The design and synthesis of a novel phosphatase probe to investigate potential new drug targets in *Leishmania Mexicana*

By Luc Robinson

1. Introduction:

The aim of this project is to advance our understanding of the disease leishmaniasis, which is a deadly parasitic infection. This project will be broken up into the following sections: Section 2 of this report will cover a review of the literature and give the scientific background to the research that was carried out. Section 2.1 will first cover the concept of neglected tropical diseases (NTDs) to give an understanding of the pressing global health challenge that leishmaniasis represents. Next, the two main approaches to the process of drug discovery will be outlined (section 2.2), including the use of chemical probes in that context (section 2.3). Section 2.4 will explain the biological concept of phosphorylation and its importance for the survival of the organism. Section 2.5 will then describe the role of kinases and phosphatases in this process, and how this makes them attractive drug targets. Section 2.6 will cover the chemical basis for the design of the probe, and finally the aims of the project in this context will then be explained in section 2.7

The strategies employed in the synthesis of the probe and its initial testing will be explored in section 3, while the results are presented in section 4. In section 5, the data and its implications will be discussed. Finally, the conclusions drawn from this research will make up section 6.

2. Review:

2.1 NTDs and Leishmaniasis

The World Health Organization has compiled a list of 20 infectious diseases that mainly affect the world's most vulnerable people, and cause devastating health, social and economic consequences; these are known as the neglected tropic diseases (NTDs). Combined, they bring misery and suffering to over 1/6th of the world's population.[1] Of these, the parasitic infection Leishmaniasis has one of the highest disease burdens.[2] It is caused by a family of parasites of the genus *Leishmania.* The disease can express itself in multiple forms, with the most common being mucosal, visceral and cutaneous.[3] Endemic to most parts of the developing world, there are estimated to be greater than 1 million new infections annually, with those affected by poverty, malnutrition, displacement, and poor housing most at risk.[4, 5]

Current treatments are inadequate; being highly toxic, require painful daily injections, have low efficacy and are expensive.[6, 7] Recently, increasing resistance to the few available medications has been observed.[8] Thus, it is clear new, safe, topical or oral, well tolerated, and affordable treatments are needed. However, low commercial return means that it receives little attention from pharmaceutical companies. Additionally, the high failure rate of drug discovery programmes discourages commercial companies from investing. Academic research can help de-risk such investments, making it more likely that pharmaceutical companies will pursue developing new drugs.

2.2 Phenotypic vs Target-Based Drug Discovery

Traditionally there have been two main approaches to drug discovery: phenotypic and target based.[9] The former involves the identification of potential drugs by observing their direct effect on organisms, without knowledge of the mechanisms by which they work. In contrast, target-based drug discovery relies on first identifying a molecular target, often a protein, associated with a disease. Target-based approaches are often simpler to execute, faster, and less costly, with the added benefit of enabling researchers to gain a better understanding of the disease, aiding in the development of structure-activity relationships and identification of new biomarkers.[10]

2.3 Introduction to Chemical Probes

When it comes to the process of target-based drug discovery small chemical compounds (also known as molecular probes) can used to understand protein function and help researchers determine whether a drug is binding to a target as intended.[11] They are made up of a 'warhead' consisting of a reactive functional group, and a reporter tag connected via a linker. The 'warhead' is what interacts with the active site of a protein while the linker helps identify the location of the probe, usually through fluorescence.

2.4 Phosphorylation

The genetic code for all living things encodes the information needed to produce proteins, which play many important roles and are vital to the survival of the organism. However, the picture is more complex. Once synthesised, many proteins are modified to attenuate their activity - this is known as post-translational modification. Of these, protein phosphorylation is one of the most prevalent.[12] Dysfunction of this process in humans has been implicated in many diseases such as Huntingdon's, cancer and heart disease.[13-15] The leishmania parasite is also vulnerable in this regard: dysregulation can have wide ranging effects, for instance preventing it from progressing in its cell cycle, ultimately leading to its death.[16] Thereby, if a compound were to be found that was able to disrupt the addition or removal of phosphate groups, it would represent an exciting novel antileishmanial compound.

2.5 Kinases and Phosphatases

The phosphorylation of many molecules, including proteins, is often achieved by another class of proteins known as kinases.[17] As they often contain highly conserved regions, kinases have long been identified as promising drug targets in a variety of diseases, including leishmaniasis.[18] The leishmanial kinome has been relatively well characterised, enabling the development of novel antileishmanial compounds.[19] However, there exists another class of proteins, the phosphatases, whose role in biological organisms is the dephosphorylation of proteins. Thus, phosphatases can be seen to have an equally important, yet antithetical role to that of kinases. Despite this, they receive much less research attention, and indeed much less is known about the leishmanial phosphatome.[20]

Phosphatases are a diverse set of proteins. In humans there are 6 main 'super-families', each characterised by a unique domain, of which the protein tyrosine phosphatase (PTP) family is the largest.[21] PTPs catalyse the dephosphorylation of tyrosine residues and themselves can

be subdivided into 4 further groups, 3 of which are based on cysteine and 1 which is based on aspartic acid.[22] Many of these PTPs are highly conserved between different species making comparisons to PTPs in the leishmania parasite possible.[22]

2.6 Design of the Probe

In previous work on human PTPs, a number of simple electrophilic aromatic derivatives have been identified as non-specific inhibitors of the proteins.[23] Thus, it was decided that this would be a good starting point for the investigation of leishmanial PTPs. As a proof-of-concept investigation, the lack of specificity was seen as a positive, given that the more PTPs targeted, the more widely applicable the probe would be to further investigations. Hence, the decision was made to create a derivative of an established human PTP probe. Modifications were made to the design of the probe to make it compatible with established 'click-chemistry', which would allow for a fluorescent tag to be added.

Mechanistic studies have shown that PTPs react with aryl vinyl sulfones through their active site thiol found in the cysteine residue, and that reactions occur through a mechanism known as conjugate addition.[24] This is important as it means that a molecule without this reactive vinyl functionality would not be expected to undergo specific binding to a PTP, leishmanial or otherwise.

As many proteins can have a wide range of interactions, a control probe lacking the reactive functional group is often used as a control to help determine whether any observed interactions between the active probe and proteins are due to non-specific binding, which can complicate the data interpretation.

2.7 Aims of the Project

The aim of this project was to develop an active probe that selectively binds to PTPs in *Leishmania Mexicana*, in addition to a control probe (see figure 1). This would facilitate further research into PTPs as potential drug targets, bringing us closer to a cure for this disease.

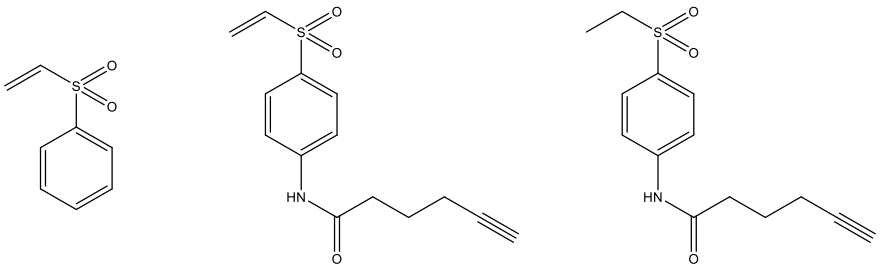

Figure 1: Established human PTP probe (Left) and proposed active leishmanial PTP probe (Middle) and proposed leishmanial PTP control probe (Right)[23]

3. Strategies

A variety of standard chemical techniques were employed in this project during the synthesis of the probes:

Each reaction was monitored by thin layer chromatography (TLC) using a hexane, ethyl acetate mixture in order to determine when the reaction was complete. This was done by taking a sample of reaction mixture at regular intervals and performing TLCs until no spot was visible for the starting material (see figure 3). Most products and reagents were UV active and hence could be seen under a source of UV light, however some weren't and a $KMnO_4$ stain had to be used to observe them.

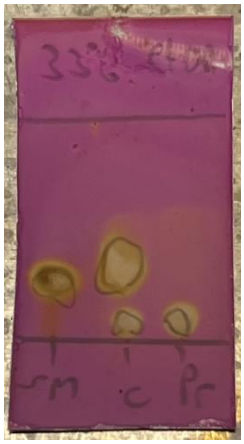

Figure 2: TLC of the crude reaction mixture of reaction 2

Liquid chromatography mass spectrometry (LCMS) was used to analyse the crude reaction mixture to determine whether the correct product had formed. It was also used to give an initial insight into the purity of the product (see figure 4). Once purified, LCMS was also used to characterise the product (see figure 5).

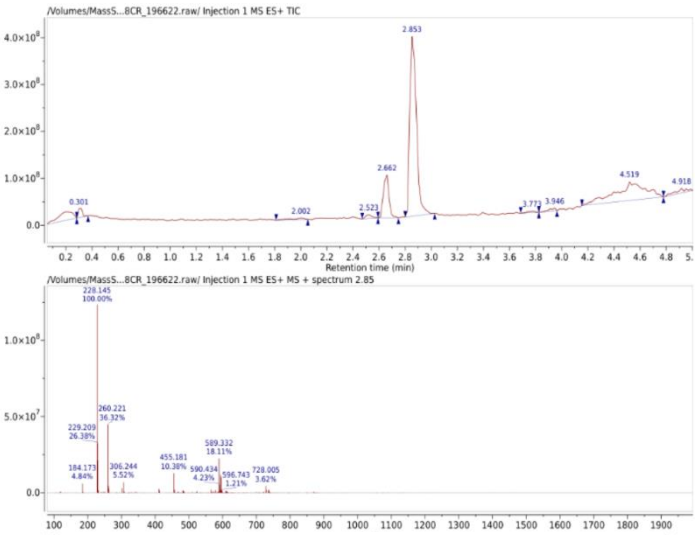

Figure 4: LCMS spectra of the crude product of reaction 4

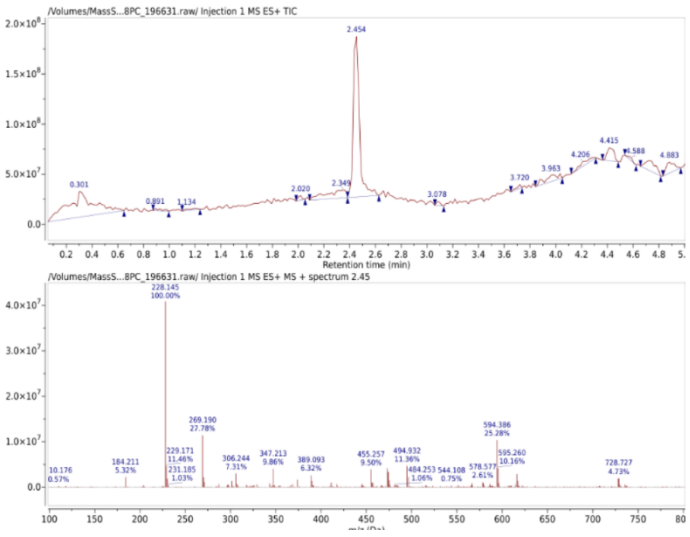

Figure 3: LCMS spectra of the purified product of reaction 4

Similarly, nuclear magnetic resonance (NMR) imaging was used to characterise the purified product (see figure 6). If LCMS of the crude was inconclusive, and the sample was of a high enough purity, NMR was also used (see figure 7).

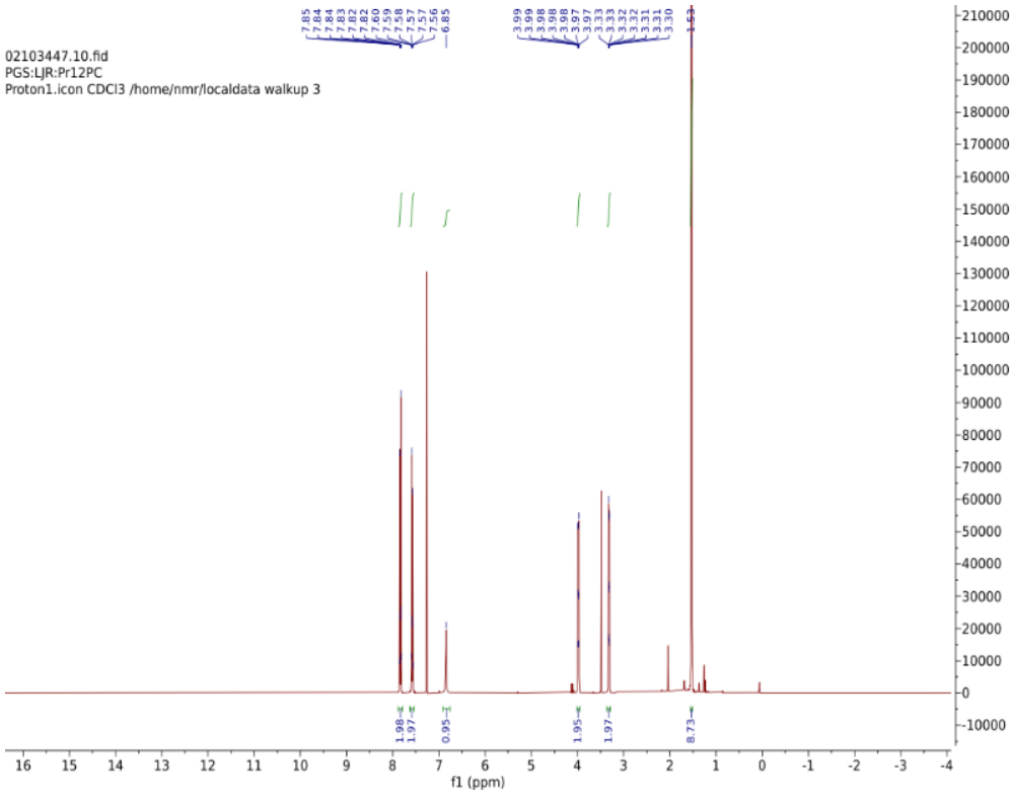

Figure 5: NMR spectra of the purified product of reaction 3

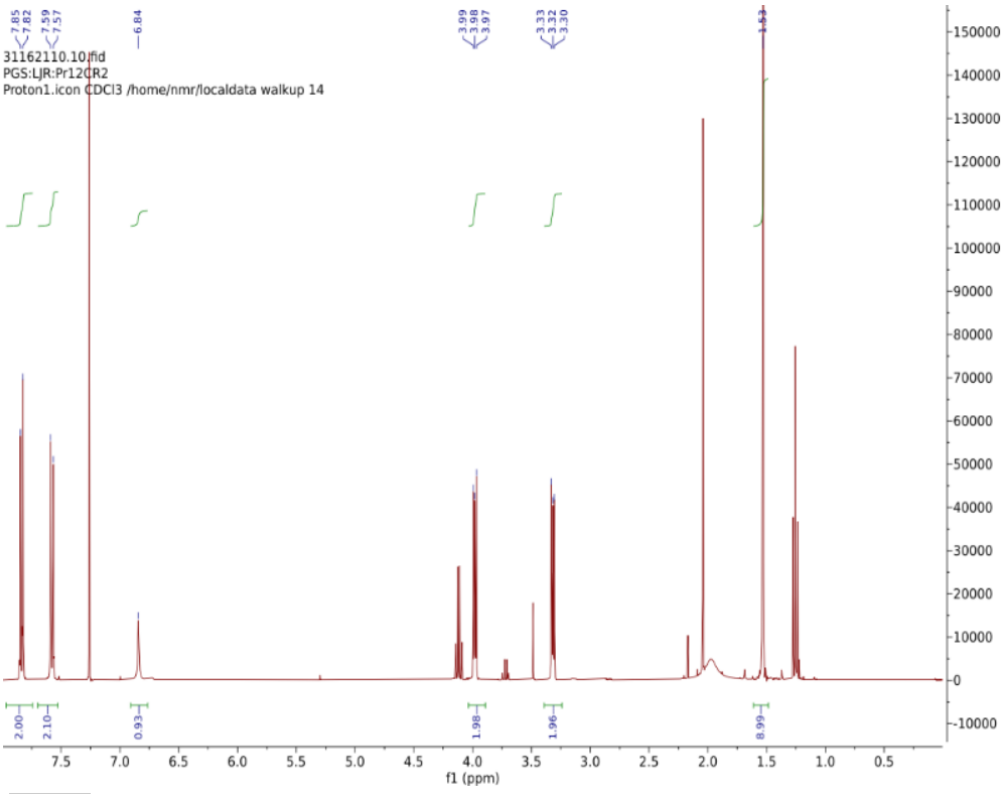

Figure 6: NMR spectra of the crude product of reaction 3

If the crude product contained impurities, it was purified by normal phase flash column chromatography using an ethyl acetate petroleum ether mixture (see figure 9). If this was not sufficient then a reverse phase column was subsequently performed using a water methanol mixture (see figure 10). TLC was used to determine the optimate initial solvent ratio.

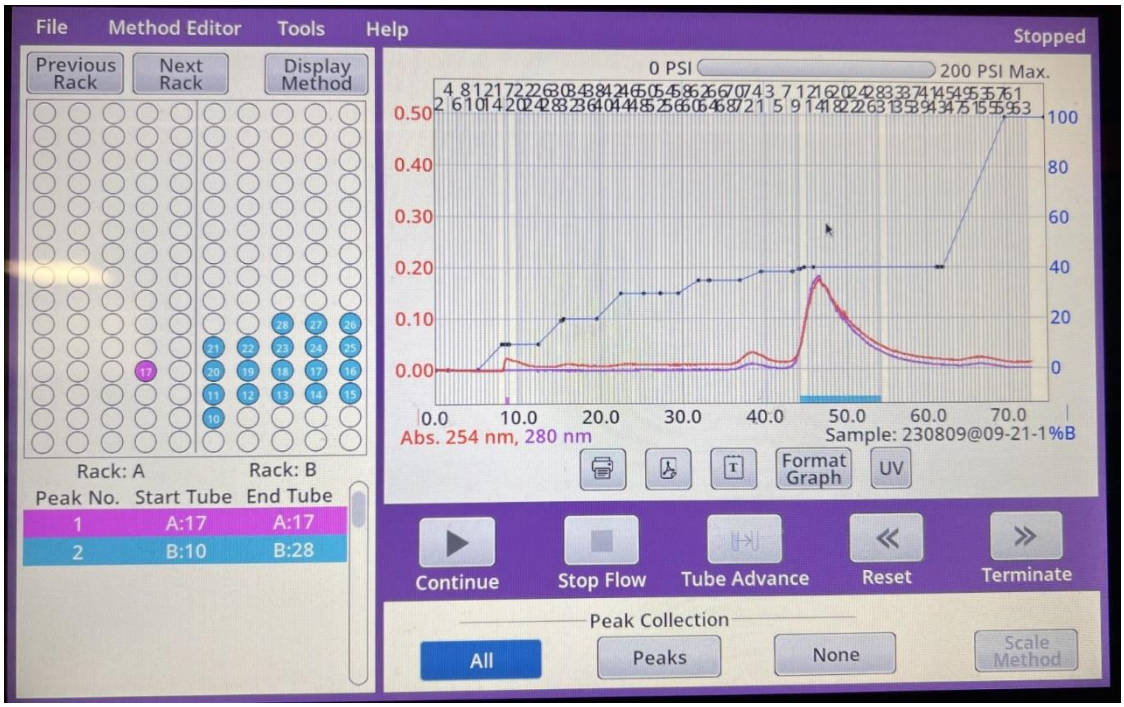

Figure 7: Results of normal phase flash column chromatography of the crude product from reaction 7

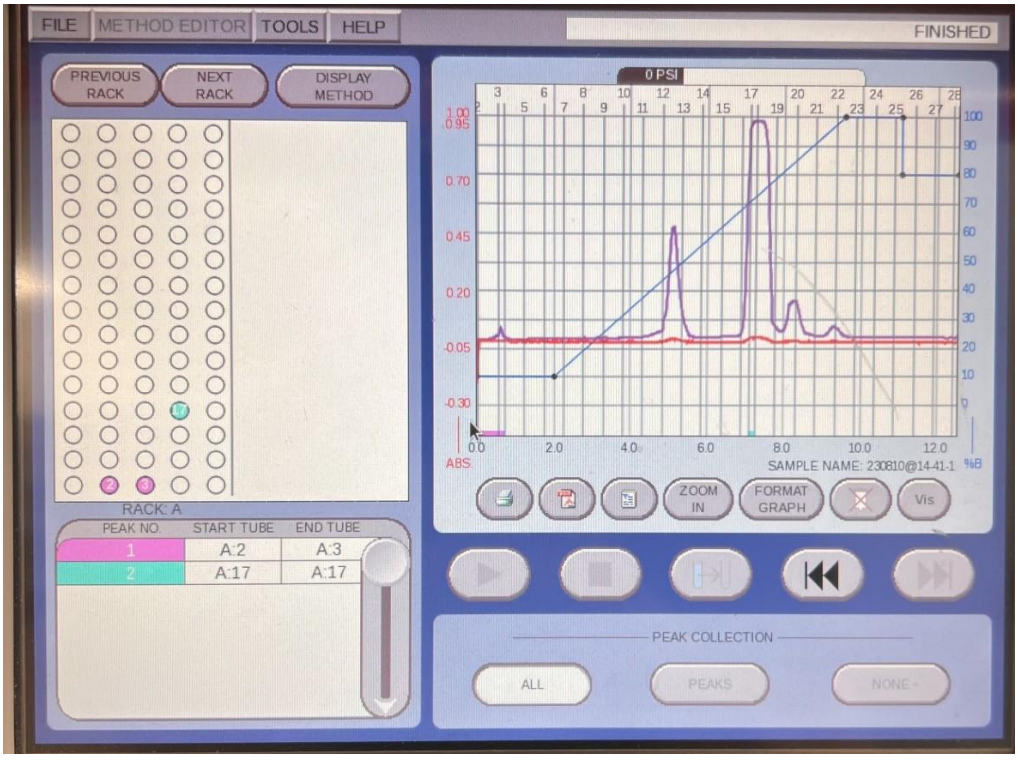

Figure 8: Results of inverse phase flash column chromatography of the crude product from reaction 7

4. Results

Figure 2 summarises the reactions used to produce the control and active probes.

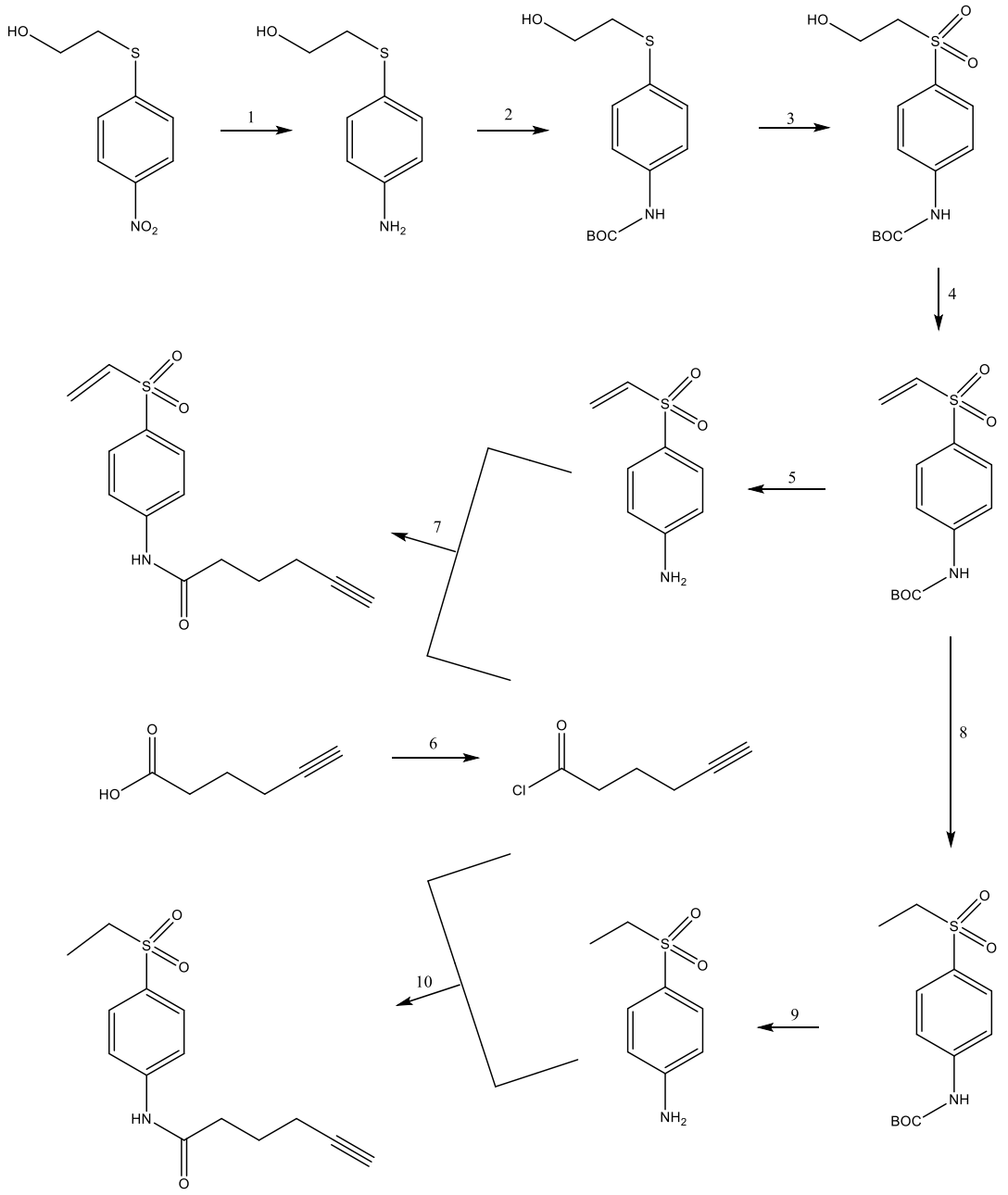

Figure 9: The synthetic scheme for the formation of the active leishmanial PTP probe and PTP control probe

2-Hydroxyethyl 4-nitrophenyl sulfide was identified as a commercially available starting material. Its nitro group was reduced using a stannous(II) chloride catalyst in ethanol to produce a para-substituted aniline (1).[25] The newly formed aniline was protected by a boc group through reaction with boc anhydride using triethylamine (2).[26] Oxidation of the sulphur atom by hydrogen peroxide, catalysed by tungstate, yielded the sulfone (3).[27] Mesylation of the alcohol using mesylchloride and triethylamine, followed by spontaneous dehydration, produced the aryl vinyl sulfone (4), which was then deprotected using TFA, yielding the free amine (5).[26, 28] The acyl chloride was produced in parallel from 6-hexyonic acid and oxalyl chloride, with a catalytic amount of DMF (6).[29] The aryl vinyl sulfone and acyl chloride where finally coupled using triethylamine to produce the active probe (7).[30] Some of the aryl vinyl sulfone was also reduced using hydrogen gas and a 10% Pd/C catalyst (8), before being deprotected with TFA (9).[26, 31] This aryl sulfone was then coupled with the acyl chloride using triethylamine, yielding the control probe (10).[30]

Each stage of the chemical synthesis was successful with yields slightly below what is found for similar reactions in the literature, with the key exceptions of the amide coupling reactions (see Table 1). Experimental yields for reactions 5, 6 and 9 were not obtained as these compounds were reacted directly on without purification.

Table 1: The experimental yields of different stages in the synthesis of the active and control probes

Reaction Number	Experimental Yield
1	92%
2	63%
3	80%
4	78%
7	31%
8	98%
10	49%

Analysis of the NMR and LCMS data shows that the purity of most compounds synthesised was good, with the exception of reaction. In any case, the purity was deemed sufficient to progress with the biological assay of the leishmanial proteome using the two probes.

A fluorescent tag was then added to the probes using ‘click chemistry’. Specifically, a copper-catalysed azide-alkyne cycloaddition (CuAAC) reaction was used.

Protocols previously established by Professor Steel’s lab group were then used to treat the parasite *in vivo*, extract the proteins and analyse the interactions of the probes with the parasite proteome.[32] The latter stage was achieved through a technique known as SDS-PAGE that separates proteins based on size. A fluorescent scanner was then used to visual the protein bands in the labelled proteome.

The results from the SDS-PAGE experiment are shown in figure 10. Wells 1, 2, 6 and 10 were empty. Well 3 contained what’s known as a protein ladder: a mixture of proteins of known molecular weights that serves as reference points on the gel. Wells 4 and 7 contained the solution of control probe and leishmanial proteins, while wells 5 and 8 contained a solution

of the active probe and leishmanial proteins. Each band in the gel corresponds to one or more proteins, thus it is clear that a substantial number of proteins in the parasite were bound to the active probe. In contrast, the absence of bands in wells 4 and 7 indicate that no proteins in the parasite were bound to the control probe. Due to the poor resolution of the machine, it is difficult to the determine accurately the number of bands present or their intensity. The results also show that proteins of a wide range of masses were targeted by the active probe.

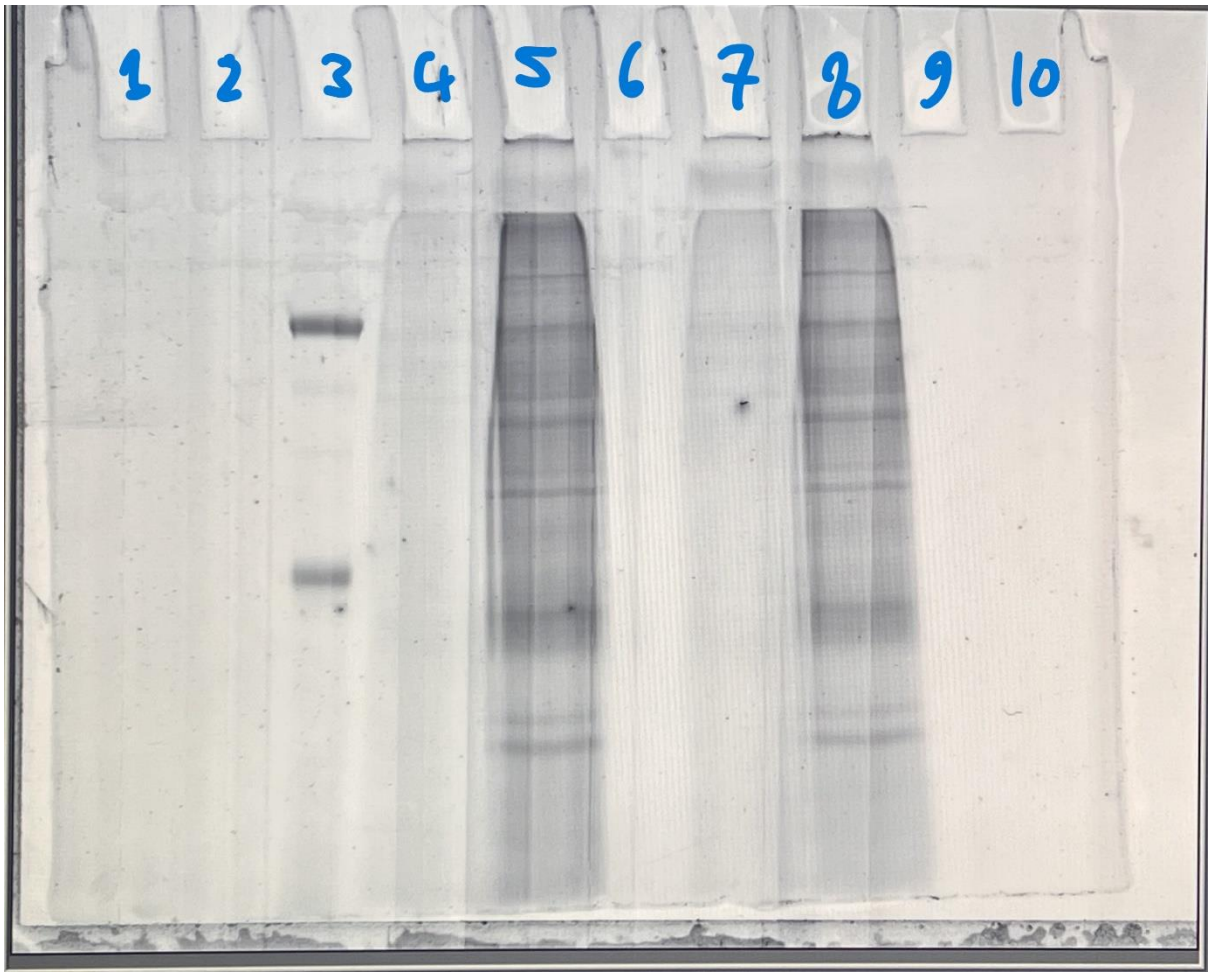

Figure 10: The fluorescent readout of the gel from the SDS_PAGE analysis of the leishmanial proteome after incubation with the control and active probes

5. Discussion:

The objective of successfully synthesising both probes was met, and all steps worked as expected. The yields of the reactions were broadly in line with similar examples in the literature, with the exception of the amide coupling reactions (reactions 7 and 10). This is potentially due to the sensitivity of acyl chlorides to hydrolysis by trace amounts of water. Inexperience with the handling of such sensitive compounds lead to errors during the synthesis. Additionally, due the sensitive nature of the acyl chloride, it was difficult to determine when its formation had gone to completion. TLC and LCMS were not suitable methods, and the crude mixture had too many compounds in it to make NMR analysis feasible. Infrared (IR) spectroscopy would have been a suitable option, however no machine was available.

There were challenges when it came to the purification of the final products. This mainly stemmed from the 'sticky' nature of the unreacted started material in reactions 7 and 10, which meant it was difficult to remove from the crude mixture. This can be seen in figure 8 where it is the first peak. It started eluting from the normal phase column early on, but then continued to do so up until, and indeed past, the point at which the desired product started eluting. This meant there was significant contamination of the product with unreacted starting material. In order to further purify the compound, reverse phase flash column chromatography was used (see figure 9). This was very successful, and the product was isolated in good purity. To try prevent the same issue arising in the formation of the control probe, the method was adapted to include a wash of the crude product with 0.01M HCl. This was somewhat successful, with most of the unreacted starting material being extracted, however trace amounts remained that were not able to be removed by normal phase flash column chromatography. If an inverse phase column was performed it would have greatly helped increase the purity of the final compound. Unfortunately, time did not allow for this.

The results of the SDS-PAGE experiment are promising and meet the objectives of the project. The active probe successfully targeted a considerable number of proteins as seen by the many fluorescent bands in the gel, while none were targeted by the control probe (see figure 10). This reveals that interactions with these proteins are dependent on the vinyl functionality, implying a nucleophilic target, consistent with the hypothesis that this probe would target leishmanial PTPs. However, PTPs are not the only class of protein with an active site cysteine residue, thus further work is needed to characterise the proteins targeted by the probe. This can be achieved through protein mass spectrometry, either analysing individual bands from the gel, or the entirety of the labelled proteome at once using highly advanced bioinformatics tools.

If the proteins targeted were proven to be leishmanial PTPs, then further work would need to be done to determine the optimal incubation period for the parasite with the probe, the optimal concentration of probe, and its toxicity towards to parasite. This represents a significant scientific undertaking but is crucial for validating the probe.

Additional work could also be done to optimise the structure of the probe, to enhance its selectivity towards certain proteins. This would involve repeating the experiments outlined

in this paper with different synthetic analogues of the probe. For instance, the length of the linker (the part connecting the ring to the fluorescent tag) could be altered, extra groups could be added to attenuate the electronic characteristics of the structure, or the warhead could be replaced with a similar functionality (see figure 11). This further investigation would present many more months of work and was significantly beyond the scope of this project.

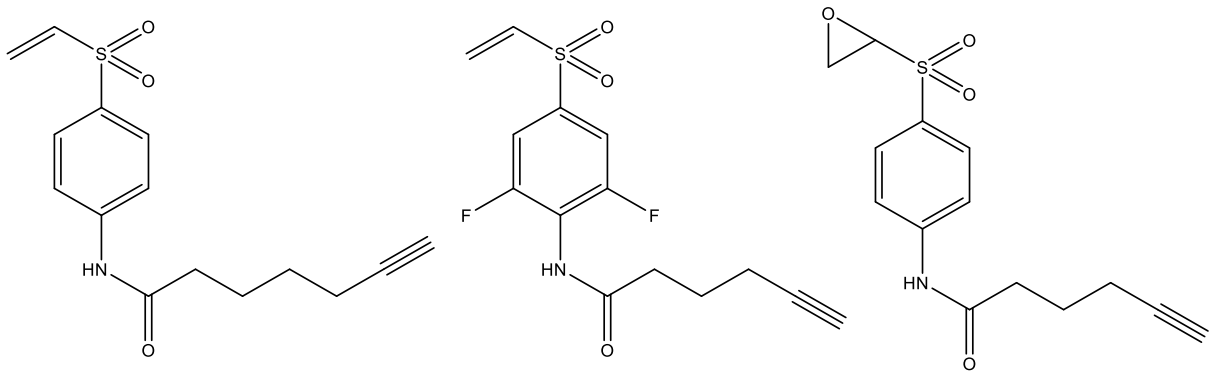

Figure 11: A probe with a modified linker (left), altered electronic characterisitics (middle), different warhead functionality (right)

6. Conclusions:

The aim of this project was to advance our understanding of the disease leishmaniasis. This was done by synthesising a novel leishmanial chemical probe based off previous work on human protein tyrosine phosphatases. It exhibited affinity for a number of proteins in the leishmanial proteome, and through the use of a control probe, it was demonstrated these interactions were solely via the reactive functionality of the warhead.

The findings of this project are the first step towards developing an effective leishmanial phosphatase probe. Further work can be done to determine what specific proteins have been targeted which is required to validate the probe for use in further proteomics studies. If it is proven that the proteins targeted are leishmanial PTPs, it will open the door for the investigation of new phosphatase drug targets using chemical proteomics experiments. Ultimately, this will lead to better understanding of the biology of the parasite, and in the long-term may lead to the development of a new class of life saving drugs.

References:

1. D. A. Álvarez-Hernández, L. Rivero-Zambrano, L. A. Martínez-Juárez and R. García-Rodríguez-Arana, *Ther Adv Infect Dis*, 2020, **7**, 2049936120966449.
2. C. J. Scheufele, R. L. Giesey and G. R. Delost, *J Am Acad Dermatol*, 2021, **84**, 1203-1205.
3. WHO Leishamaniasis, https://www.who.int/news-room/fact-sheets/detail/leishmaniasis, (accessed 19/08, 2023).
4. E. Torres-Guerrero, M. Quintanilla-Cedillo, J. Ruiz-Esmenjaud and R. Arenas, *F1000Research*, 2017, **6**.
5. J. Alvar, S. Yactayo and C. Bern, *Trends Parasitol*, 2006, **22**, 552-557.
6. eBioMedicine, *eBioMedicine*, 2023, **87**, 104440.
8. S. L. Croft, S. Sundar and A. H. Fairlamb, *Clin Microbiol Rev*, 2006, **19**, 111-126.
9. D. C. Swinney, *Clin Pharmacol Ther*, 2013, **93**, 299-301.
10. G. E. Croston, *Expert Opinion on Drug Discovery*, 2017, **12**, 427-429.
11. ChemicalProbes.Org, https://www.chemicalprobes.org/faq, (accessed 19/08, 2023).
12. S. Qausain, H. Srinivasan, S. Jamal, M. Nasiruddin and M. K. A. Khan, in *Protein Modificomics*, eds. T. A. Dar and L. R. Singh, Academic Press, 2019, DOI: https://doi.org/10.1016/B978-0-12-811913-6.00003-5, pp. 69-86.
13. P. Coopman, *Biomolecules*, 2022, **12**.
14. I. Mees, H. Tran, A. Roberts, L. Lago, S. Li, B. R. Roberts, A. J. Hannan and T. Renoir, *Molecular Neurobiology*, 2022, **59**, 2456-2471.
15. S. T. Rapundalo, *Cardiovasc Res*, 1998, **38**, 559-588.
16. P. Hassan, D. Fergusson, K. M. Grant and J. C. Mottram, *Molecular and Biochemical Parasitology*, 2001, **113**, 189-198.
17. J. V. B. Borba, A. C. Silva, P. I. P. Ramos, N. Grazzia, D. C. Miguel, E. N. Muratov, N. Furnham and C. H. Andrade, *Comput Struct Biotechnol J*, 2019, **17**, 352-361.
18. A. Efstathiou and D. Smirlis, *Microorganisms*, 2021, **9**.
19. N. Baker, C. M. C. Catta-Preta, R. Neish, J. Sadlova, B. Powell, E. V. C. Alves-Ferreira, V. Geoghegan, J. B. T. Carnielli, K. Newling, C. Hughes, B. Vojtkova, J. Anand, A. Mihut, P. B. Walrad, L. G. Wilson, J. W. Pitchford, P. Volf and J. C. Mottram, *Nature Communications*, 2021, **12**, 1244.
20. A. L. Freitas-Mesquita, A. L. A. Dos-Santos and J. R. Meyer-Fernandes, *Front Cell Infect Microbiol*, 2021, **11**, 633146.
21. F. Sacco, L. Perfetto, L. Castagnoli and G. Cesareni, *FEBS Lett*, 2012, **586**, 2732-2739.
22. A. Alonso, J. Sasin, N. Bottini, I. Friedberg, I. Friedberg, A. Osterman, A. Godzik, T. Hunter, J. Dixon and T. Mustelin, *Cell*, 2004, **117**, 699-711.
23. S. h. Hong, S. Y. Xi, A. C. Johns, L. C. Tang, A. Li, M. N. Hum, C. A. Chartier, M. Jovanovic and N. H. Shah, *ChemBioChem*, 2023, **24**, e202200706.
24. S. Liu, B. Zhou, H. Yang, Y. He, Z. X. Jiang, S. Kumar, L. Wu and Z. Y. Zhang, *J Am Chem Soc*, 2008, **130**, 8251-8260.
25. V. V. Kalashnikov and V. V. Samukov, *Chemistry of Natural Compounds*, 1988, **24**, 624-629.
26. Fischer Scientific BOC-Protection/Deprotection, https://www.fishersci.co.uk/gb/en/scientific-products/lab-reporter-europe/chemicals/amine-protection-deprotection.html#:~:text=The%20deprotection%20of%20a%20BOC,are%20the%20acids%20of%20choice., (accessed 05/08, 2023).
27. C. A. A. Claesen, R. P. A. M. Segers and G. I. Tesser, *Recueil des Travaux Chimiques des Pays-Bas*, 1985, **104**, 119-122.
28. J. W. Lee, C.-W. Lee, J. H. Jung and D. Y. Oh, *Synthetic Communications*, 2000, **30**, 2897-2902.
29. C. Xu, J. M. Raible and P. H. Dussault, *Organic Letters*, 2005, **7**, 2509-2511.
30. S. Li, J. Wawrzyniak, Y. Queneau and L. Soulère, *Molecules*, 2017, **22**, 2090.

31. S. Nishimura, *Handbook of heterogeneous catalytic hydrogenation for organic synthesis*, Wiley, New York ; Chichester, 2001.
32. E. Porta, J. Isern, K. Karunakaran and P. Steel, *Frontiers in Pharmacology*, 2022, **13**.