity
3.8	Singlet	CH3
4.1	Quartet	CH2
7.0	Doublet	CH
7.2	Doublet	CH
7.4	Singlet	CH
7.5	Singlet	CH
13.8	Singlet	NH? Or impurity

Appendix 27 – Ligand 28 Proton QC



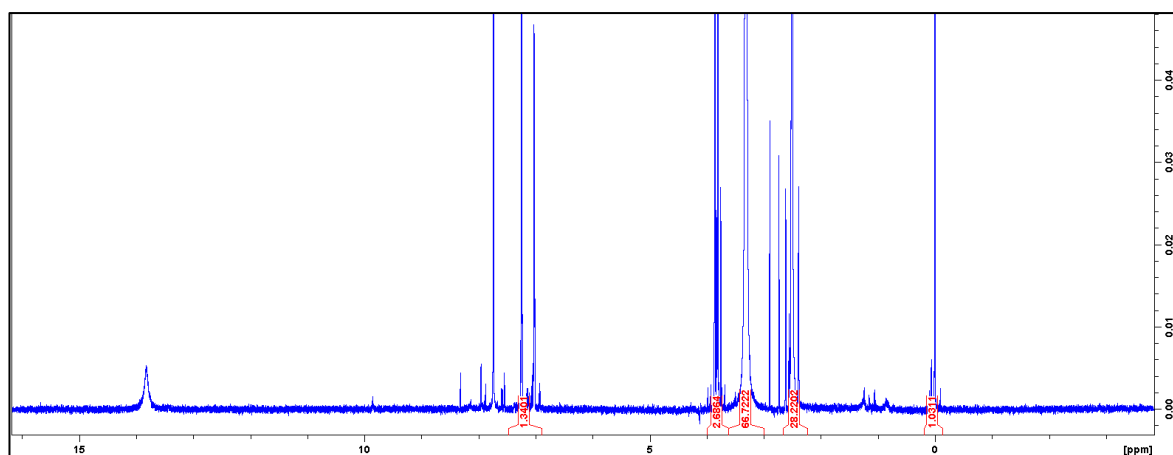
Shift	Splitting	Corresponds to
0	Singlet	TMS
1.3	Triplet	CH3
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.8	Singlet	CH3
4.1	Quartet	CH2
7.0	Doublet	CH
7.1	Doublet	CH
7.4	Singlet	CH
7.7	Singlet	CH
13.8	Singlet	NH? Or impurity

Appendix 28 – Ligand 30 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
3.4	Singlet	Impurity
3.8	Singlet	CH3
7.0	Doublet	CH
7.2	Doublet	CH
7.3	Singlet	CH
7.4	Singlet	CH
14.8	Singlet	NH? Or impurity

Appendix 29 – Ligand 31 Proton QC



Shift	Splitting	Corresponds to
0	Singlet	TMS
2.4	Singlet	DMSOD6
2.5-2.8	Singlets	Impurity?
3.4	Singlet	Impurity
3.9	Singlet	CH3
3.95	Singlet	CH3
7.0	Doublet	CH
7.1	Quartet	CH
7.2	Doublet	CH
7.6	Singlet	CH
13.9	Singlet	NH? Or impurity

Appendix 30 – X-Ray Crystallography manual hit overview table and key

Ligand #	Ligand ID	Soak Details				
		40mM / 24 hours	40mM / 48 hours	30mM / 24 hours	50mM / 72 hours	50mM / 96 hours
1	CHEMBL31877	1 - /	3 - CPS4346 - 5 - Res 1.39	1 - CPS4347 - 5 - Res 1.67		
2	CHEMBL88119	1 - CPS4347 - 9 - Res 1.63	3 - CPS4346 - 6 - Res 1.47	1 - CPS4347 - 8 - Res 1.99		4 - CPS4347 - 9 - Res 1.59
3	CHEMBL131164	1 - CPS4347 - 11 - Res 1.99	3 - CPS4346 - /	1 - CPS4347 - 10 - Res 2.26		
4	CHEMBL124706	1 - CPS4347 - 12 - Res 3.58	3 - CPS4346 - 8 - /	1 - CPS4347 - 13 - Res 1.88		
5	CHEMBL760	/	3 - CPS4346 - /	/	/	/
6	CHEMBL279898	1 - /	3 - CPS4346 - 10 - Res 1.56	1 - CPS4347 - 14 - Res 1.86		
7	CHEMBL189	2 - CPS4346 - 1 - Res 1.78	2 - CPS4347 - 9 - Res 1.87			4 - CPS4347 - 10 - Res 1.49
8	CHEMBL58355	2 - CPS4346 - 2 - Res 2.94	2 + 3 - /			4 - CPS4347 - 11 - Res 1.93
9	CHEMBL131181	2 - CPS4346 - 3 - Res 1.71	2 - CPS4347 - 10 - Res 1.83			
10	CHEMBL330581	2 - CPS4346 - 4 - Res 1.8	2 - CPS4347 - 11 - Res 2.14			4 - CPS4347 - 12 - Res 1.75
11	CHEMBL100112	2 - CPS4346 - 5 - Res 1.97	3 - CPS4346 - 12 - Res 1.52			4 - CPS4347 - 15 - Res 1.55
12	CHEMBL1779264	/	3 - CPS4346 - 13 - Res 1.48			
13	CHEMBL1779265	/	3 - CPS4346 - /		/	/
14	CHEMBL619	/	3 - CPS4346 - /		/	/
15	CHEMBL255611	2 - CPS4346 - 6 - Res 1.64	2 - CPS4347 - 12 - Res 2.11			4 - CPS4347 - 13 - Res 1.56
16	CHEMBL119506	2 - CPS4346 - 7 - Res 1.76	2 - CPS4347 - 13 - Res 1.84		4 - CPS4348 - 5 - Res 1.89	4 - /
17	CHEMBL118193	2 - CPS4346 - 8 - Res 1.76	3 - CPS4346 - 7 - Res 1.79			
18	CHEMBL372055	2 - CPS4346 - 9 - Res 1.64	3 - CPS4346 - 9 - Res 1.19			4 - CPS4347 - 14 - Res 1.75
19	CHEMBL451589	2 - /	2 - CPS4347 - 14 - Res 1.85		/	/
20	CHEMBL255421		3 - CPS4346 - 15 - Res 1.65		4 - /	4 - /
21	CHEMBL190		3 - CPS4346 - /		4 - CPS4348 - 9 - Res 1.75	4 - /
22	CHEMBL2172707		3 - CPS4346 - 16 - Res 1.79			
23	CHEMBL3622905		3 - CPS4349 - 1 - Res 1.47			
24	CHEMBL45854		3 - CPS4349 - 2 - Res 1.69			4 - CPS4347 - 16 - Res 1.47
25	CHEMBL66732		3 - CPS4349 - 8 - Res 1.52		/	/
26	CHEMBL270636		3 - CPS4349 - 3 - Res 1.44		/	/
27	CHEMBL1779267		3 - CPS4349 - 4 - Res 1.27		4 - CPS4348 - 11 - Res 1.48	4 - /
28	CHEMBL270635		3 - CPS4349 - 5 - Res 1.48		/	/
29	CHEMBL484928		3 - CPS4349 - 6 - Res 1.52		/	/
30	CHEMBL255130		3 - CPS4349 - 7 - Res 1.38		/	/
31	CHEMBL255821		3 - CPS4349 - /		/	/

Key	
	Done - non binder by X-Ray
	Successful hit - ligand modeled
	Ligand crashing out of solution
	Wrong space group - poor quality data
	Crystal lost
	? - Check notes / come back to
Trip x - Puck code - Puck Position - Resolution	
Trip 1 - 16/02/18 - Code mx16207 - Beamline io4-1	
Trip 2 - 11/03/18 - Code mx16207 - Beamline io4-1	
Trip 3 - 23/04/18 - Code mx16207 - Beamline io4-1	
Trip 4 - 21/05/18 - Code mx16207 - Beamline io4-1	
Trip 5 -	
Unit Cell : a = 115 b = 115 c = 68	

Appendix 31 – X-Ray Crystallography XCHEM hit overview table and key

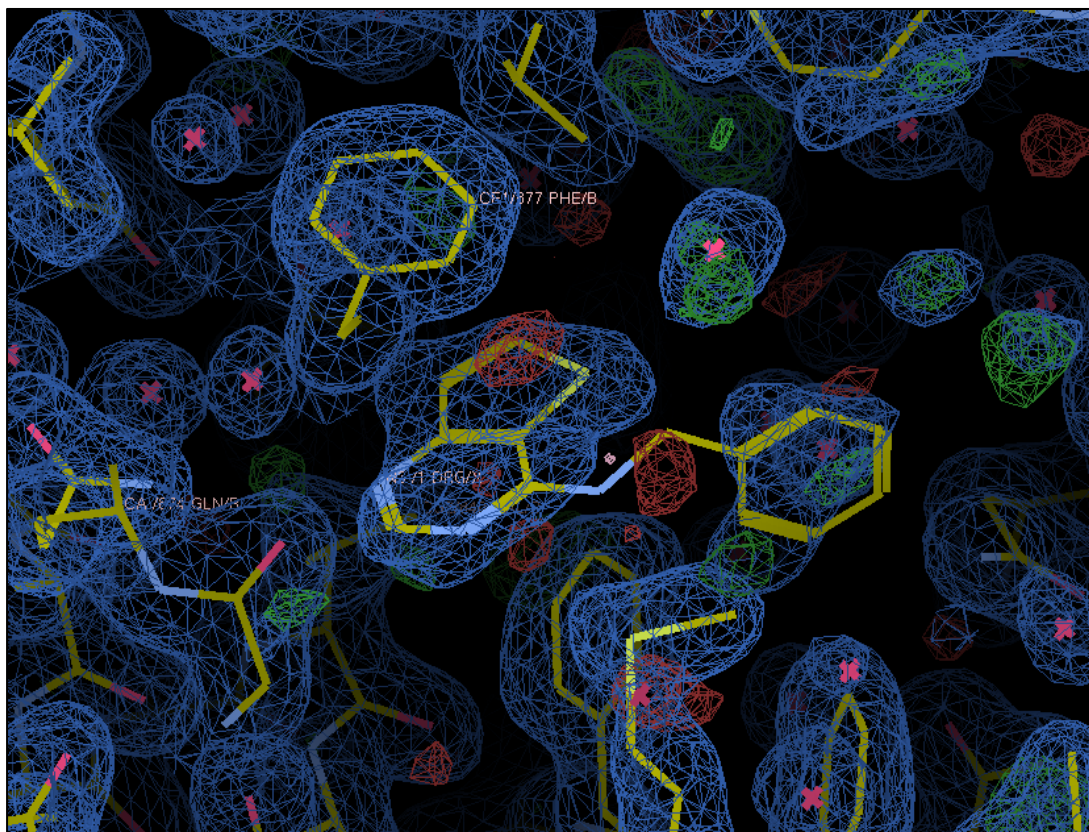
Ligand #	Ligand ID	Soak Details	
		20mM / 3 hours	30mM / 3 hours
1	CHEMBL31877	XCHEM - x0428_1 - DF006 - 1 - Res 2.05	XCHEM - x0455_1 - DF0019 - 12 - Res 8.00
2	CHEMBL88119	XCHEM - x0429_1 - DF006 - 2 - Res 4.04	XCHEM - x0456_1 - DF0019 - 13 - Res 7.14
3	CHEMBL131164	XCHEM - Melted	XCHEM - Melted
4	CHEMBL124706	XCHEM - x0430_1 - DF006 - 3 - Res 2.04	XCHEM - x0457_1 - DF0019 - 14 - /
5	CHEMBL760	XCHEM - x0431_2 - DF006 - 4 - Res 2.33	XCHEM - x0458_1 - DF0019 - 15 - Res 3.01
6	CHEMBL279898	XCHEM - Melted	XCHEM - Melted
7	CHEMBL189	XCHEM - x0432 - DF006 - 5 - /	XCHEM - Melted
8	CHEMBL58355	XCHEM - x0433_1 - DF006 - 6 - Res 5.05	XCHEM - Melted
9	CHEMBL131181	XCHEM - x0434_1 - DF006 - 7 - Res 2.13	XCHEM - x0459_1 - DF0019 - 16 - Res 2.43
10	CHEMBL330581	XCHEM - x0435_2 - DF006 - 8 - Res 6.85	XCHEM - x0460 - DLS599 - 1 - /
11	CHEMBL100112	XCHEM - Melted	XCHEM - Melted
12	CHEMBL1779264	XCHEM - x0436_1 - DF006 - 9 - Res 1.98	XCHEM - x0461 - DLS599 - 2 - Res 1.83
13	CHEMBL1779265	XCHEM - x0437_2 - DF006 - 10 - Res 4.68	XCHEM - x0462 - DLS599 - 3 - /
14	CHEMBL619	XCHEM - Melted	XCHEM - Melted
15	CHEMBL255611	XCHEM - x0438_1 - DF006 - 11 - Res 2.12	XCHEM - Melted
16	CHEMBL119506	XCHEM - x0439_1 - DF006 - 12 - Res 3.71	XCHEM - x0463 - DLS599 - 4 - /
17	CHEMBL118193	XCHEM - x0440_1 - DF006 - 13 - Res 2.09	XCHEM - Melted
18	CHEMBL372055	XCHEM - Melted	XCHEM - x0464 - DLS599 - 5 - Res 1.24
19	CHEMBL451589	XCHEM - x0441_1 - DF006 - 14 - Res 1.86	XCHEM - Melted
20	CHEMBL255421	XCHEM - Melted	XCHEM - Melted
21	CHEMBL190	XCHEM - Melted	XCHEM - x0465 - DLS599 - 6 - Res 3.30
22	CHEMBL2172707	XCHEM - x0442_1 - DF006 - 15 - Res 4.60	XCHEM - x0466 - DLS599 - 7 - /
23	CHEMBL3622905	XCHEM - x0443_1 - DF006 - 16 - /	
24	CHEMBL45854	XCHEM - Melted	
25	CHEMBL66732	XCHEM - x0444_1 - DF019 - 1 - Res 2.27	
26	CHEMBL270636	XCHEM - x0445_1 - DF019 - 2 - Res 4.10	
27	CHEMBL1779267	XCHEM - x0446_2 - DF019 - 3 - Res 2.17	
28	CHEMBL270635	XCHEM - Melted	
29	CHEMBL484928	XCHEM - x0447_1 - DF019 - 4 - Res 3.27	
30	CHEMBL255130	XCHEM - x0448_1 - DF019 - 5 - Res 2.13	
31	CHEMBL255821	XCHEM - Melted	

Key	
	Done - non binder by X-Ray
	Successful hit - ligand modeled
	Not recorded due to error, same with 40mM
	Wrong space group - poor quality data
	Couldn't pick crystal
	? - Check notes
	Bad resolution - poor map
XCHEM - ISPyB ID - Puck code - Puck Position - Resolution	
XCHEM - 01/02/18 - Code ib18402 - Beamline io4-1	
Almost all densities in chain B - exceptions noted	
Unit Cell : a = 115 b = 115 c = 68	

Note that the following structure density maps are not fully refined and can undergo further modifications.

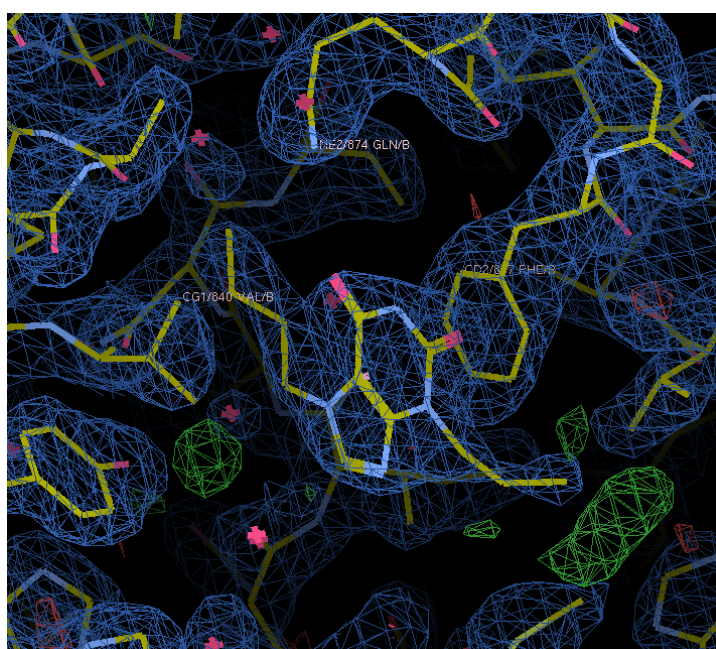
Appendix 32 – Ligand 1 in TbrPDEB1 binding pocket. 30mM soak for 24 hours.

Resolution 1.67Å



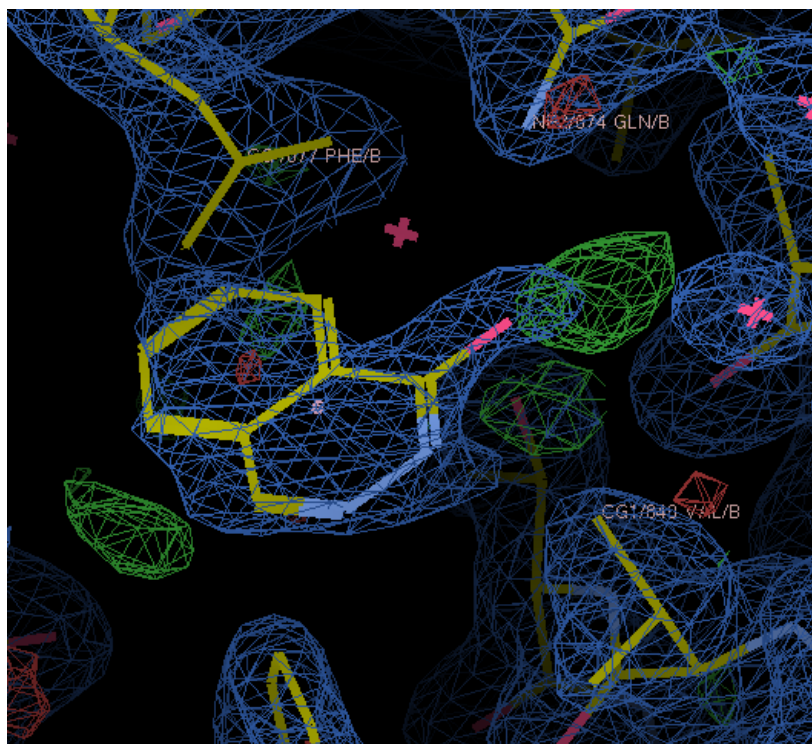
Appendix 33 – Ligand 3 in TbrPDEB1 binding pocket. 30mM soak for 24 hours.

Resolution 2.26Å



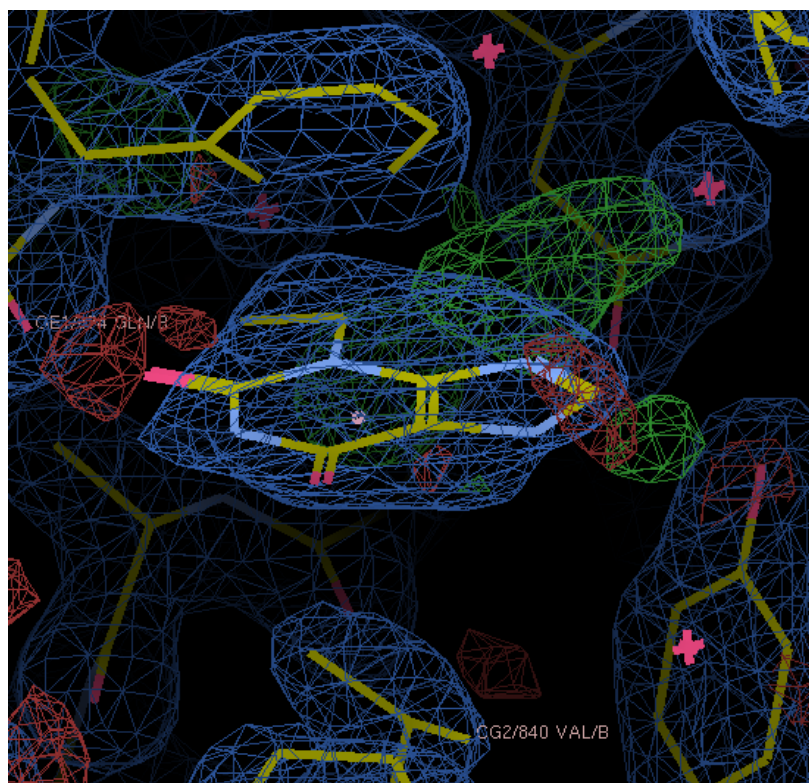
Appendix 34 – Ligand 4 in TbrPDEB1 binding pocket. 30mM soak for 24 hours.

Resolution 1.88Å



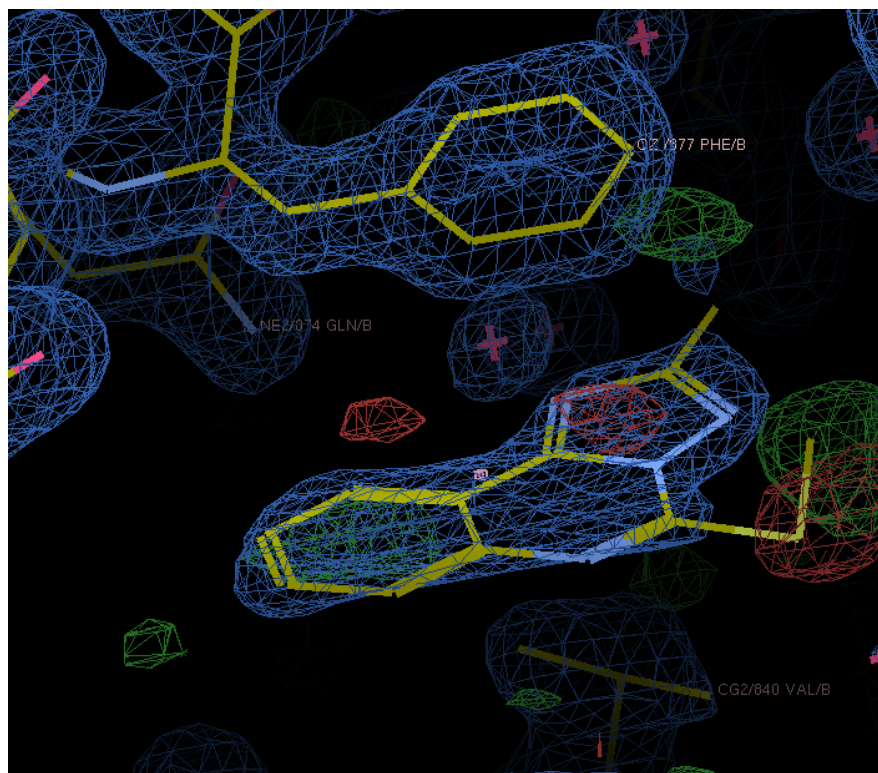
Appendix 35 – Ligand 9 in TbrPDEB1 binding pocket. 40mM soak for 48 hours.

Resolution 1.83Å



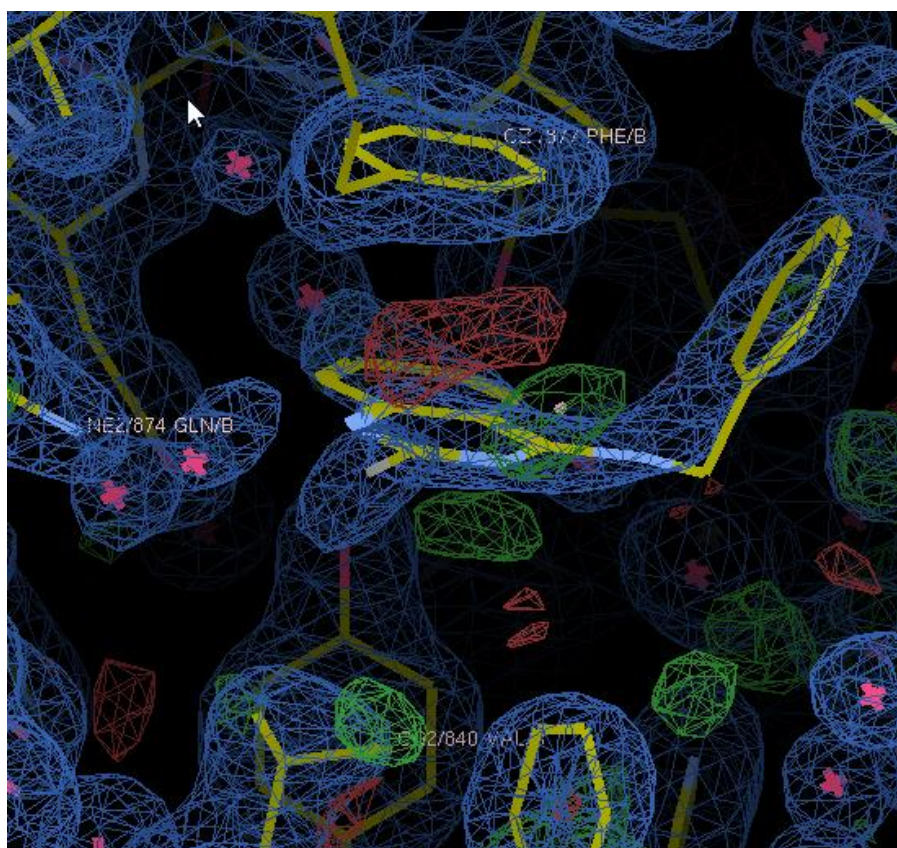
Appendix 36 – Ligand 12 in TbrPDEB1 binding pocket. 40mM soak for 48 hours.

Resolution 1.48Å

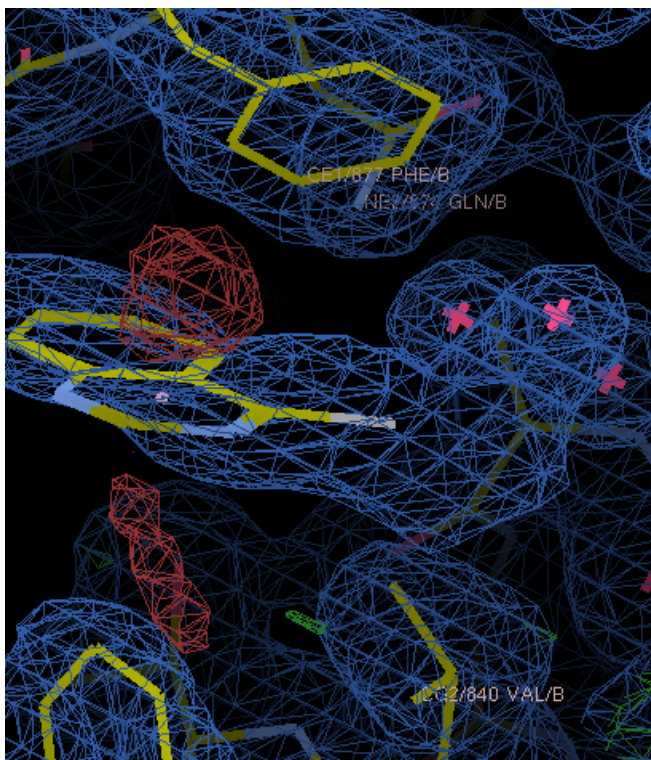


Appendix 37 – Ligand 17 in TbrPDEB1 binding pocket. 40mM soak for 48 hours.

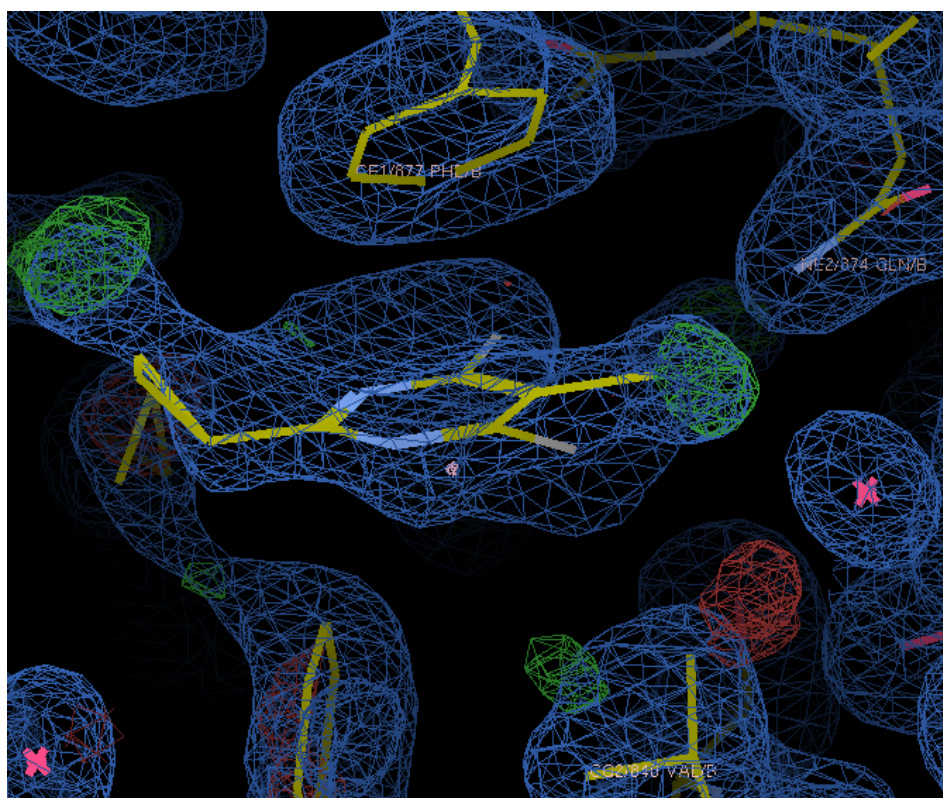
Resolution 1.79Å



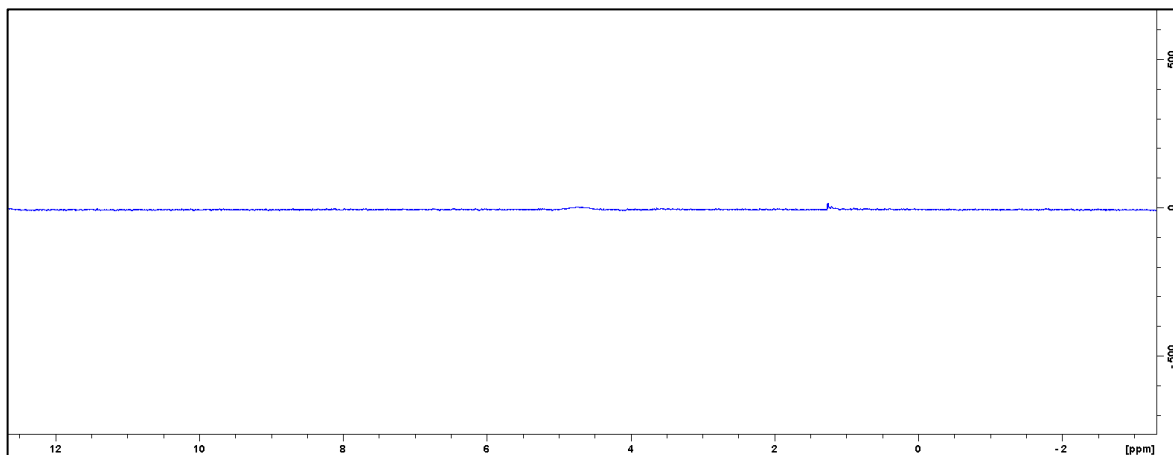
Appendix 38 – Ligand 22 in TbrPDEB1 binding pocket. 40mM soak for 48 hours.
Resolution 1.79Å



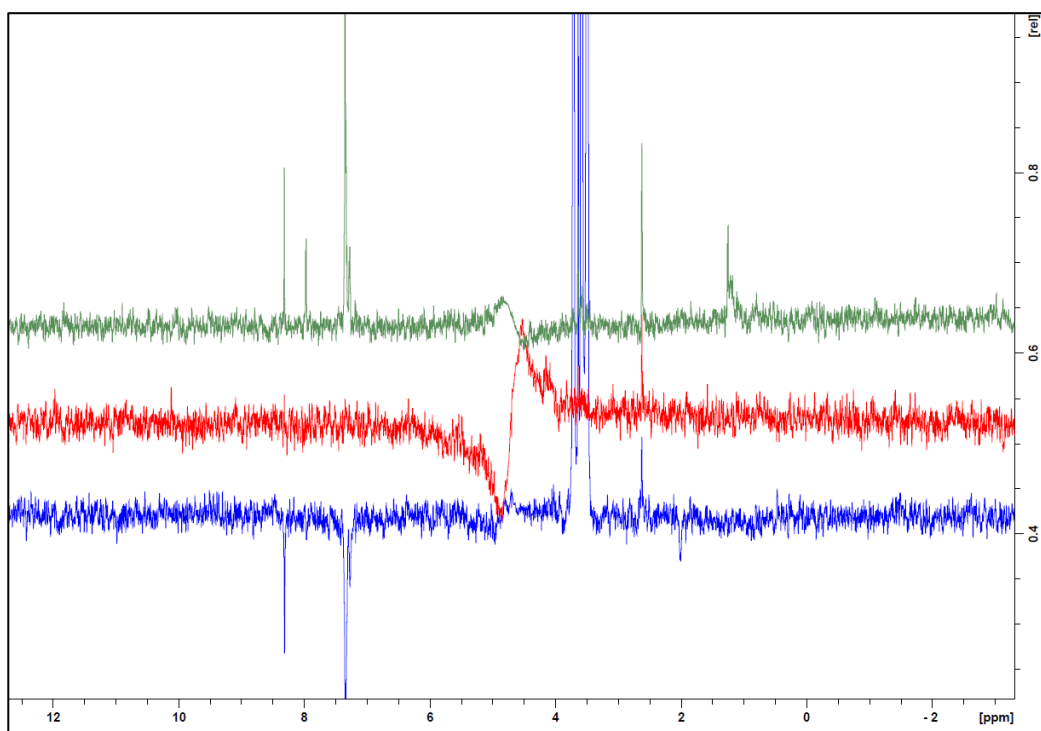
Appendix 39 – Ligand 23 in TbrPDEB1 binding pocket. 40mM soak for 48 hours.
Resolution 1.47Å



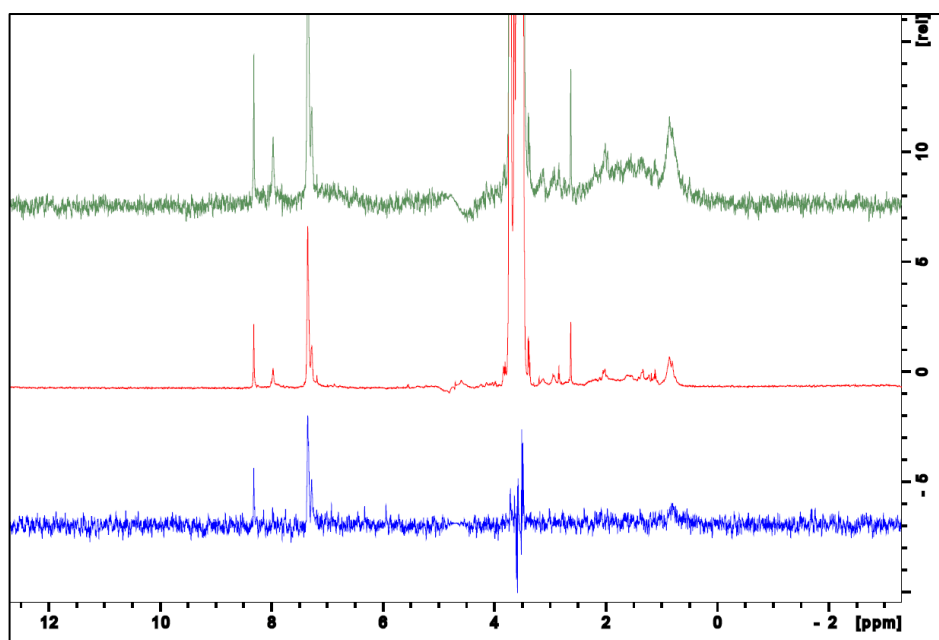
Appendix 40 – Phosphate buffer solvent used to make up Protein-Ligand solutions Proton
1D NMR



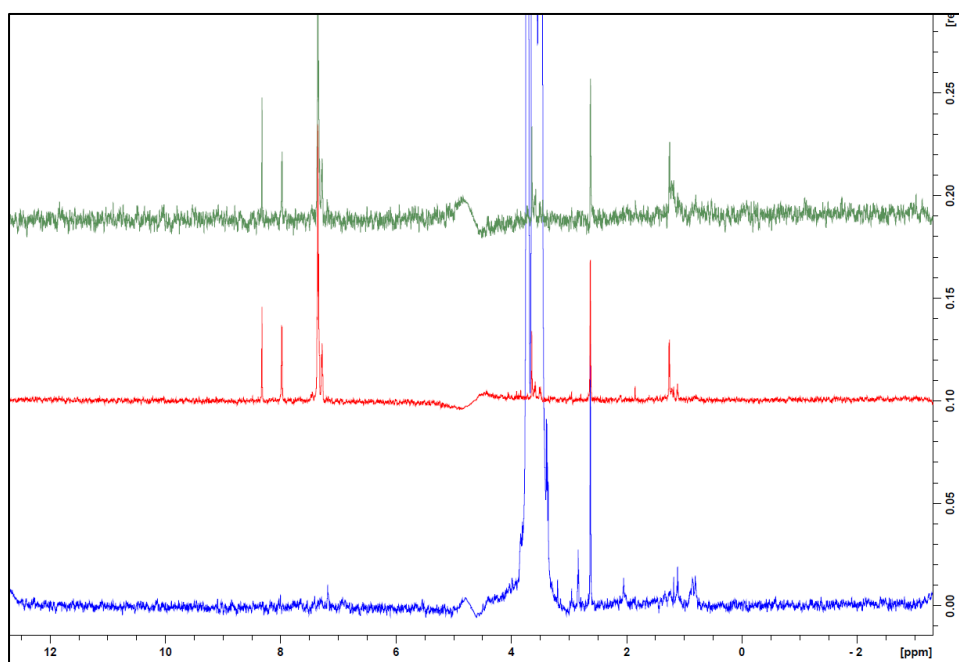
Appendix 41 - Ligand 1-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D
Spectrum (Top), Ligand 1 Control WaterLOGSY spectrum (Middle), Ligand-Protein
WaterLOGSY spectrum (Bottom)



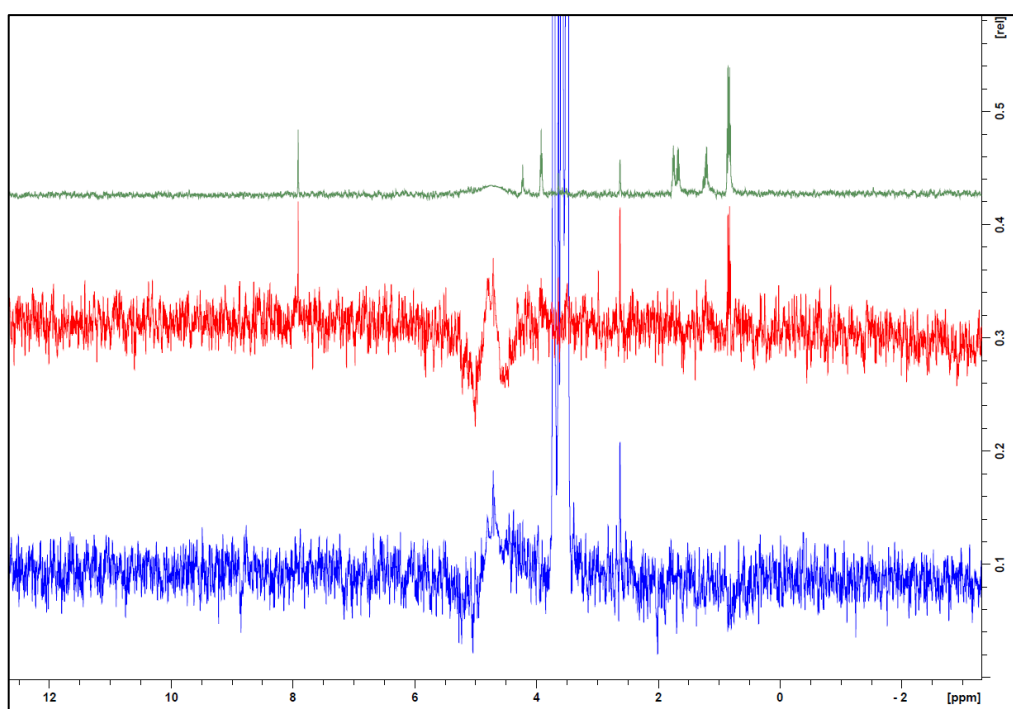
Appendix 42 - Ligand 1-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 1 on-resonance control spectrum (Middle), difference spectrum (Bottom)



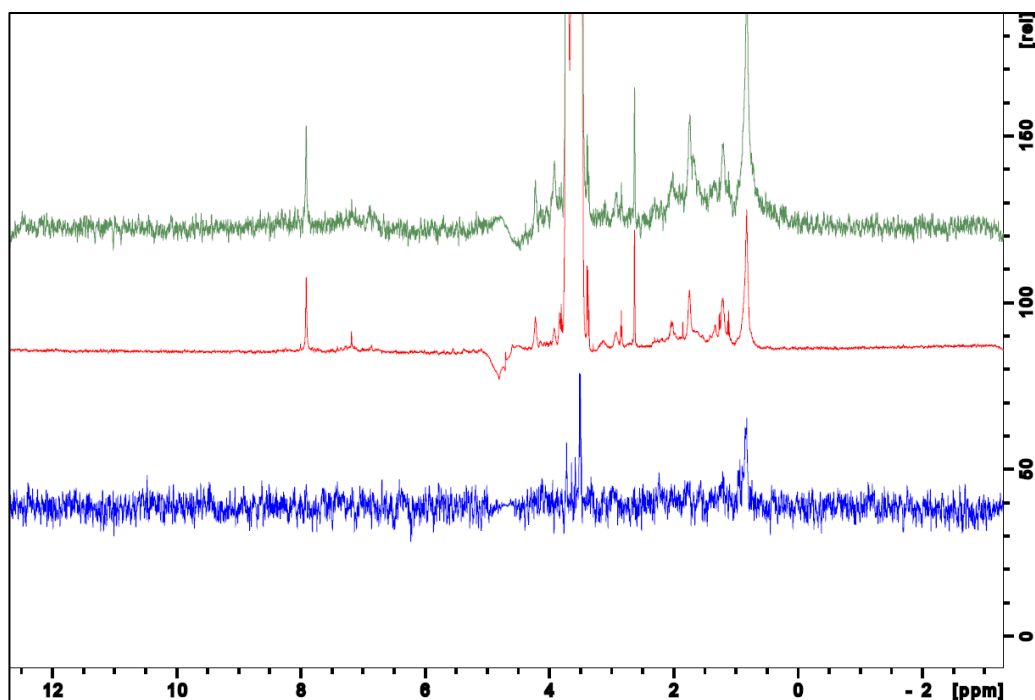
Appendix 43 – Ligand 1-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 1 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



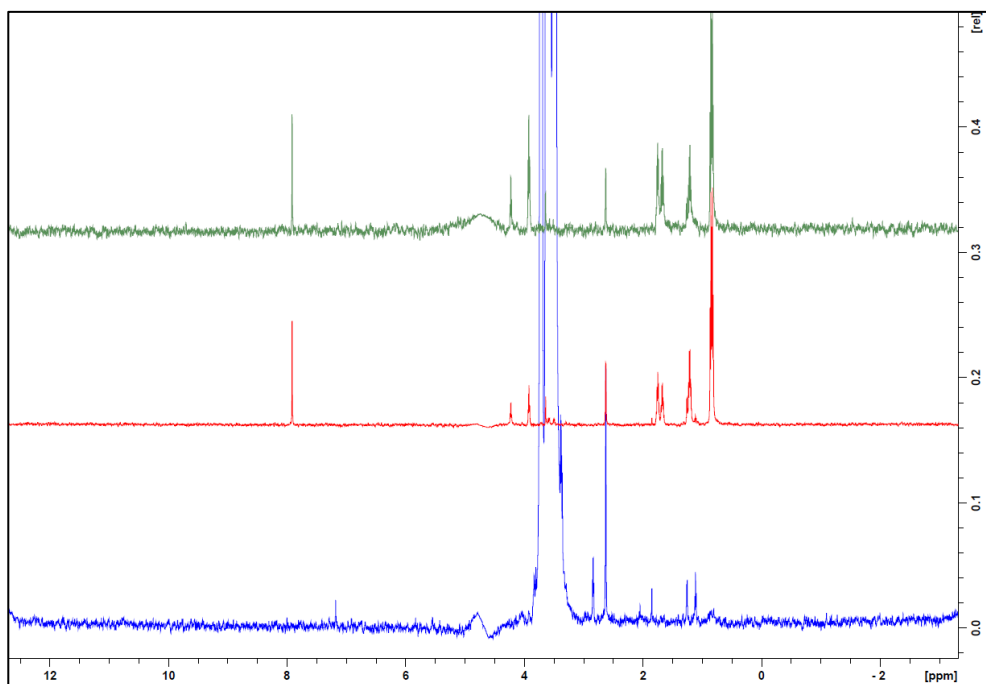
Appendix 44 - Ligand 3-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 3 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



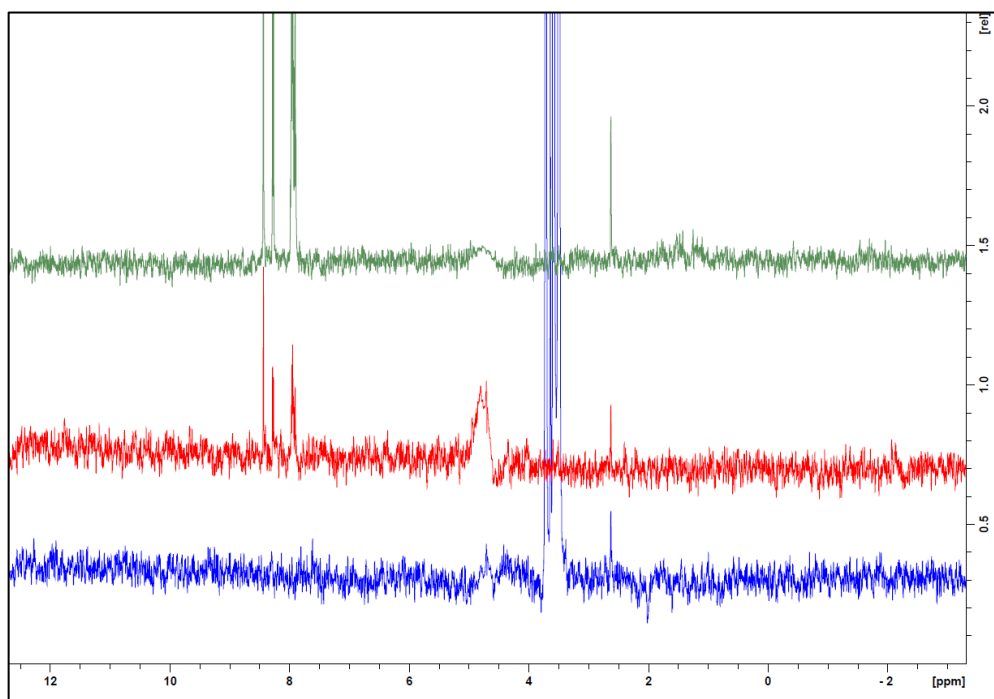
Appendix 45 - Ligand 3-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 3 on-resonance control spectrum (Middle), difference spectrum (Bottom)



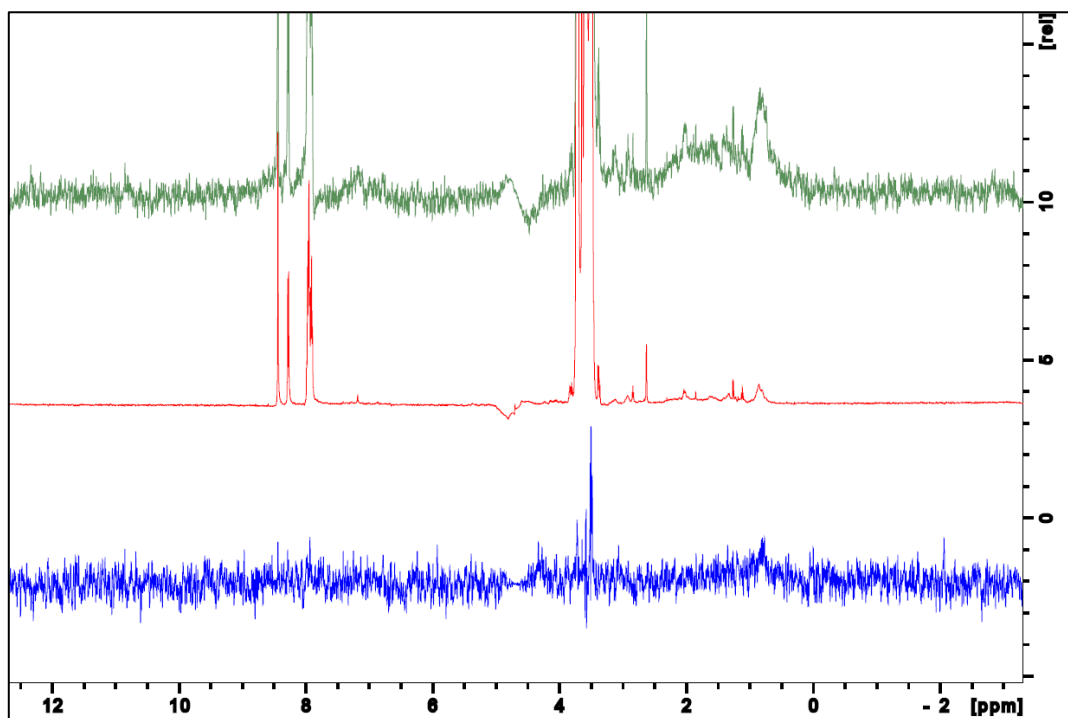
Appendix 46 – Ligand 3-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 3 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



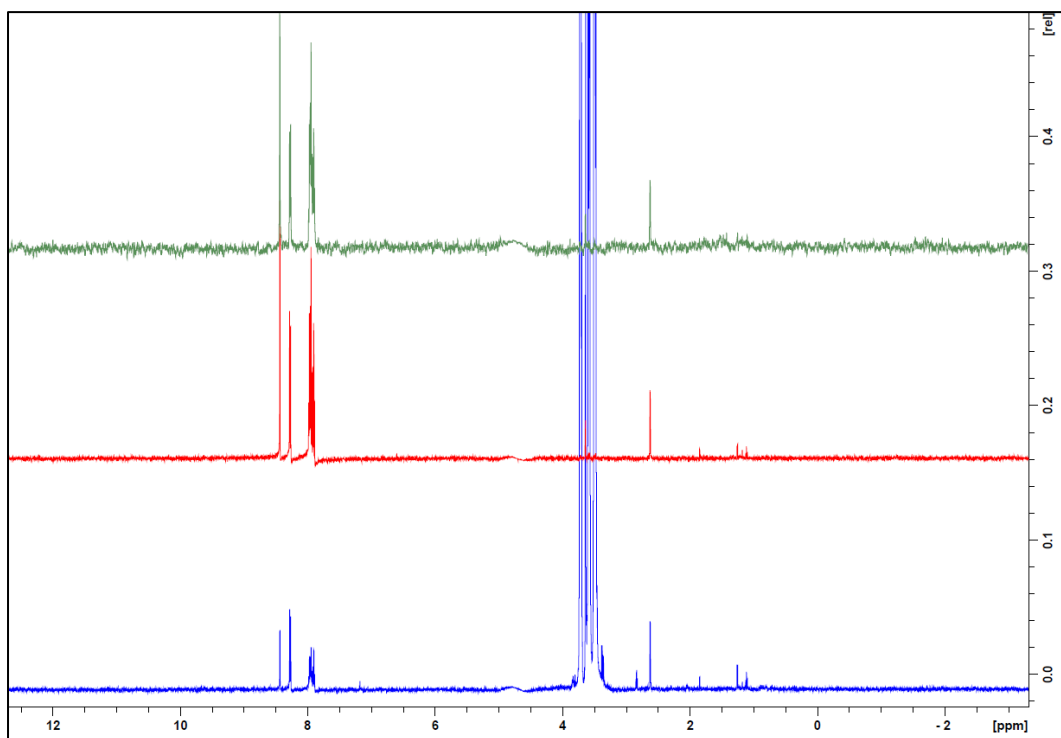
Appendix 47 - Ligand 4-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 4 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



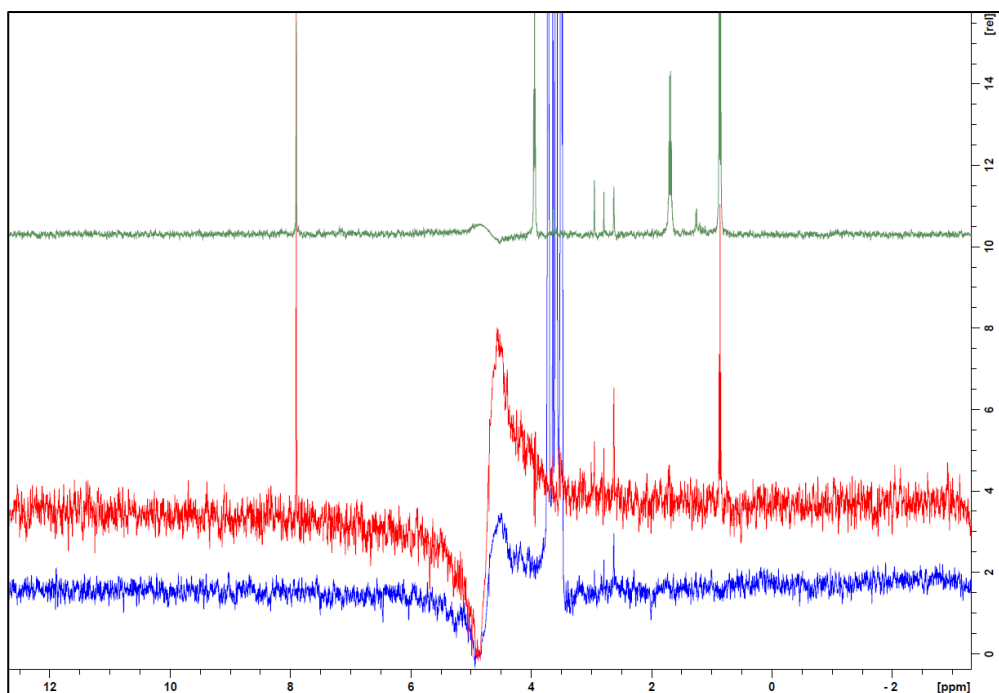
Appendix 48 - Ligand 4-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 4 on-resonance control spectrum (Middle), difference spectrum (Bottom)



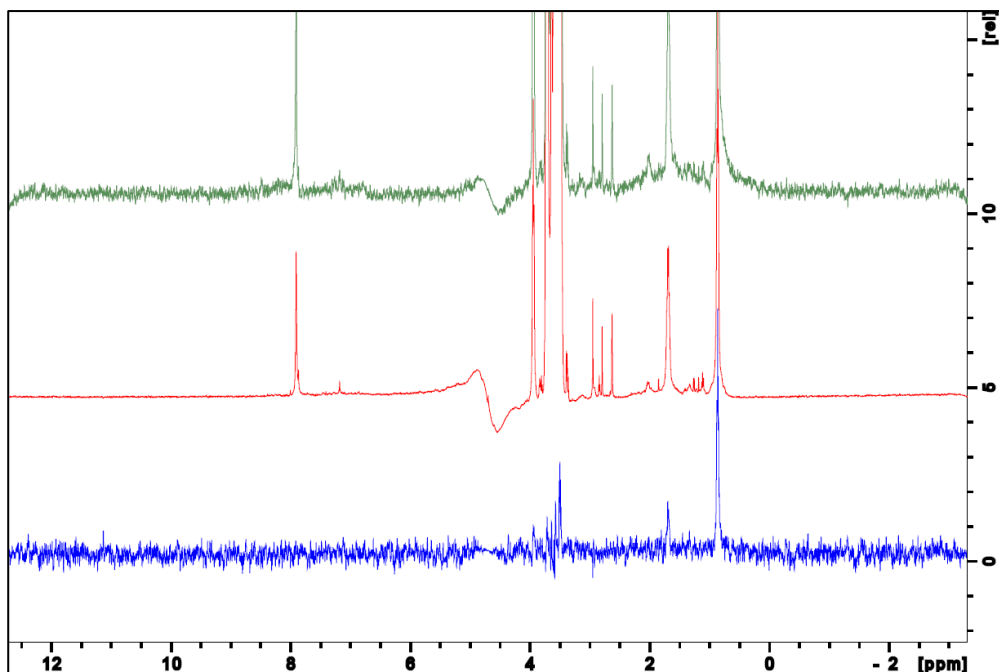
Appendix 49 - Ligand 4-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 4 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



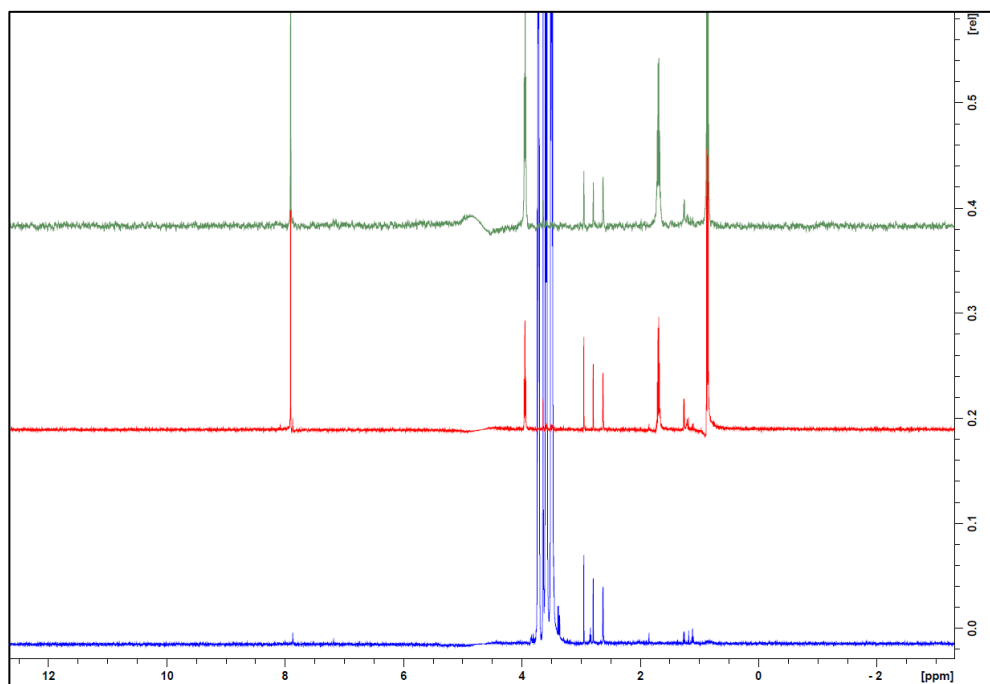
Appendix 50 - Ligand 6-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 6 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



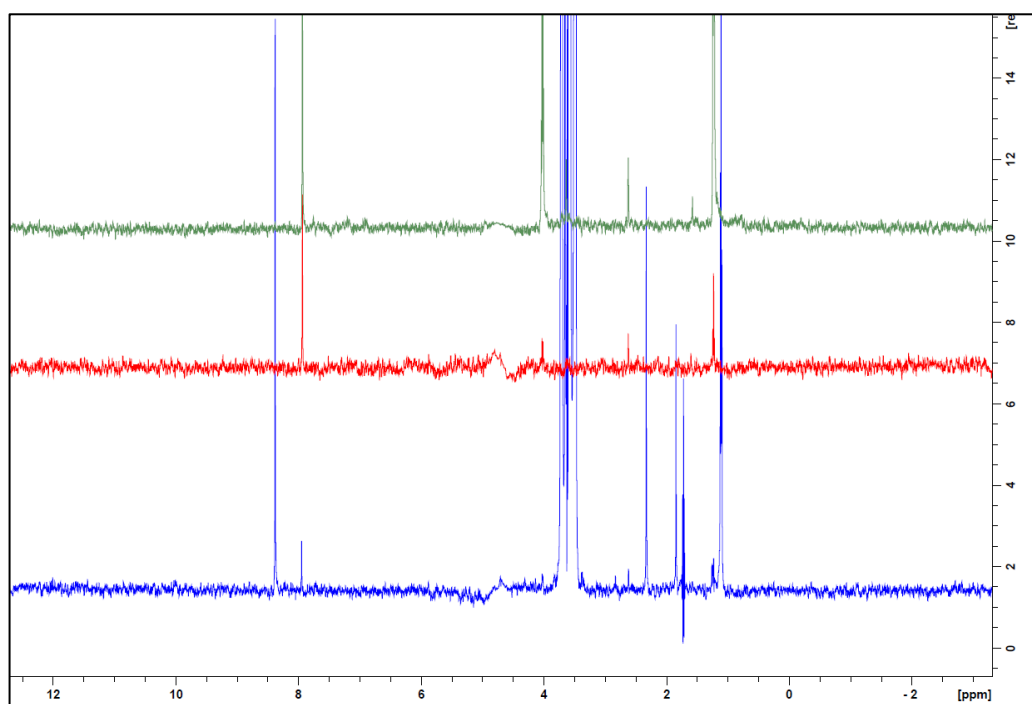
Appendix 51 - Ligand 6-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 6 on-resonance control spectrum (Middle), difference spectrum (Bottom)



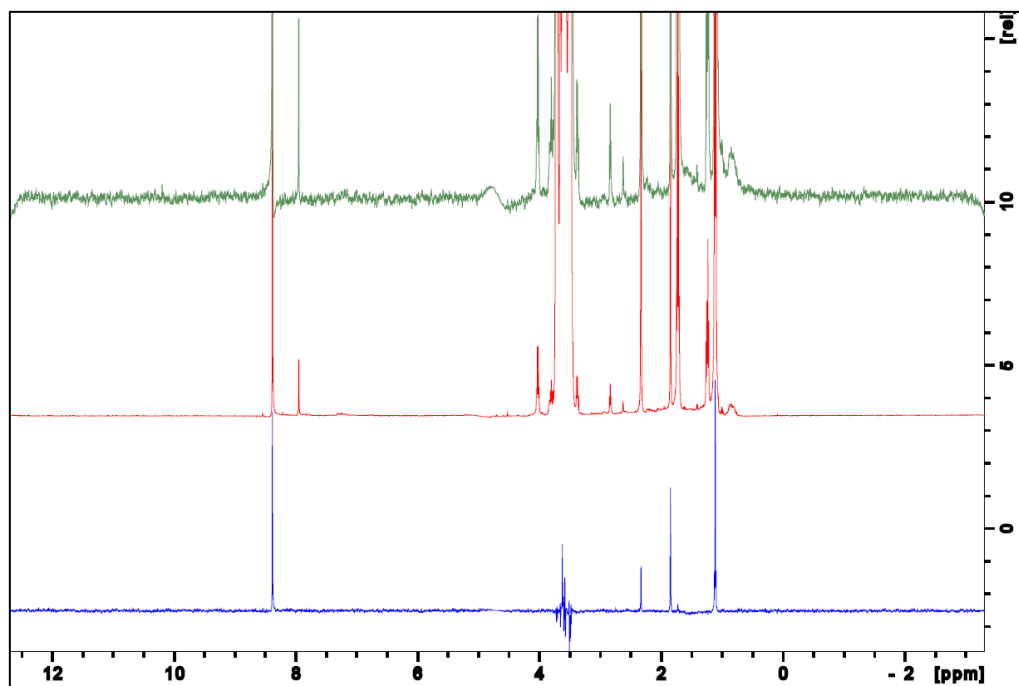
Appendix 52 - Ligand 6-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 6 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



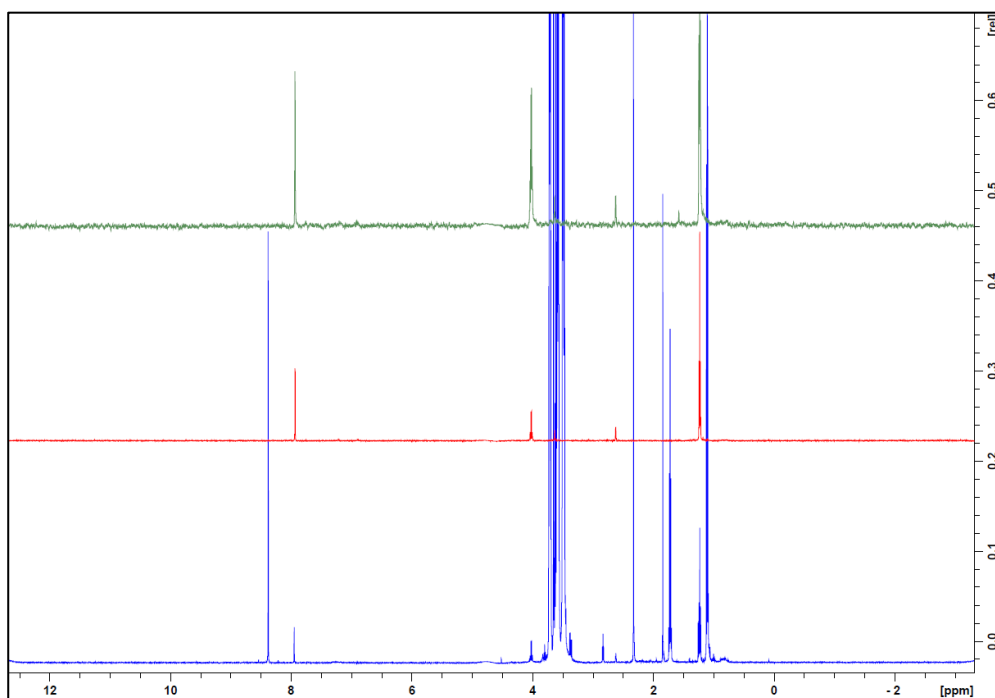
Appendix 53 - Ligand 9-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 9 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



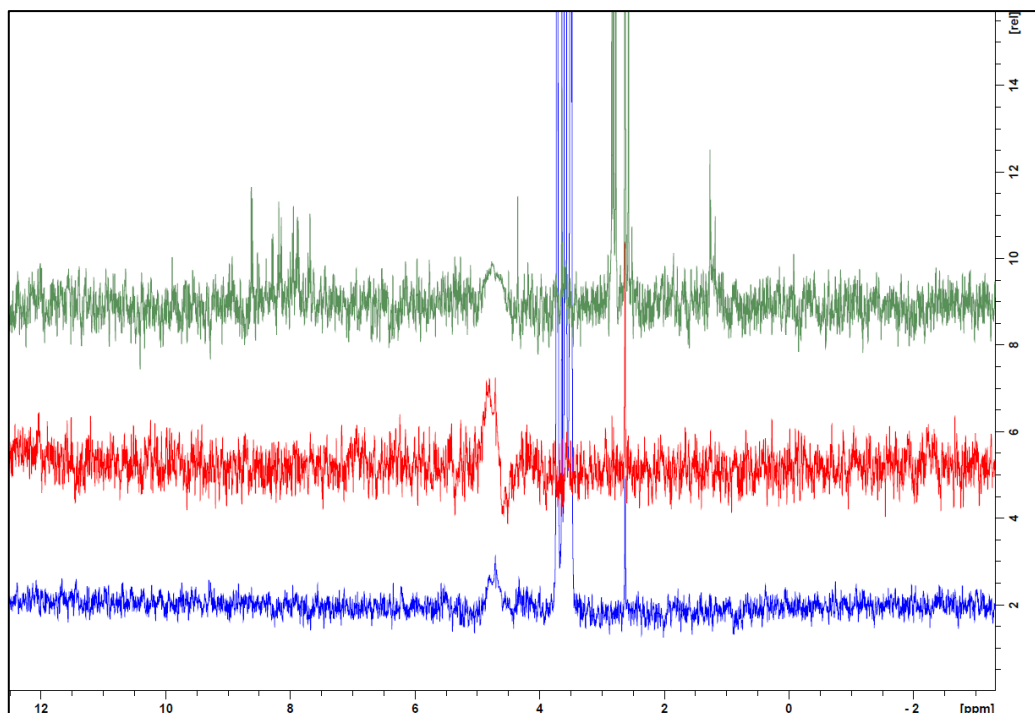
Appendix 54 - Ligand 9-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 9 on-resonance control spectrum (Middle), difference spectrum (Bottom)



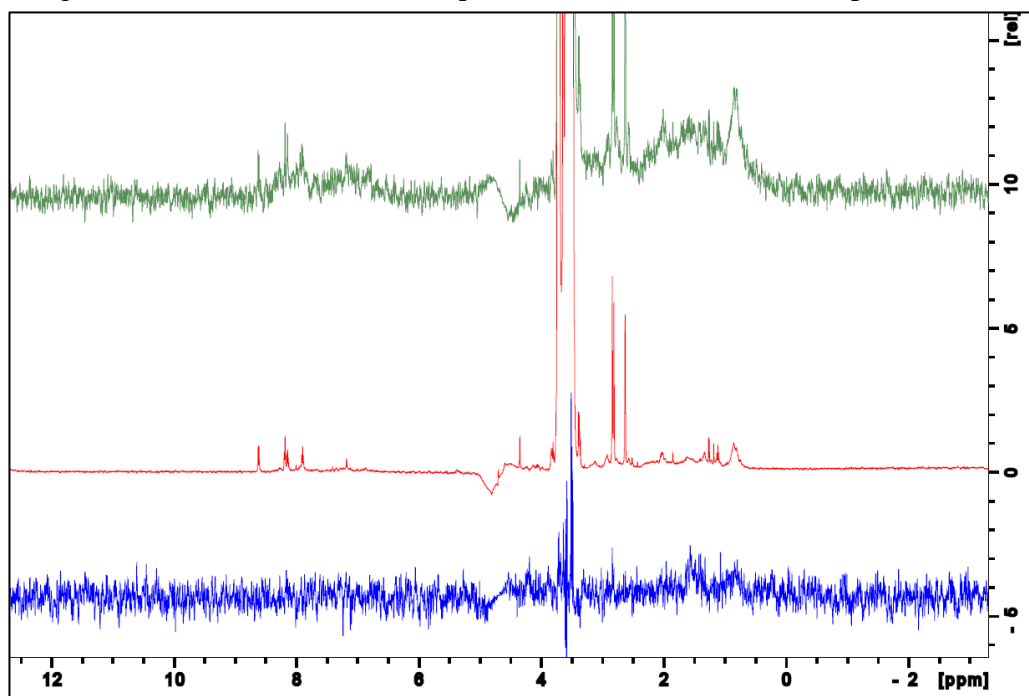
Appendix 55 - Ligand 9-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 9 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



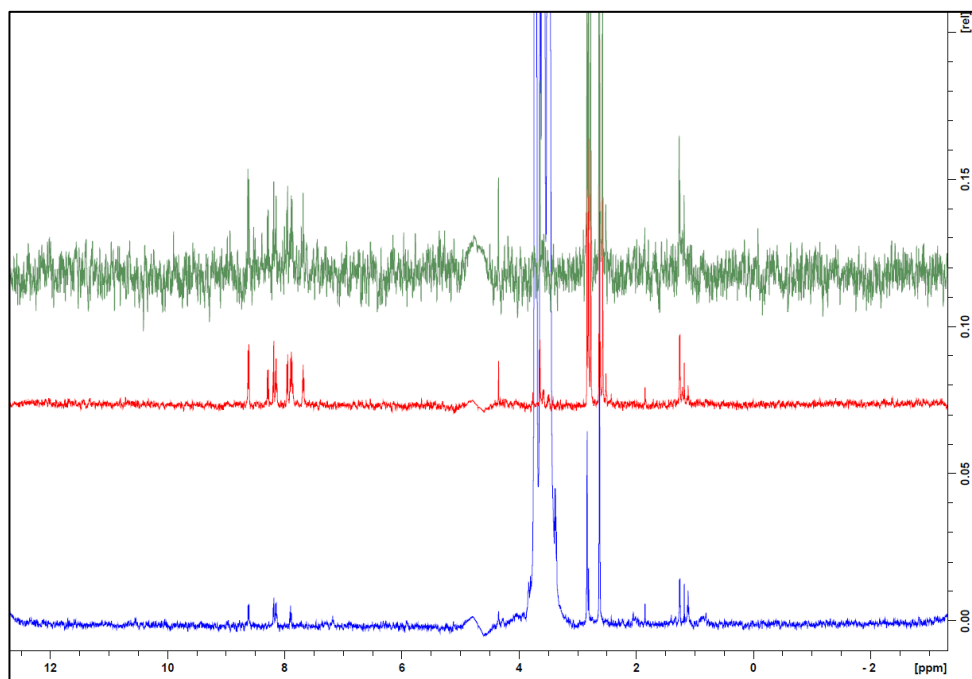
Appendix 56 - Ligand 12-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 12 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



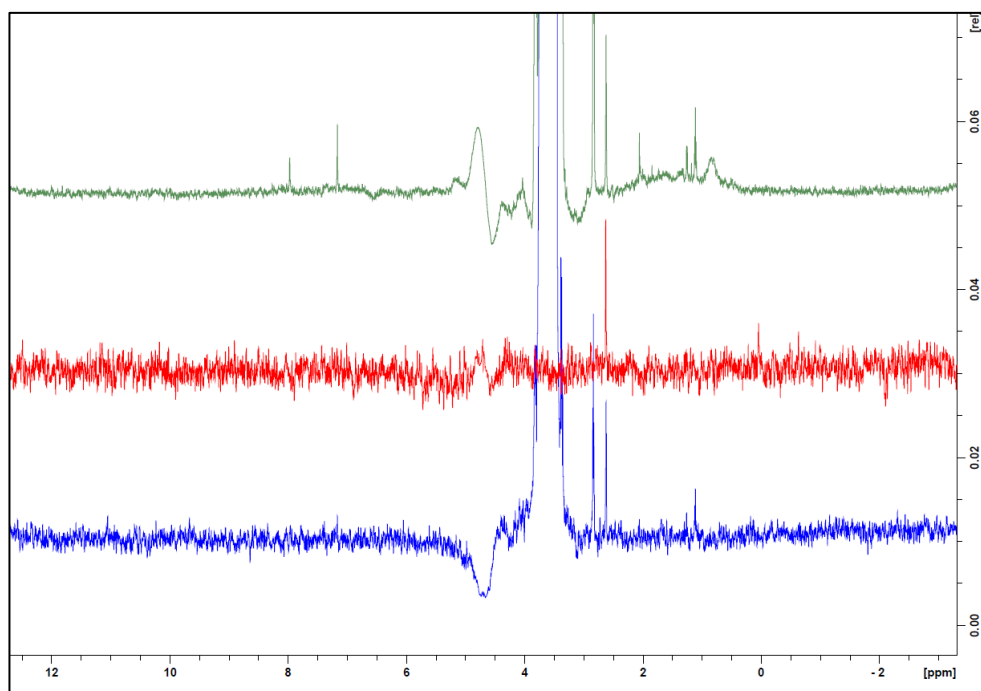
Appendix 57 - Ligand 12-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 12 on-resonance control spectrum (Middle), difference spectrum (Bottom)



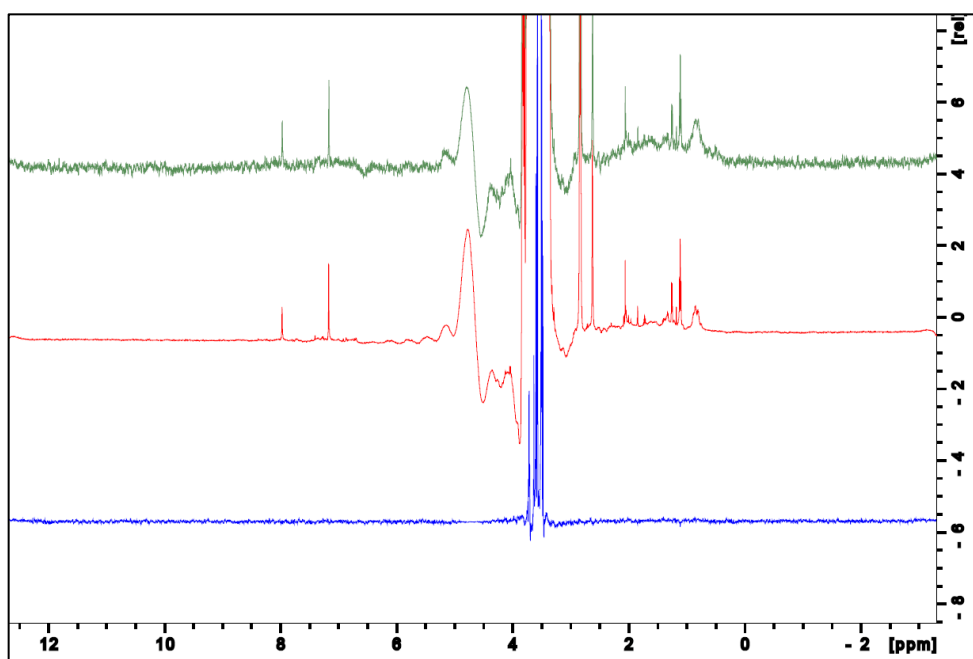
Appendix 58 - Ligand 12-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 12 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



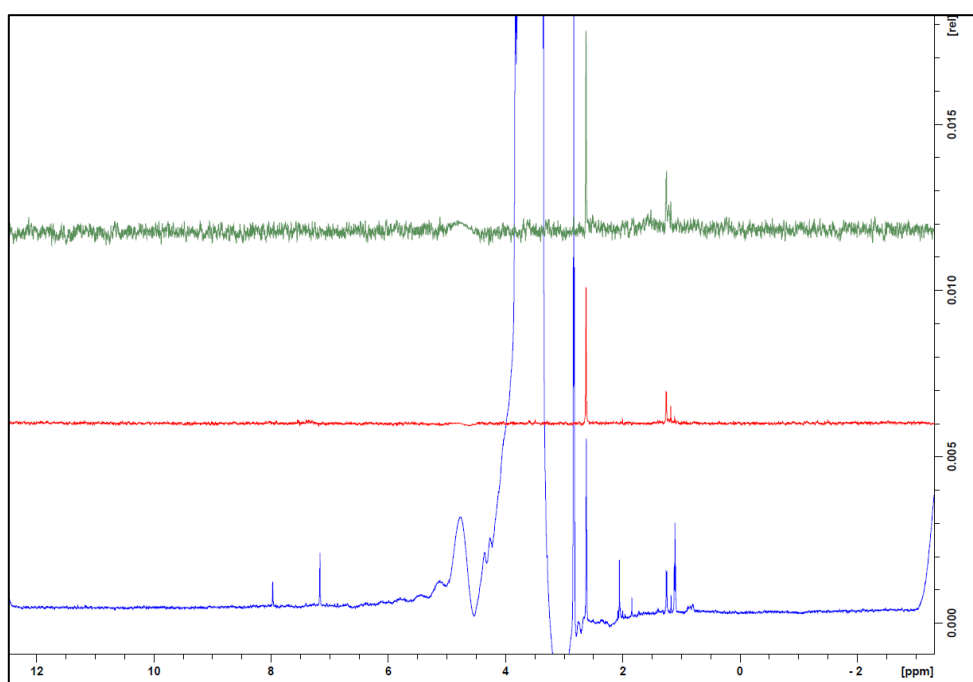
Appendix 59 - Ligand 17-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 17 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



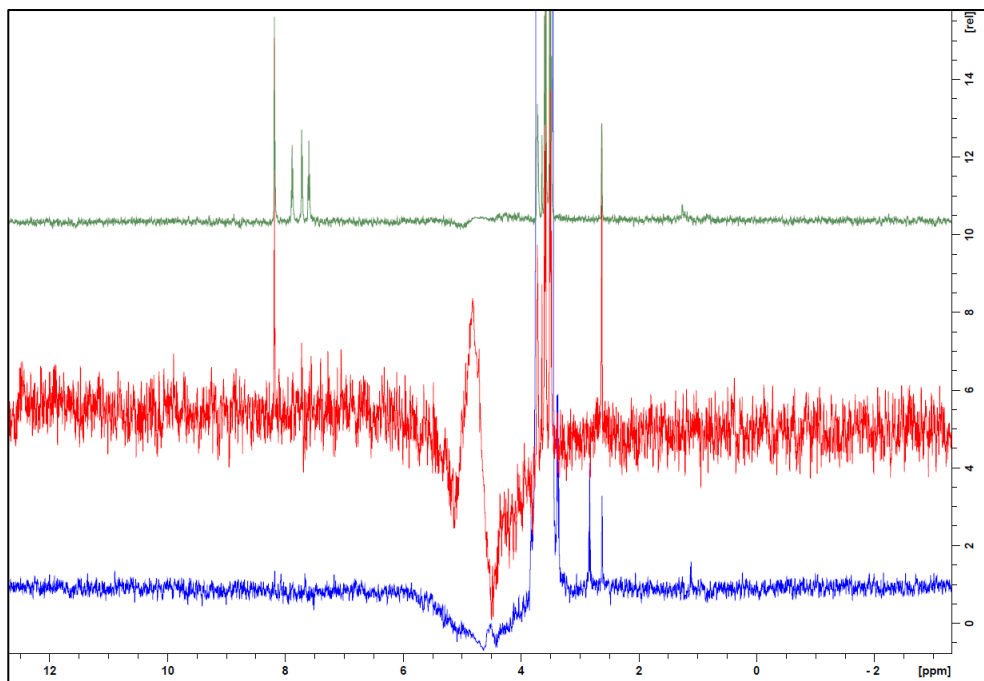
Appendix 60 - Ligand 17-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 17 on-resonance control spectrum (Middle), difference spectrum (Bottom)



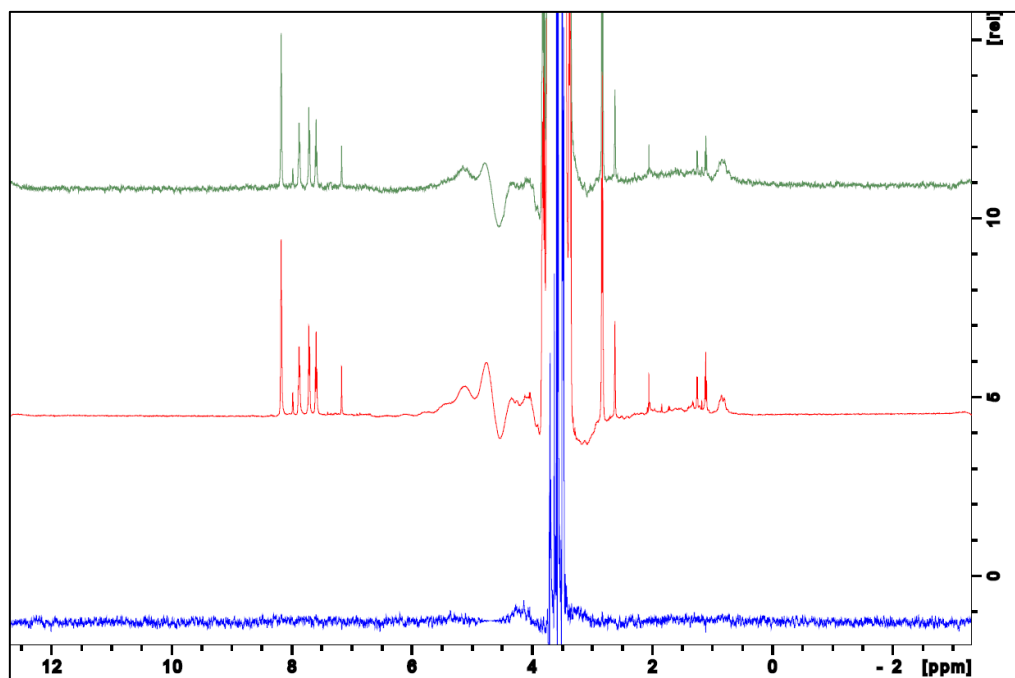
Appendix 61 - Ligand 17-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 17 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



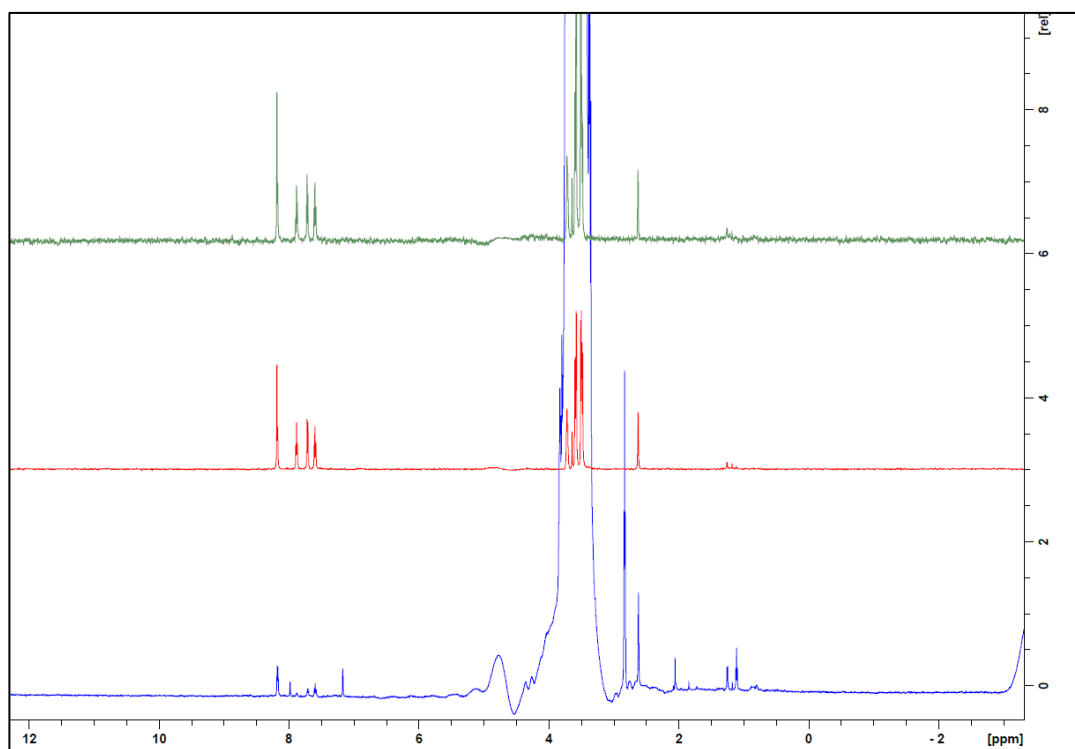
Appendix 62 - Ligand 22-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 22 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)



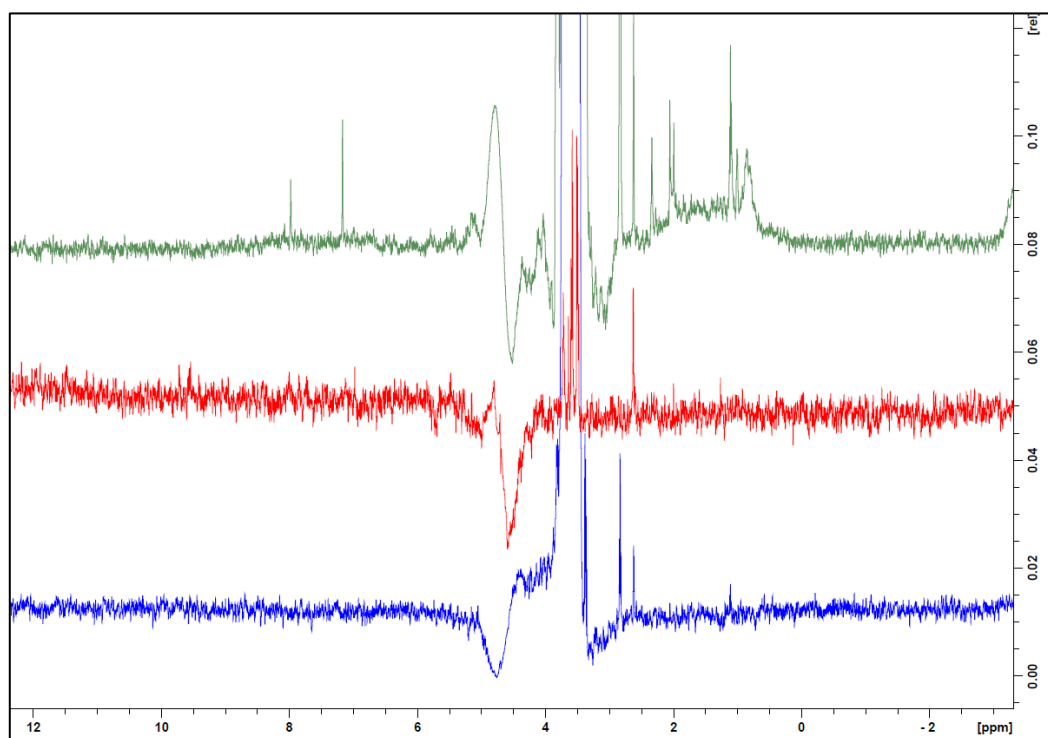
Appendix 63 - Ligand 22-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 22 on-resonance control spectrum (Middle), difference spectrum (Bottom)



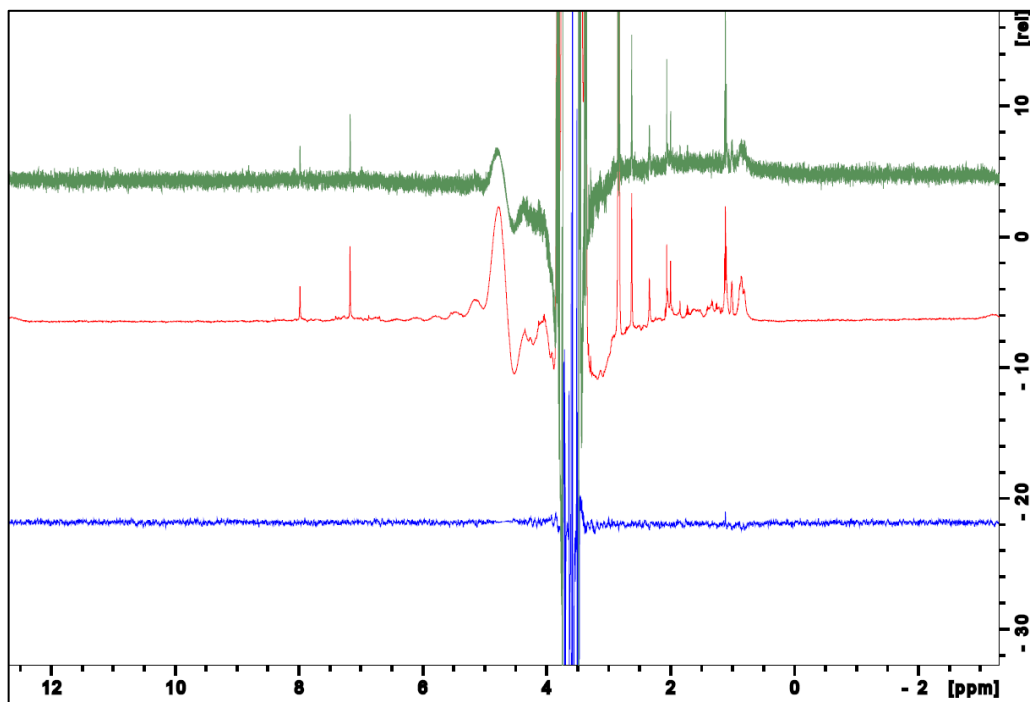
Appendix 64 - Ligand 22-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 22 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



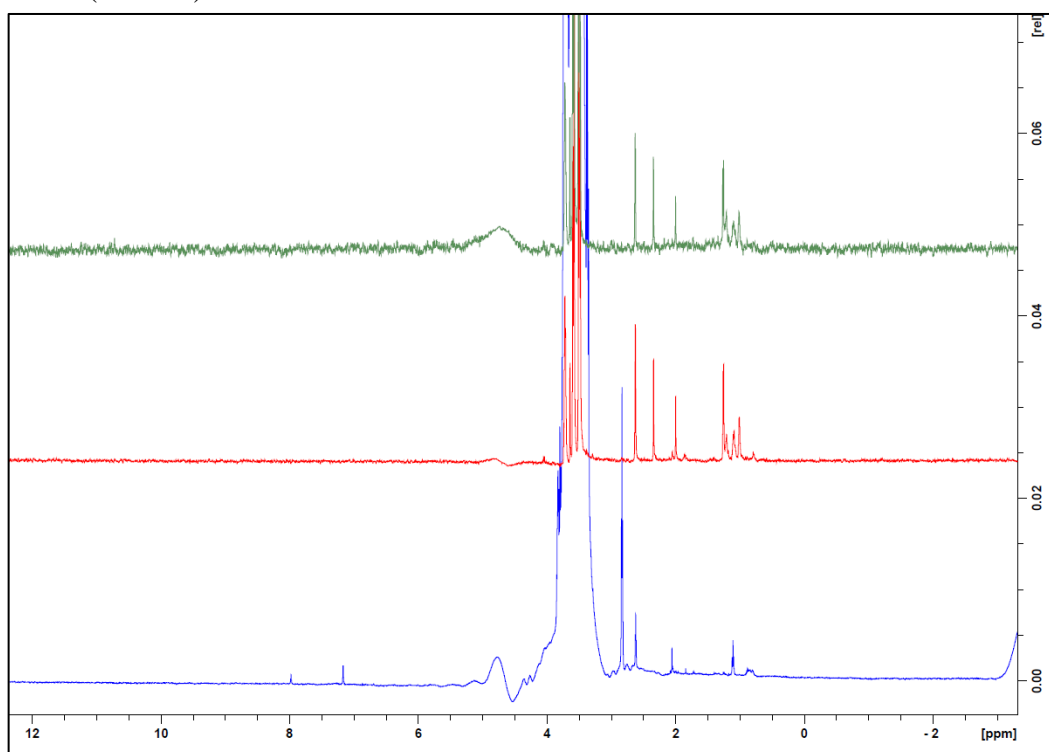
Appendix 65 - Ligand 23-TbrPDEB1 waterLOGSY NMR Spectra, Reference Proton 1D Spectrum (Top), Ligand 23 Control WaterLOGSY spectrum (Middle), Ligand-Protein WaterLOGSY spectrum (Bottom)

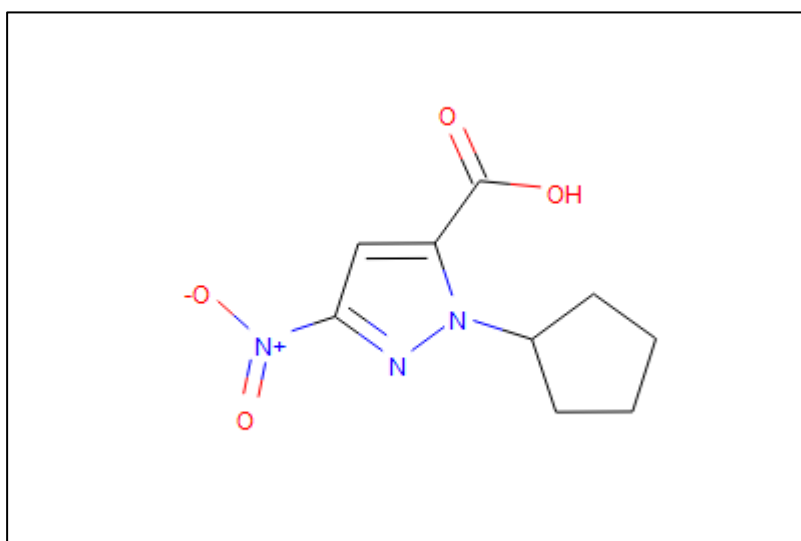
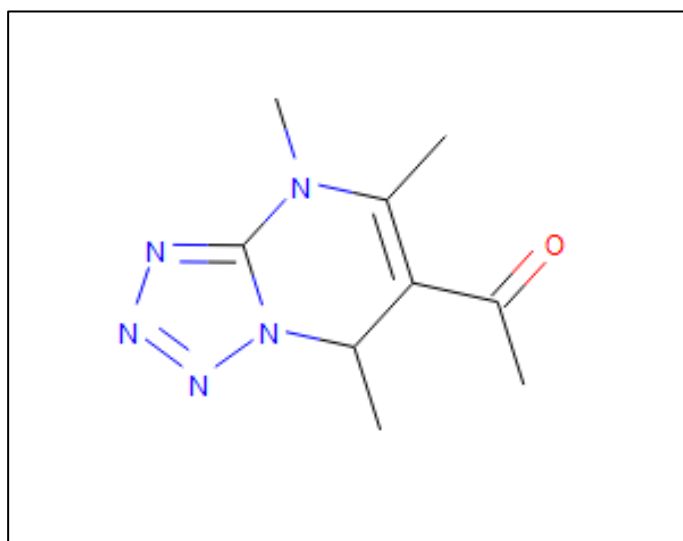


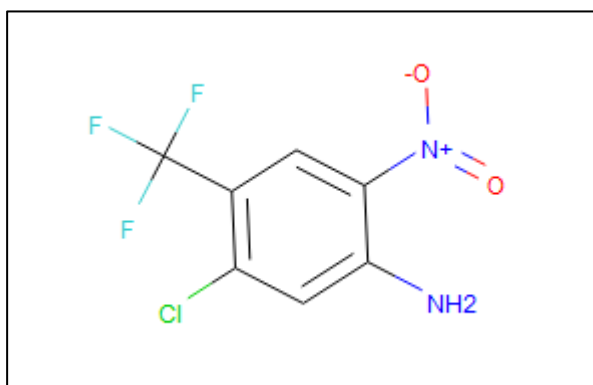
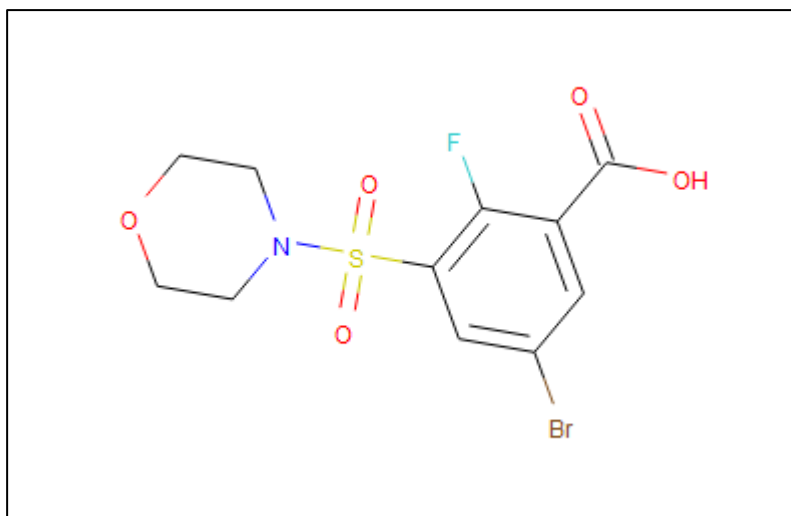
Appendix 66 - Ligand 23-TbrPDEB1 STD NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 23 on-resonance control spectrum (Middle), difference spectrum (Bottom)



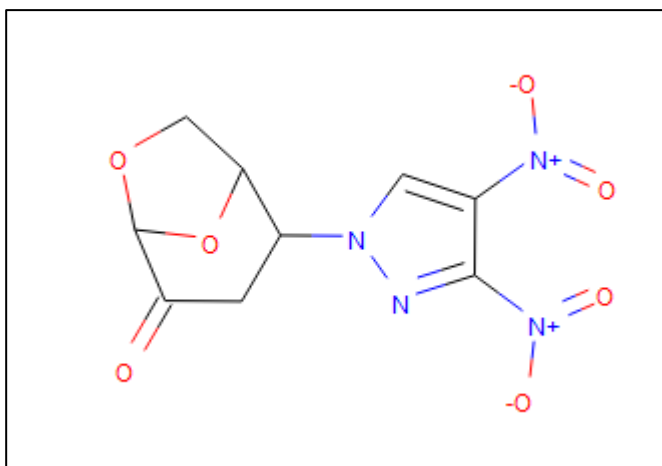
Appendix 67 - Ligand 23-TbrPDEB1 CPMG NMR Spectra, Reference Proton 1D spectrum (Top), Ligand 23 Control CPMG spectrum (Middle), Ligand-Protein CPMG spectrum (Bottom)



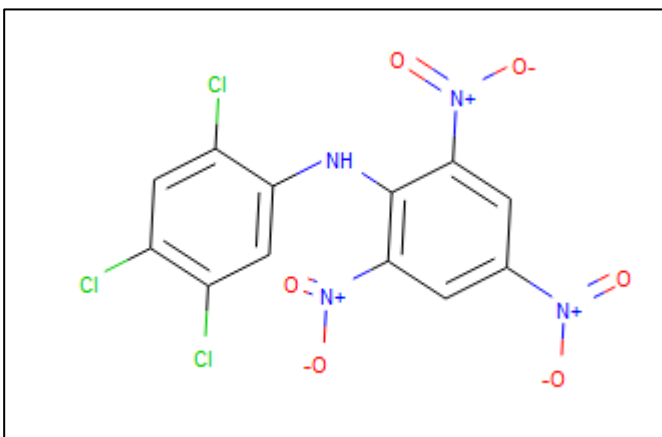




Appendix 72 – IB_STOCK1S-01606:VM_STL320887 ROCS Similarity search,
ComboScore 1.355



Appendix 73 – TT-ST51255295 ROCS Similarity search, ComboScore 1.323



Appendix 74 – CD_8010-5614:TT_ST4055854:VM_STK762870 ROCS Similarity search,
ComboScore 1.321

