Peer-reviewed version available at *Molecules* **2019**, *24*, 142; doi:10.3390/molecules24010142

Article

Identification of novel inhibitors of pteridine reductase 1 in *Trypanosoma brucei* via computational structure-based approaches and *in vitro* inhibition assays

Magambo Phillip Kimuda[1], Dustin Laming[2,3], Heinrich C. Hoppe[2,3], and Özlem Tastan Bishop[1,*]

[1] Research Unit in Bioinformatics (RUBi), Department of Biochemistry and Microbiology, Rhodes University, P.O. Box 94, Grahamstown, 6140, South Africa; pjkanywa@gmail.com

[2] Department of Biochemistry and Microbiology, Rhodes University, Grahamstown 6140, South Africa; dustinlaming89@gmail.com; H.Hoppe@ru.ac.za

[3] Centre for Chemico- and Biomedicinal Research, Rhodes University, Grahamstown 6140, South Africa; dustinlaming89@gmail.com; H.Hoppe@ru.ac.za

* Correspondence: o.tastanbishop@ru.ac.za; Tel.: +27-084-826-5789 (O.T.B)

Abstract: Pteridine reductase 1 is a trypanosomatid multifunctional enzyme that provides a mechanism for escape of Dihydrofolate reductase (DHFR) inhibition. This is because PTR1 can reduce pterins and folates. Trypanosomes require folates and pterins for survival and are unable to synthesize them *de novo*. Currently there are no anti-folate based Human African Trypanosomiasis (HAT) chemotherapeutics in use. Thus, successful dual inhibition of *Tb*DHFR and *Tb*PTR1 has implications in the exploitation of anti-folates. We carried out molecular docking of a ligand library of 5742 compounds against *Tb*PTR1 and identified 18 compounds showing promising binding modes. The protein-ligand complexes were subjected to Molecular dynamics to characterize their molecular interactions and energetics followed by *in vitro* testing. In this study, we identified five potential *Tb*PTR1 inhibitors that showed low micromolar Trypanosome growth inhibition in *in vitro* experiments with no significant human cell cytotoxicity. Compounds RUBi004, RUBi007, RUBi014, and RUBi018 displayed moderate to strong antagonism when used in combination with the known *Tb*DHFR inhibitor, WR99210. This gave an indication that the compounds might inhibit both *Tb*PTR1 and *Tb*DHFR. RUBi016 showed an additive effect in the isobologram assay. Our results provide a basis for scaffold optimization for further studies in the development of HAT antifolates.

Keywords: Human African Trypanosomiasis; pteridine reductase 1; PTR1; DHFR; antifolates; anti-trypanosomal agents; molecular dynamics; dynamic residue network analysis; binding free energy; isobologram assay.

1. Introduction

African trypanosomes are flagelleted hemo-parasites, transmitted by Tsetse flies, and cause zoonotic infection in mammalian hosts [1]. In animals the disease is known as Nagana while in humans it is known as Human African Trypanosomiasis (HAT) [2, 3]. Acute HAT is caused by *Trypanosoma brucei rhodesiense* (*Tbr*) while chronic HAT by *Trypanosoma brucei gambiense* (*Tbg*). This neglected tropical disease (NTD) remains of considerable public health and animal production concern [4, 5].

Trypanosomes are unable to synthesize folates and pterins *de novo* [6]. Reduced folate and pterin cofactors are essential for parasite survival where they are critical in pathways such as protein and nucleic acid biosynthesis [7]. In order to survive, trypanosomes scavenge extracellular folate and pterin precursors from their hosts [8, 9], hence the pathway is an interesting drug target. Drugs targeting the folate pathway have been used in the treatment of several infections, most notably in the treatment of bacterial and malarial infections [10]. However, their use in the treatment and management of HAT has not been successful to date.

The key enzymes involved in trypanosome folate metabolism are dihydrofolate reductase (DHFR) and pteridine reductase 1 (PTR1) (Figure 1) [11–13]. DHFR (EC 1.5.1.3) is an NADPH dependent enzyme that catalyzes reduction of folate to dihydrofolate (H_2F), and H_2F to tetrahydrofolate (H_4F) (Figure 1) [11, 12]. Folate is essentially a pteridine that has been conjugated to *p*-aminobenzoic acid (pABA) that is glutamylated (Figure 1) [14]. DHFR is a validated and primary target of most antifolate drugs [12]. However the use of traditional antifolates against DHFR in trypanosomatids such as *Trypanosoma* and *Leishmania* has been largely unsuccessful [12, 13, 15, 16].

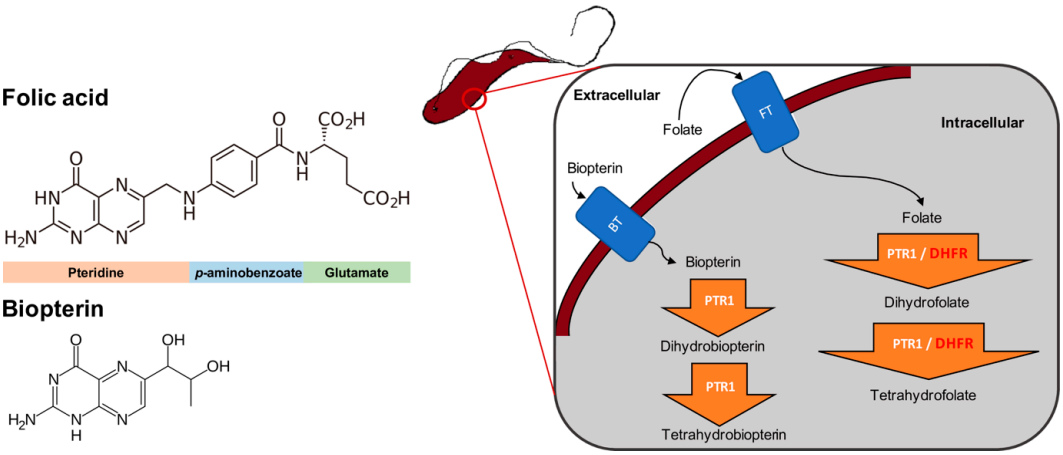


Figure 1. The role of DHFR and PTR1 in trypanosome folate and pterin metabolism. Trypanosomes, which are auxotrophic for folates and pterins, salvage them from the host. The structures of folate and biopterin are shown to the left. Folates and pterins are taken up by transporters (folate-biopterin transporter superfamily, includes biopterin transporter 1 [BT1] and folate transporter 1 [FT1]) after which they are reduced to their functional cofactors (right).

PTR1 (EC 1.5.1.33), which is a short-chain dehydrogenase reductase family member and an NADPH dependent enzyme, is unique to trypanosomatids [8]. It is important in the reduction of biopterin to dihydrobiopterin (H_2B), and H_2B to tetrahydrobiopterin (H_4B) (Figure 1). PTR1 also reduces folate to H_2F, and H_2F to H_4F (Figure 1) [8]. In trypanosomatids, PTR1, which is less susceptible to traditional antifolate inhibition, contributes about 10% to total folate metabolism [13]. It is important to note that studies have shown that under DHFR inhibition PTR1 is over-expressed, thus promoting antifolate resistance in *Leishmania major* and *Trypanosoma cruzi* [8, 13, 15, 16]. This has been proposed as the key mechanism by which trypanosomatids are able to resist antifolates targeting DHFR [8, 13, 15, 16]. Gene knock down and knock out studies in *T. brucei* have shown that

PTR1 is essential for parasite survival. As such, its inhibition alone might be sufficient to negatively impact parasite survival [17, 18].

There are several studies that have reported successful combination of PTR1 and DHFR inhibitors in order to achieve synergistic inhibition of the trypanosomatid folate pathway in *T. cruzi*, *L. major* and *T. brucei* [18–22]. However, identifying a single inhibitor motif that can target both enzymes has remained largely elusive. This has been hampered by poor selectivity against human DHFR as has been the case with PTR1 inhibitors that contain functional groups derived from DHFR inhibitors, such as 2,4 diaminoquinazoline, 2,4 diaminopteridine, or 2,4 diaminopyrimidine moieties [18, 20, 21]. Further, the current drugs used to treat HAT are old, toxic and reducing in efficacy due to resistance [23, 24]. A recent development in African HAT chemotherapy is the promising oral drug fexinidazole that is currently in clinical testing for the treatment of late stage chronic HAT (*T b gambiense*) [25].

In this study, we sought to identify novel *T. brucei* PTR1 (*Tb*PTR1) inhibitors that can be used in conjunction with known DHFR inhibitors or single inhibitors that target both enzymes with minimal human toxicity. Here, we performed structure based virtual screening of 5742 small ligand molecules against *Tb*PTR1 and its orthologues from *T. cruzi* (*Tc*PTR1), *L. major* (*Lm*PTR1) and human (*Hs*DHRS4). *In silico* docking experiments identified 18 compounds preferentially bound to the trypanosomatid PTR1s and not the human DHRS4 orthologue. These promising 18 potential hits complexed with *Tb*PTR1 were then subjected to molecular dynamics (MD) simulations, molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) and free energy calculations as well as dynamic residue network (DRN) analysis. Based on their computational binding modes, selectivity, dynamic stability during MD simulation, DRN analysis, free energy of binding, and commercial availability, 13 compounds were subsequently subjected to a blood stream form (BSF) *T. brucei in vitro* inhibition assay and *H. sapiens in vitro* cytotoxicity assay. Five of the compounds, named RUBi004, RUBi007, RUBi014, RUBi016 and RUBi018, out of 13 exhibited anti-trypanosomal activities against trypanosomes in culture with IC_{50} values of 4.1 μM, 32.6 μM, 16.3 μM, 31.9 μM, and 12.3 μM, respectively, with no significant human cell cytotoxicity at 20 μM.

Compounds RUBi004, RUBi007, RUBi014 and RUBi018 showed reduced inhibition when used in combination with a known DHFR inhibitor (WR99210), which may be suggestive of competitive inhibition of *Tb*DHFR confounded by the upregulation of *Tb*PTR1 expression as a result of the simultaneous inhibition. Compound RUBi016 showed an additive effect when used in combination with WR99210 suggesting that it preferentially inhibits *Tb*PTR1 and the addition of WR99210 further contributes to the reduction of available reduced folates resulting in reduced parasite viability. In summary, the current study reports five compounds with inhibitory activities at low μM levels and their scaffolds may be further optimized to design safe and effective HAT chemotherapeutics targeting the folate pathway.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 November 2018 doi:10.20944/preprints201811.0334.v1

Peer-reviewed version available at Molecules 2019, 24, 142; doi:10.3390/molecules24010142

2. Results

Please note that unless otherwise indicated, *Tb*PTR1 numbering is used throughout the article.

2.1. Overview of PTR1 structure and conservation

PTR1 is a short-chain dehydrogenase/reductase (SDR) family member that has the ability to catalyze the NADPH dependent two-stage reduction of biopterins to their 7,8-dihydro, and 5,6,7,8-tetrahydro forms as well as folates to H_2F and H_4F forms [8] (Figure 1). PTR1 is a tetramer (Figure 2A) with each single α/β- domain subunit constructed around an NADPH binding Rossman-fold repeat that is composed of seven parallel β-sheets that are between three α-helices on either side (Figure 2A) [16].

Multiple sequence alignment (MSA) of the *Tb*PTR1, *Tc*PTR1, *Lm*PTR1 and *Hs*DHRS4 orthologue protein sequences showed conservation of several key residues within the SDR family signature (residues ASP161-ALA193) as well as in the substrate binding loop (residues SER207-GLU215) present in trypanosomatids (Figure2B). Pterin and folate substrates along with inhibitors interact with PTR1 complexes quite similarly, often via binding in a π-sandwich between the NADPH nicotinamide ring and residue PHE97 [19, 26]. The NADPH cofactor is known to be essential in creating both the substrate binding site as well as the catalytic center [19, 27]. ARG14, SER95, PHE97, ASP161, and TYR174 are important residues that interact with the folate and pterin substrates, and are well conserved among the trypanosomatids [19, 26] (Figure 2B). The structural superimposition of *Tb*PTR1 (PDB:2X9N) [26] with *T. cruzi* PTR2 (PDB:1MXH) [28], *Lm*PTR1(PDB:1E92) [16], and *H. sapiens* DHRS4 (PDB: 304R) gave RMS values of 0.4, 0.5, and 1.6, indicating that the trypanosomatid PTR1s are structurally very similar.

Figure 2: A cartoon representation of the TbPTR1 protein structure (PDB: 2X9N) and a multiple sequence alignment of *T. brucei* PTR1, *T. cruzi* PTR1, *T. cruzi* PTR2, *L. major* PTR1 and *H. sapiens* dehydrogenase/reductase (SDR family) member 4 (DHSR4). A) The protein is colored by chain, the NADPH cofactor blue, and cyromazine orange. *Tb*PTR1 is a tetramer and the α/β single domain subunit (chain A) is shown in green. The substrate binding loop is colored red and was composed of

residues SER207-GLU215, while the SDR family signature which is colored brown was composed of residues ASP161-ALA193. B) The MSA showed notable conservation in the SDR family signature as shown by the sequence logo of the extracted motif. The MSA also showed that within the substrate binding loop there was a 4 residue deletion that was present only among the trypanosomatids.

2.2. Eighteen potential hits out of 5742 compounds are identified via virtual screening

*Tb*PTR1 (PDB:2X9N) [26], *Lm*PTR1(PDB:1E92) [16], *Hs*DHRS4 (PDB: 304R) and a homology model of *Tc*PTR1 were used in *in silico* large-scale docking experiments. All structures included the NADPH cofactor that is essential for arrangement of the substrate binding site and catalytic center. 5742 compounds were docked against four proteins using Autodock Vina as described in the Methodology section in order to identify potential hits. Top compounds were selected based on their Autodock Vina energy score of > -8.0 kcal/mol and their hydrogen bonding profiles. Eighteen compounds showed good selectivity for trypanosomatid PTR1 as presented in Table 1. Out of 18 compounds, only RUBi006 bound to the *Hs*DHSR4 active site but with a weaker binding energy than the trypanosomatids. The docked complexes were analyzed using PyMOL [29] and Discovery Studio [30]. The docking energy scores were also evaluated using Xscore [31]. A summary of the compounds with the top binding modes and their corresponding energies are shown in the Table S3. The top binding modes involved ligand interactions with residues that are known to be of catalytic importance *i.e*. ARG14, SER95, PHE97, ASP161 and TYR174 (Table S3) [16, 19]. Residues ARG14, SER95, PHE97, and ASP161 were conserved among all the trypanosomatids, while TYR174 was conserved in all the PTR1 orthologues (Figure 2). Furthermore, residues ASP161 and TYR174 are located within the SDR family signature that is important in the NADPH cofactor and substrate binding (Figure 2).

Table 1. The IUPAC names of the top *Tb*PTR1 docking compounds and Autodock Vina molecular docking results

COMPOUND INFORMATION			DOCKING BINDING ENERGY (kcal/mol)			
Code name	**IUPAC name**	**Database ID**	***T. brucei* PDB:2X9N**	***T. cruzi* Homology model**	***L. major* PDB: 1E92**	***H. sapiens* PDB:3O4R**
RUBi001	2-(2,3-dihydroxyphenyl)-6-hydroxychromen-4-one	ZINC00057846	-10.1	-9.6	-9.2	-
RUBi002	N-(2,6-dimethylphenyl)-2-(1-hydroxy-7-methoxy-9-oxo-xanthen-3-yl)oxy-acetamide	ZINC08992677	-10.2	-10.1	-9.8	-
RUBi003	2-(4-Hydroxyphenyl)ethyl 4-hydroxy-3-methoxybenzoate	SANC00368	-9.1	-8.6	-8.1	-
RUBi004	N'-{[1-(2,4-dimethylphenyl)-3-methyl-5-oxo-1,5-dihydro-4H-pyrazol-4-ylidene]methyl}-3-nitrobenzohydrazide	ZINC00809143	-10.3	-10.1	-9.1	-
RUBi005	2-nitro-N-({2-[(2-methylphenoxy)acetyl]hydrazino}carbothioyl)benzamide	ZINC02690799	-9.0	-8.8	-8.6	-
RUBi006	1,4,6-Trihydroxy-3-methoxy-8-methyl-9H-xanthen-9-one	SANC00470	-10.2	-9.5	-8.6	-7.7[1]
RUBi007	N-(4-methoxyphenyl)-2-[(4-oxo-5-phenyl-4,5-dihydro-1H-pyrazolo[3,4-d]pyrimidin-6-yl)sulfanyl]acetamide	ZINC00630525	-9.6	-8.8	-9.1	-
RUBi008	[3-Methoxy-4-(3-methyl-benzyloxy)-benzyl]-(1H-[1,2,4]triazol-3-yl)-amine	ZINC06556964	-8.5	-8.9	-8.6	-
RUBi009	N-{2-[(2-chlorobenzyl)oxy]ethyl}-2-(1,3-dimethyl-2,6-dioxo-1,2,3,6-tet	ZINC02177983	-8.9	-8.3	-8.2	-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 November 2018 doi:10.20944/preprints201811.0334.v1

Peer-reviewed version available at Molecules 2019, 24, 142; doi:10.3390/molecules24010142

	rahydro-9H-purin-9-yl)acetamide					
RUBi010	2-{[2-oxo-2-(1-pyrrolidinyl)ethyl]sulfanyl}-1,3-benzothiazol-6-ylamin					
	e	ZINC00359797	-6.9	-7.6	-7.8	-
RUBi011	2-[2-(4-amino-1,2,5-oxadiazol-3-yl)-1H-benzimidazol-1-yl]-N-(3-fluor					
	o-4-methylphenyl)acetamide	ZINC00677623	-9.7	-9.8	-9.7	-
RUBi012	1-cyclohexyl-3-(4-sulfamoylphenyl)urea	ZINC01003765	-9.1	-8.1	-7.9	-
RUBi013	3-{[(5-methyl-3-isoxazolyl)amino]sulfonyl}-N-(1,3,4-thiadiazol-2-yl)b					
	enzamide	ZINC02184332	-8.7	-9.4	-7.9	-
RUBi014	2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-3,4-dihydro-2H-1-benzopyra	ZINC0058117 /				
	n-4-one	SANC00320	-9.7	-9.1	-9.7	-
RUBi015	N-(3-hydroxyphenyl)-2-(1-oxido-3-oxo-3,4-dihydro-2H-1,4-benzothia					
	zin-2-yl)-2-oxoacetamide	ZINC04671320	-9.1	-9.1	-8.3	-
RUBi016						
	2-(1,3-dimethyl-2,6-dioxo-1,2,3,6-tetrahydro-9H-purin-9-yl)-N-(4-hyd					
	roxyphenyl)acetamide	ZINC00612219	-8.9	-8.9	-7.7	-
RUBi017	2-nitro-N-{2-[(2-nitrophenyl)formamido]propyl}benzamide	ZINC04523829	-8.8	-9.1	-8.4	-
RUBi018						
	2-[2-(cyclopentylidenehydrazono)-4-hydroxy-2,5-dihydro-1,3-thiazol					
	-5-yl]-N-phenylacetamide	ZINC04313814	-8.4	-8.4	-8.8	-

[1] RUBi006 was the only compound that bound to the human orthologue.

2.3. Drug likeness

A PCA of the compounds based on their molecular descriptors as listed in the Methodology section showed that the top docking compounds clustered well with known FDA approved CNS permeable drugs (Figure 3) occupying the same chemical space as FDA approved drugs as well as FDA approved CNS permeable drugs. The compounds are of 'drug-like' desirability and are likely to cross the BBB making them good candidates for HAT chemotherapeutics [32].

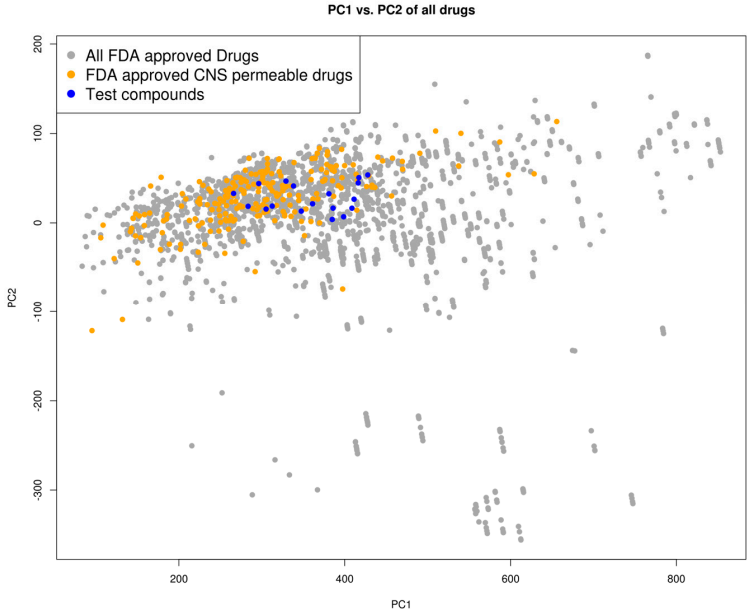


Figure 3. Chemical space distribution of the test compounds (n = 18), FDA approved drugs (n= 3180), and FDA approved CNS permeable drugs (n= 221). The PCA analysis of the molecular descriptors

showed that the test compounds clustered well with FDA approved CNS permeable drugs. The PCA scatter plot was based on the first (PC1) and second (PC2) components as shown on the two axes. PC1 explained 76% of the variance while PC2 explained 18%. All FDA approved drugs are shown in grey dots, while FDA approved CNS permeable drugs are shown as yellow dots and the test compounds are shown as blue dots.

2.4. Five hit compounds show anti-trypanosomal activity in vitro

As a next step, a total of 18 *Tb*PTR1-ligand complexes were subjected to 200 ns all atom MD simulations followed by MM-PBSA free energy calculations. All the compounds showed linear stable MD trajectories as observed in RMSD, Rg calculations (Figures S2 and S3), and promising hydrogen bonding features (Figure S4). RMSF values (Figure S5) and binding free energies (Figure S6) were also calculated. PCA was also carried out to investigate the overall dynamics of the protein systems [33]. Overall, *in silico* analysis indicated that all 18 compounds could be further studied for *in vitro* analysis. However, only 13 of these were commercially available and tested for anti-trypanosomal activity against *T. brucei* BSF in culture. Compounds RUBi003, RUBi006, RUBi009, RUBi013, and RUBi017 could not be purchased, and as such were not used in the *in vitro* inhibition assays, even though they showed similar binding modes to folate, pterins and known *Tb*PTR1 inhibitors. In spite of good binding modes and stable MD trajectories, compounds RUBi001, RUBi002, RUBi005, RUBi008, RUBi010, RUBi011, and RUBi012 did not show anti-trypanosomal activity in *in vitro* experiments when used at 20 µM. The five compounds shown in Figure 4 that showed significant inhibition of *T. brucei* viability at this concentration were subjected to dose-response assays to derive their IC_{50} values against *T. brucei* (Figure 5) and are discussed further below.

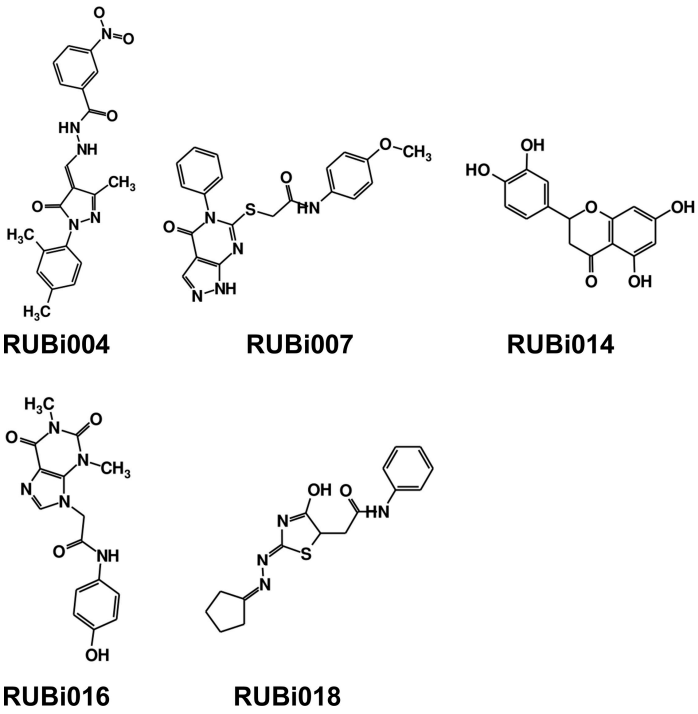


Figure 4. Structures of the five *Tb*PTR1 hit compounds

2.4.1. *Tb*PTR1 hit compounds have either antagonist or additive activity when used in combination with a *Tb*DHFR inhibitor

RUBi016 inhibited blood stream form (BSF) trypanosome growth in culture with IC_{50} value of 31.9 µM (Figure 5), and it was the only compound which showed an additive effect when used in combination with WRR99210 (Figure 6A). RUBi004, RUBi007, RUBi014, and RUBi018 inhibited BSF

trypanosome growth in culture with IC_{50} values of 4.1 μM, 32.6 μM, 16.3 μM, and 12.3 μM respectively (Figure 5). When used in combination with the known DHFR inhibitor WR99210, these compounds showed antagonism - RUBi018 to the least and RUBi014 to the highest extent (Figure 6).

As a next step, all the hit compounds were also docked against *Tb*DHFR active site to see potential binding activity. RUBi004 (Figure 8A), RUBi007 (Figure 8B), RUBi014 (Figure 8C), and RUBi018 (Figure 8E) molecular *Tb*DHFR dockings showed high binding affinity towards the protein with -9.4 kcal/mol, -9.5 kcal/mol, -8.7 kcal/mol, and -7.9 kcal/mol respectively. RUBi016 bound to *Tb*DHFR with a binding affinity of -7.6 kcal/mol (Figure 8D). Interestingly, *Tb*DHFR-compound binding energy scores were in agreement with isobologram results. Three compounds (RUBi004, RUBi007 and RUBi014) with the greatest antagonist activities had also the highest binding energies towards *Tb*DHFR. The possibility that RUBi004, RUBi007 and RUBi014 also inhibit DHFR may explain why they show more marked antagonism when used in combination with DHFR inhibitor: WR99210 and each of these three compounds both inhibit *Tb*DHFR, thus possibly enhancing the upregulation of *Tb*PTR1.

Compound RUBi018, with the least antagonist activity, had a similar *Tb*DHFR binding energy to RUBi016 that showed an additive effect with WRR99210. RUBi018 also displayed the highest binding energy for *Tb*PTR1 among all five compounds (-8.4 kcal/mol; Table 1). Interestingly, with the exception of RUBi018 which was more active than predicted, the anti-trypanosomal IC_{50} values of the remaining 4 compounds correlated with their *Tb*PTR1 binding energies – RUBi004 with the highest affinity (-10.4 kcal/mol) displayed the most potent activity (IC_{50} 4.1 μM).

The use of *Tb*PTR1 inhibitors in combination with *Tb*DHFR inhibitors has long been proposed as a viable avenue for the generation of a new anti-trypanosomal anti-folate drugs [19]. It is important to note, however, that studies have shown that under DHFR inhibition PTR1 is over-expressed, thus promoting antifolate resistance in *Leishmania major* and *Trypanosoma cruzi* [8, 13, 15, 16]. This likely contributes to the antagonism of the RUBi compounds' activity by WR99210 – an effect that might be exacerbated by the additional inhibition of *Tb*DHFR by RUBi004, RUBi007, RUBi014, and RUBi018.

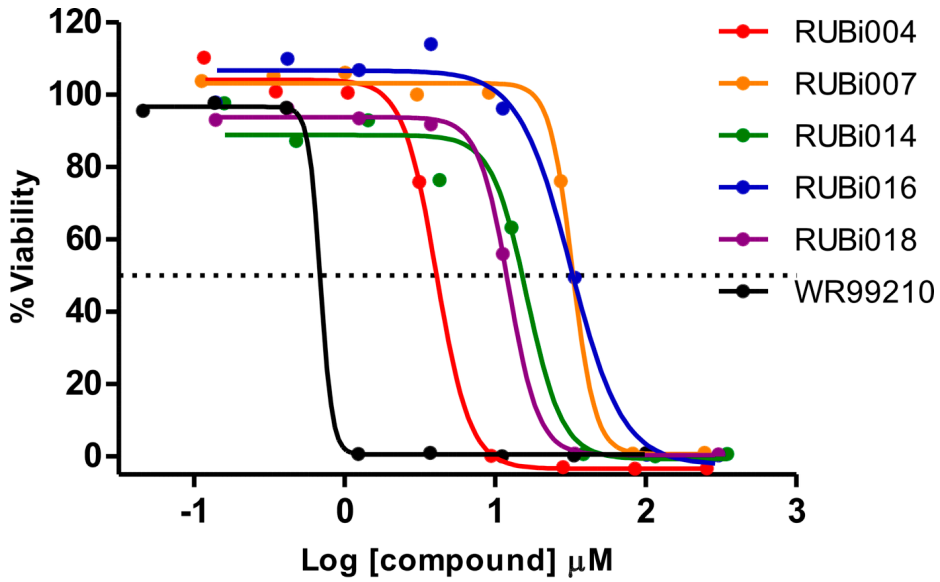


Figure 5. Inhibition constant determination of the RUBi004 (red), RUBi014 (orange), RUBi007 (green), RUBi016 (blue), RUBi018 (purple), WR99210 (maroon), and pentamidine (black) which was used as the positive control. The IC_{50} was determined using the resazurin method. The IC_{50} of RUBi004, RUBi007, RUBi014, RUBi016, and RUBi018 were determined to be 4.1 μM, 32.6 μM, 16.3 μM, 31.9

μM, and 12.3 μM respectively. The IC_{50} of WR99210 and pentamidine (not shown) were determined to be 0.69 μM and 0.005 μM respectively.

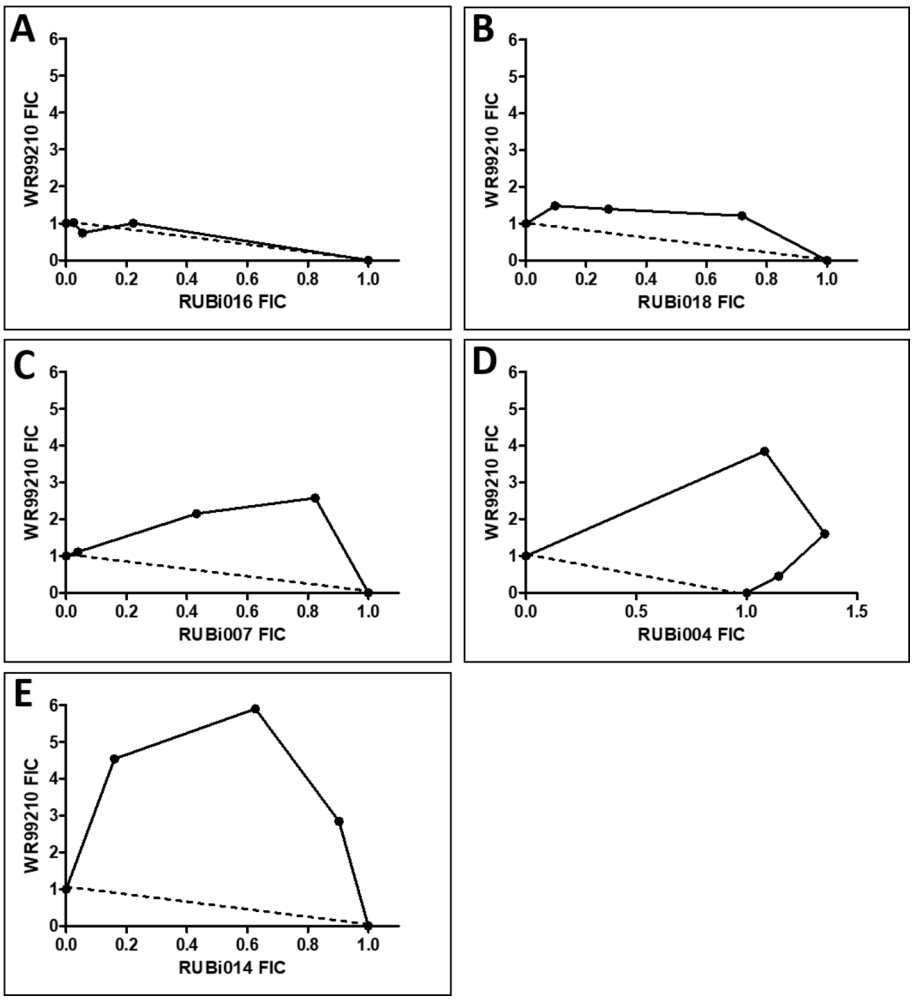


Figure 6. Isobologram analysis of compounds RUBi016 (A), RUBi018 (B), RUBi007 (C), RUBi004 (D) and RUBi014 (E) in combination with WR99210. RUBi compounds and WR99210 were employed alone at starting concentrations of 100 μM and 20 μM respectively, and in combination ratios of 75:25, 50:50 and 25:75. IC_{50} values obtained for the RUBi compounds and WR99210 alone and in combination were used to calculate and plot their fractional inhibitory concentrations (FIC). The dotted line in the graphs denotes an additive effect; FIC values above the line indicate compound antagonism, while values below the line indicate synergism.

*Tb*PTR1 and *Tb*DHFR both catalyze the reduction of folate to H_2F and H_4F however the substrate binding sites are differently ordered [34]. Additionally, *Tb*DHFR is capable of undergoing significant conformational changes when in complex with thymidine synthase (TS) while *Tb*PTR1 is more rigid [35–37]. We observed that all the ligands in the *Tb*DHFR-complexes formed a π-π interactions with residue PHE58 that is involved in both WR99210 and pyrimethamine inhibitor binding (Figure 8) [38]. In *Tb*PTR1 RUBi004, RUBi007, RUBi014 and RUBi018 formed similar π-π interactions with residues PHE97, TYR174, TRP221, and PHE171 (Figure 7). RUBi016 appears to be the only ligand that formed a hydrogen bond with the NADPH cofactor (Figure 7D). All the compounds appear to bind in similar patterns to pterin, folate, and known *Tb*PTR1 inhibitors interacting with the NADPH cofactor and the *Tb*PTR1 protein [19]. The detailed analysis for each compound is presented below.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 November 2018 doi:10.20944/preprints201811.0334.v1

Peer-reviewed version available at Molecules 2019, 24, 142; doi:10.3390/molecules24010142

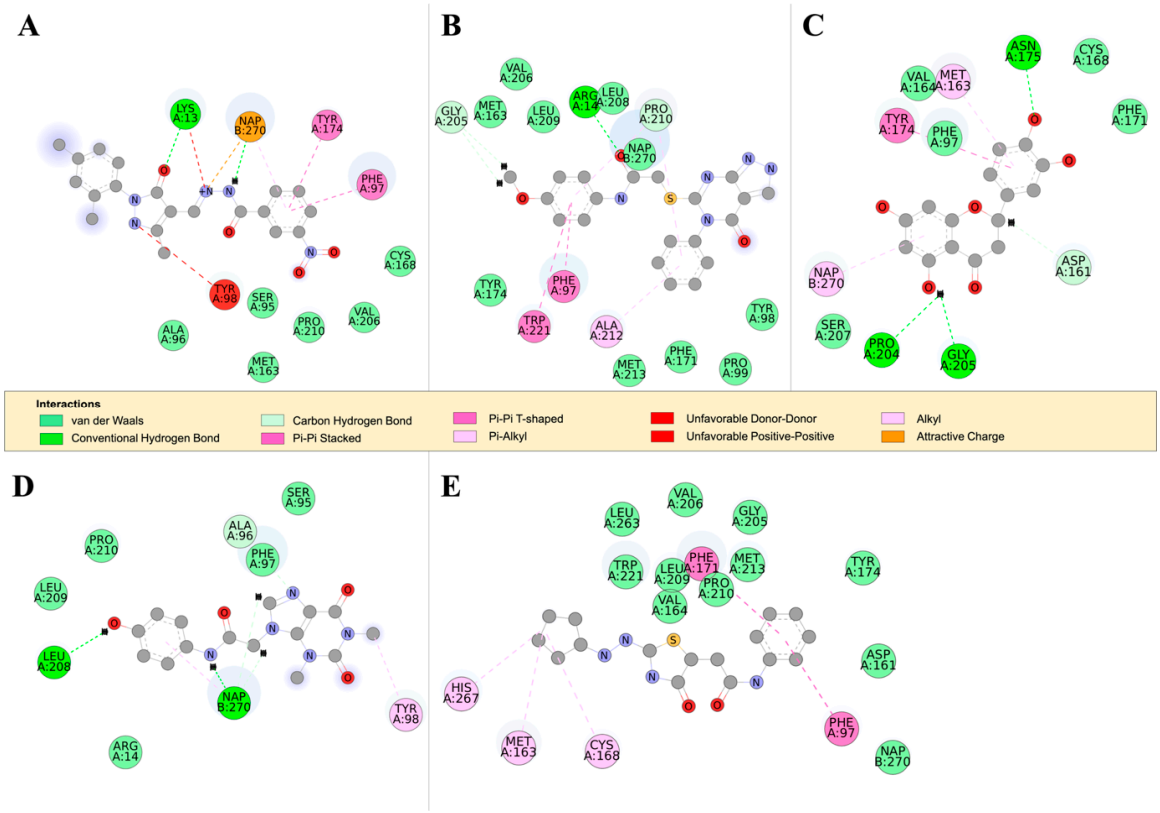


Figure 7. The 2D structure of RUBi004, RUBi007, RUBi014, RUBi016, and RUBi018 along with their binding modes in the *Tb*PTR1 protein. A) A 2D representation of RUBi004 and its binding mode with *Tb*PTR1; B) A 2D representation of RUBi007 and its binding mode with *Tb*PTR1; C) A 2D representation of RUBi014 and its binding mode with *Tb*PTR1; D) A 2D representation of RUBi016 and its binding mode with *Tb*PTR1; D) A 2D representation of RUBi018 and its binding mode with *Tb*PTR1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 November 2018 doi:10.20944/preprints201811.0334.v1

Peer-reviewed version available at Molecules 2019, 24, 142; doi:10.3390/molecules24010142

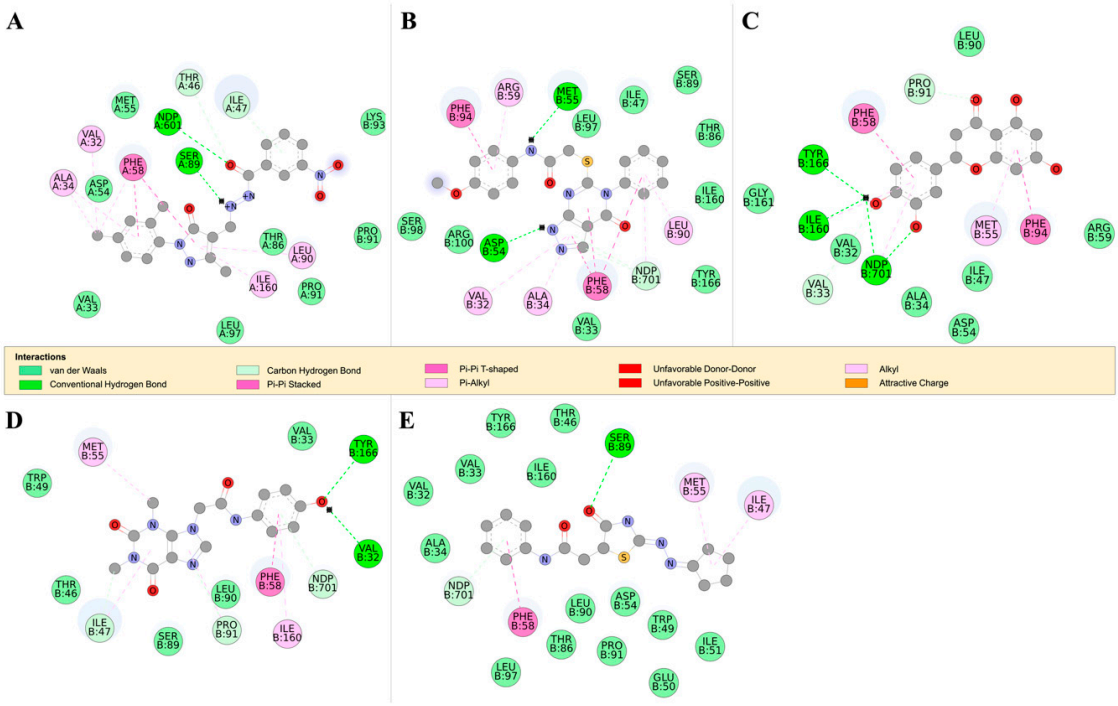


Figure 8. The 2D structure of RUBi004, RUBi007, RUBi014, RUBi016, and RUBi018 along with their binding modes in the *Tb*DHFR protein. A) A 2D representation of RUBi004 and its binding mode with *Tb*DHFR; B) A 2D representation of RUBi007 and its binding mode with *Tb*DHFR; C) A 2D representation of RUBi014 and its binding mode with *Tb*DHFR; D) A 2D representation of RUBi016 and its binding mode with *Tb*DHFR; D) A 2D representation of RUBi018 and its binding mode with *Tb*DHFR.

2.4.2. Compound RUBi004 analysis

Ligand docking: Docking analysis of the *Tb*PTR1-RUBi004 complex showed that the RUBi004 formed a π-sandwich between the NADPH nicotinamide ring and PHE97 (Figure 7A). It also formed a T-shaped π-π interaction with TYR174 and two hydrogen bonds with LYS13 (RUBi004 O1 acceptor; LYS13 HZ1 donor) and NADPH (RUBi004 H8 donor; NADPH O2N and O2A acceptors) (Figure 7A). RUBi004 formed vdW interactions with SER95, ALA96, CYS168, VAL206, and PRO210 (Figure 7A). Unfavorable interactions (colored in red) included positive-positive interaction between LYS13 NZ and RUBi004 N2, along with a donor-donor interaction between TYR98 HH and RUBi004 N1 (Figure 7A).

Furthermore, when RUBi004 was docked to *Tb*DHFR it formed hydrogen bonds with the NADPH cofactor and *Tb*DHFR residues SER89, THR46, and ILE47 (Figure 8A). RUBi004 also formed π-π interactions with *Tb*DHFR residues VAL32, ALA34, THR46, PHE58, LEU90, PRO91, and ILE169 (Figure 8A). RUBi004 bound to the *Tb*DHFR active site with a docking energy of -9.4 kcal/mol. *Tb*DHFR residues ALA34, VAL32, MET55, PHE58, SER89, PHE94, TYR166, and the NADPH cofactor are known to be involved in the binding of *Tb*DHFR inhibitors such as pyrimethamine and WR99210 [38]. This docking analysis shows that RUBi004 binds to *Tb*DHFR in a similar binding mode to known *Tb*DHFR inhibitors [38].

Molecular dynamics: The RMSD calculations revealed that the protein backbone and the NADPH cofactor in the *Tb*PTR1-RUBi004 complex to be stable with slight conformational changes observed in the ligand (Figure S2E). The binding of the ligand led the protein to become more compact as observed in the Rg analysis (Figure S3D). We observed slight increases in the flexibility of residues LYS13, PHE97, TYR98 and TYR174 via RMSF calculations (Figure S5D). Loop residues MET169 and ALA170 showed increased flexibility while helix residues ALA188 and ALA189 showed reduced flexibility (Figure S5D). The substrate binding loop SER207 – GLU217 was stable

(Figure S5D). From the PCA we observed differences in motion between the apo *Tb*PTR1 (Figure 9) and the ligand bound proteins (Figure 10-13). In both systems the largest motions were in the substrate binding loop (residues SER207-GLU215), the α6 helix (residues GLY214-VAL225), CYS160-TY174 loop region, C-terminal residues HIS267-ALA268, the modelled loop 1 (residues GLN104-GLY113), and modelled loop 2 (residues LYS143-SER151) (Figure 9 and Figure 10). The modelled missing residues in modelled loop 1 and modelled loop 2 showed a lot of variability in motion (Figure 9 and Figure 10). From the per residue energy decomposition shown in Figure 15A we can see that the RUBi004 compound has favorable binding energy contributions from substrate binding loop residue VAL206, α6 helix residue LYS218, and forms an unfavorable interaction with ASP161 resulting in their altered motions (Figure 10). On average the ligand formed two hydrogen bonds during simulation with *Tb*PTR1 residue LYS13 and the NADPH cofactor (Figure S4D and Figure 7A).

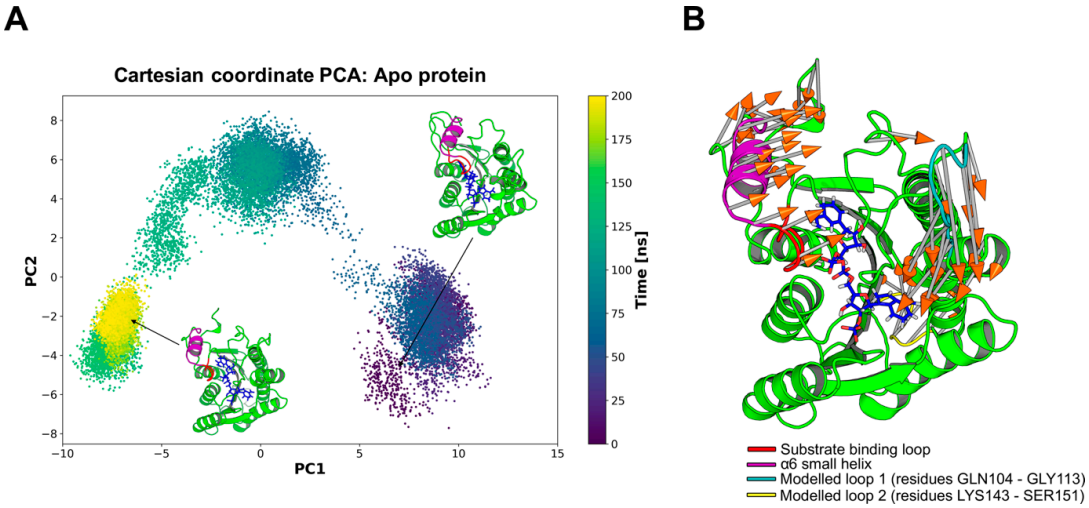


Figure 9. Principal component analysis of the *Tb*PTR1 apo protein. The motions of the protein during 200ns of all atom MD simulation are shown along the first and second principal components (PC1 and PC2). The substrate binding loop (residues SER207-GLU215) is colored red, α6 alpha helix (residues GLY214 - VAL225) is colored magenta, modelled missing residues loop 1 colored cyan, modelled missing loop residues loop2 colored yellow, and NADPH cofactor colored blue. PC1 explained 54% of the variance while PC2 explained 18%. A) Projection of the protein-ligand complex dynamics along the PC1 and PC2. B) The differential motions described by PC1 and PC2 are shown by light gray arrows with orange tips.

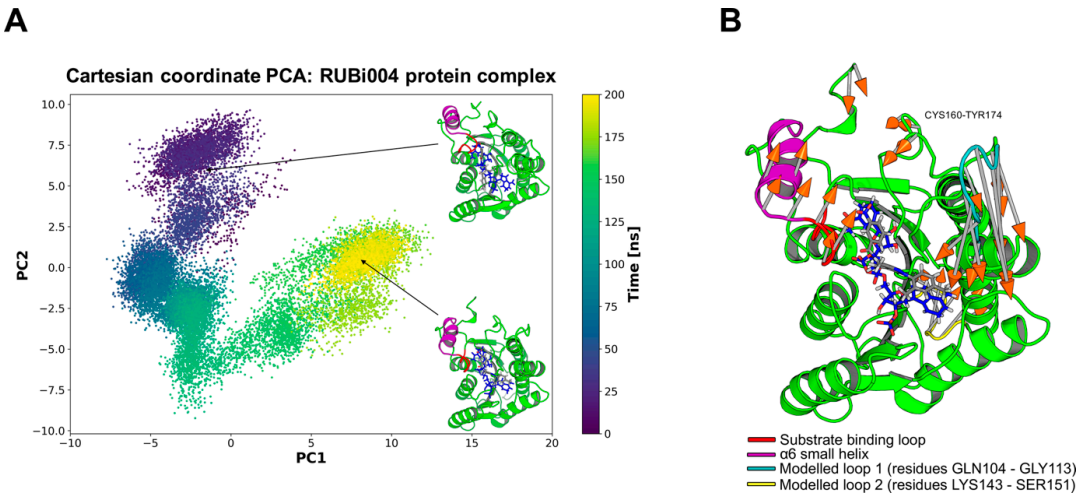


Figure 10. Principal component analysis of the RUBi004 - protein complex. The motions of the complex during 200ns of all atom MD simulation are shown along the first and second principal components (PC1 and PC2). The substrate binding loop (residues SER207-GLU215) is colored red, α6 alpha helix (residues GLY214 - VAL225) is colored magenta, modelled missing residues loop 1 colored cyan, modelled missing loop residues loop2 colored yellow, NADPH cofactor colored blue, and RUBi004 colored gray. PC1 explained 46% of the variance while PC2 explained 26%. A) Projection of the protein-ligand complex dynamics along the PC1 and PC2. B) The differential motions described by PC1 and PC2 are shown by light gray arrows with orange tips.

Binding free energy: RUBi004 bound stably to the protein throughout the MD simulation with a free binding energy of -63.127 ± 14.401 kJ/mol (Table 2). A per residue energy decomposition showed that residues ARG14, PHE97, VAL206, and LYS218 contributed -8.172 kJ/mol, -9.178 kJ/mol, -3.010 kJ/mol, and -4.170 kJ/mol respectively to binding (Figure 15A and Figure 15B-i). Unfavorable energy contributions included ASP161 and GLU231 that contributed 3.790 kJ/mol and 3.110 kJ/mol respectively (Figure 15A and Figure 15B-i). VAL206 was covalently bonded to substrate binding loop residue SER207 that also showed increased centrality and importance in communication (Figure S7A).

Dynamic residue network: *Average L* and *average BC* metrics over MD trajectory [39] were calculated for a comparative DRN analysis between ligand-bound and unbound *Tb*PTR1. These metrics were also compared to RMSF data (Figure S7A). As shown previously by Penkler *et al.*, a general trend between *Average BC*, *Average* L^{-1} and $RMSF^{-1}$ was observed [40]. Pearson correlation coefficient values are presented in the Table S4.

Table 2. A decomposition of the binding energy components obtained from MM-PBSA

Ligand	Van der Waal energy (kJ/mol)	Electrostatic energy (kJ/mol)	Polar solvation energy (kJ/mol)	SASA energy (kJ/mol)	Binding energy (kJ/mol)
RUBi001	-147.290 ± 9.684	-72.775 ± 7.789	113.647 ± 6.897	-13.131 ± 0.561	-119.549 ± 9.302
RUBi002	-158.331 ± 10.799	-54.957 ± 17.424	167.445 ± 25.372	-20.181 ± 0.987	-66.024 ± 17.960
RUBi003	-120.331 ± 9.856	-45.157 ± 16.390	88.501 ± 14.847	-12.178 ± 0.714	-89.166 ± 10.016
RUBi004	-89.386 ± 8.660	-2.494 ± 19.739	40.462 ± 29.099	-11.708 ± 1.480	-63.127 ± 14.401
RUBi005	-176.822 ± 12.683	-21.677 ± 11.102	105.804 ± 13.441	-18.946 ± 0.896	-111.641 ± 15.287
RUBi006	-129.667 ± 9.297	-2.552 ± 7.030	72.606 ± 6.948	-14.686 ± 0.722	-74.298 ± 9.423
RUBi007	-124.684 ± 8.329	-45.487 ± 7.400	97.333 ± 9.983	-15.092 ± 0.730	-87.931 ± 9.776
RUBi008	-133.753 ± 8.370	-46.219 ± 11.226	110.761 ± 15.299	-16.040 ± 0.753	-85.250 ± 10.888
RUBi009	-143.867 ± 18.665	1.809 ± 20.494	94.287 ± 13.356	-18.040 ± 1.565	-65.812 ± 21.921
RUBi010	-144.193 ± 11.717	-138.108 ± 32.060	230.551 ± 48.093	-16.209 ± 0.689	-67.960 ± 23.305
RUBi011	-158.890 ± 9.774	-22.047 ± 12.754	106.149 ± 15.520	-16.788 ± 0.854	-91.577 ± 11.944
RUBi012	-115.925 ± 12.327	41.151 ± 9.268	103.048 ± 20.694	-13.047 ± 1.114	15.227 ± 13.133
RUBi013	-139.991 ± 10.650	-48.508 ± 27.677	142.199 ± 30.711	-15.508 ± 1.132	-61.808 ± 12.718
RUBi014	-98.534 ± 17.175	-17.122 ± 19.998	70.730 ± 28.464	-11.472 ± 1.506	-56.399 ± 11.696
RUBi015	-106.414 ± 10.967	-28.566 ± 11.142	76.073 ± 17.307	-13.479 ± 1.334	-72.386 ± 12.019
RUBi016	-87.250 ± 8.743	21.304 ± 8.487	53.635 ± 10.796	-11.042 ± 0.799	-23.353 ± 10.361
RUBi017	-129.769 ± 12.116	-32.613 ± 13.539	91.673 ± 19.346	-15.262 ± 1.235	-85.971 ± 11.865
RUBi018	-132.408 ± 11.987	-14.638 ± 9.183	74.896 ± 10.635	-15.764 ± 1.018	-87.914 ± 12.026

Overall, only very slight changes, especially in the substrate binding loop (SER207-GLU215) and the small α6 helix (GLY214-VAL225), were observed for *Average L* between apo protein and *Tb*PTR1-RUBi004 complex (Figure S7A). *Average BC* is an important metric to identify residues critical for communication flow within the protein network [39]. Residues THR9, SER95, and ALA238 showed the highest *Average BC* in the *Tb*PTR1 protein in both apo and *Tb*PTR1-RUBi004 complex (Figure S7A). Residue THR9 is conserved among all the PTR1 orthologues while SER95 is conserved among the trypanosomatid PTR1 orthologues only (Figure 2). ALA238 is conserved in all the PTR1 orthologues except *Lm*PTR1 where it is replaced by a serine residue (Figure 2). In the *Tb*PTR1-RUBi004 complex VAL164, SER172, and SER207 showed increases in *Average BC* compared to the apo protein (Figure S7A). SER207 is also conserved among all the PTR1 orthologues (Figure 2).

BC related residue interaction analysis: A close examination of the residue interaction networks of important residues identified via *Average BC* calculations showed that THR9 formed vdW interactions with ALA11, ASN92, ALA94, and hydrogen bonds with VAL91 and ASN93 (Figure S8-1-A). ALA11 in turn formed vdW interactions with LYS13 that is covalently bonded to ARG14 which contributed -8.172 kJ/mol to ligand binding during simulation (Figure 15A). While SER95 formed hydrogen bonds with ALA96, ASN127, and the NADPH cofactor (Figure S8-1-B). Furthermore, SER95 formed vdW interactions with TYR174 and PHE97 both of which are involved in ligand binding (Figure S8-1-B). Lastly ALA238 formed an alkyl interaction with ALA18 that is hydrogen bonded to ILE15 that is covalently bonded to ARG14 (Figure S8-1-C). These interactions with functionally important residues indicate that average BC is helpful in identifying residues crucial to communication flow within the protein-ligand dynamic network.

The binding of the ligand resulted in slightly altered information flow across the network especially at residues VAL164, SER172, and SER207 all of which were interacting with residues

involved in ligand and NADPH cofactor binding (Figure S7A; Figure S8-2-A). VAL164 formed vdW interactions with ASN175 that was covalently bonded to TYR174 and had a hydrogen bond with ASP161. SER172 had an unfavorable donor-donor interaction with TYR174 (Figure S8-2-B). Lastly, residue SER207 formed a hydrogen bond with the NADPH cofactor and residue ARG14 that is covalently bonded to LYS13 (Figure S8-2-C). All these interactions were of note because docking analysis showed that residues LYS13, PHE97, TYR98 and TYR174 were involved in RUBi004 binding (Figure 7A).

2.4.3. Compound RUBi007 analysis

Ligand docking: Docking analysis of the *Tb*PTR1-RUBi007 complex showed that the compound formed T-shaped π-π interactions with PHE97 and TRP221. It also formed π-alkyl interactions with PRO210 and ALA212. It formed hydrogen bonds with ARG14 (RUBi007 O1 acceptor; ARG14 HH12 and HH22 donors), and GLY205 (RUBi007 H3 and H15 donors; GLY205 O acceptor) (Figure 7B). RUBi007 formed vdW interactions with TYR98, PRO99, MET163, PHE171, TY174, GLY205, VAL206, LEU208, LEU209, PRO210, MET213, and the NADPH cofactor (Figure 7B). In the *Tb*DHFR-RUBi007 complex, the compound formed hydrogen bonds with the NADPH cofactor, THR46, ILE47, and SER89 along with π-π interactions with VAL32, ALA34, PHE58, LEU90, and ILE160 (Figure 8B). Many of these residues are known to be involved in the binding of *Tb*DHFR inhibitors such as pyrimethamine and WR99210 [38].

Molecular dynamics: The RMSD calculations revealed that the *Tb*PTR1-RUBi007 complex was very stable (Figure S2H). The complex showed a linear trajectory with the protein backbone, NADPH cofactor, and the ligand showing only slight conformational changes (Figure S2H). The binding of the ligand led the protein to become more compact as observed in the Rg analysis (Figure S3G). We also observed a significant decrease in the flexibility of the substrate binding loop (residues GLY205 – TRP221) due to the ligand binding (Figure S5G). Furthermore, decreased flexibility in ARG14, PHE97, GLY205, PRO210, ALA212, and TRP221 (Figure S5G), and increased flexibility in GLY247, SER248, and ALA249 was observed (Figure S5G). As seen in the decreased flexibility of the substrate binding loop in the RMSF analysis, there were no large motions observed in the substrate binding loop in the PCA in comparison to the apo protein (Figure 9 and Figure 11). The largest motions were in the α6 helix, C-terminal residues HIS267-ALA268, the modelled loop 1, and modelled loop 2 (Figure 11). In comparison to the apo protein (Figure 9), the modelled loop 1 showed less motion than the ligand-protein complex (Figure 11) a similar trend was seen in the RMSF analysis (Figure S5G). From the per residue energy decomposition shown in Figure 15B we can see that the RUBi007 compound has favorable energy binding contributions from substrate binding loop residues LEU209 and PRO210, α6 helix residue LYS224, and forms an unfavorable interaction with ASP161. RUBi007 formed on average 2 hydrogen bonds with the protein residues ARG14 and GLY205 during the MD simulation (Figure S4G and Figure 7B).

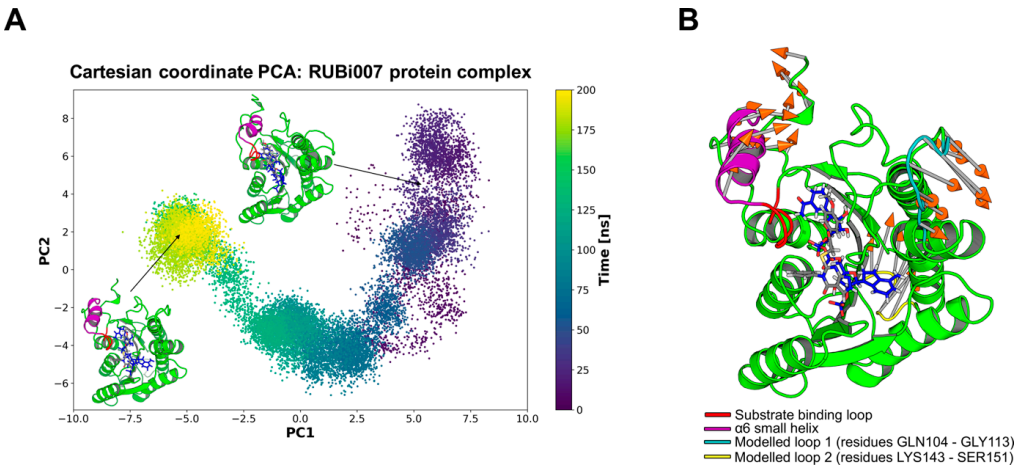


Figure 11. Principal component analysis of the RUBi007 - protein complex. The motions of the complex during 200ns of all atom MD simulation shown along the first and second principal components (PC1 and PC2). The substrate binding loop (residues SER207-GLU215) is colored red, α6 alpha helix (residues GLY214 - VAL225) is colored magenta, modelled missing residues loop 1 colored cyan, modelled missing loop residues loop2 colored yellow, NADPH cofactor colored blue, and RUBi007 colored gray. PC1 explained 40% of the variance while PC2 explained 21%. A) Projection of the protein-ligand complex dynamics along the PC1 and PC2. B) The differential motions described by PC1 and PC2 are shown by light gray arrows with orange tips.

Binding free energy: RUBi007 bound to the protein stably throughout the MD simulation with a free binding energy of -87.931 ± 9.776 kJ/mol (Table 2). A per residue decomposition showed that residues LYS13, ARG14, PHE97, LYS178, LEU209, PRO210, and LYS224 contributed -3.200 kJ/mol, -6.260 kJ/mol, -7.750 kJ/mol, -4.4 kJ/mol, -6.55 kJ/mol, -11.08 kJ/mol, and -3.30 kJ/mol (Figure 15A and Figure 15B-ii). ASP161 gave an unfavorable energy contribution of 14.32 kJ/mol (Figure 15A and Figure 15B-ii). ILE15 which was covalently bonded to ARG14 showed decreased flexibility in RMSF (Figure S5G) and increased centrality (Figure S5G). Similarly, PHE97, LEU209 and PRO210 showed decreased flexibility in RMSF (Figure S5G) and increased centrality (Figure S7B).

Dynamic residue network: When the apo protein was compared to the *Tb*PTR1-RUBi007 complex, *Average L* indicated spatial displacement of the substrate binding loop and the small α6 helix (Figure S7B). A general trend between three metrics (*Average BC*, *Average* L^{-1} and $RMSF^{-1}$) was again in agreement (Figure S7B; Table S4). Like in the *Tb*PTR1-RUBi004, the substrate binding loop and small α6 helix in the *Tb*PTR1-RUBi007 complex showed a lot of *Average L* fluctuation. In comparison to the apo protein, residues VAL211-GLY214 showed increases in *Average L* (Figure S7B). The observed increases in *Average L* correlated well with residue fluctuations (Figure S7B and Table S4) [41].

In the *Tb*PTR1-RUBi007 complex, residues ILE15, MET163, SER207, LEU208, and PRO210 showed increases in average BC (Figure S7B). Multiple sequence analysis showed that residue ILE15 was conserved in all the PTR1 orthologues except *Lm*PTR1 where it is substituted for a leucine (Figure 2). Residue MET163 was conserved only among the trypanosomatids, while LEU208 was conserved only among the *trypanosoma* (Figure 2).

BC related residue interaction analysis: As seen in the other *Tb*PTR1-ligand complexes, these residues with higher average BC values were involved in ligand binding. Residue ILE15 was covalently bound to ARG14, while MET163 formed vdW interactions with the RUBi007 ligand (Figure S8-3-A and Figure S8-3-B). Residue SER207 and LEU208 each had a hydrogen bond with the NADPH cofactor (Figure S8-3-C and Figure S8-3-D). Residue LEU208 also had vdW interactions with the RUBi007 ligand and had a hydrogen bond with both ARG14 (Figure S8-3-D). Residue PRO210 had a π alkyl interaction with PHE97, vdW interactions with ARG14, and had a hydrogen bond with the RUBi007 ligand (Figure S8-3-E).

2.4.4. Compound RUBi014 analysis

Ligand docking: RUBi014, also known as eriodictyol, has previously been reported to be selectively anti-protozoal with activity against *T. brucei* in culture [42, 43]. In the *Tb*PTR1 protein, RUBi014 formed a T-shaped π-π interaction with TYR174, π-alkyl interactions with MET163 and a π-alkyl interaction between with the NADPH nicotinamide ring (Figure 7C). It also formed hydrogen bonds with ASP161 (RUBi014 H donor; ASP161 OD2 acceptor), ASN175 (RUBi014 O3 acceptor; ASN175 HD21 donor), PRO204 (RUBi014 H10 donor; PRO204 O acceptor), and GLY205 (RUBi014 H10 donor; GLY205 O acceptor) (Figure 7C). RUBi014 formed vdW interactions with ASP161, VAL164, CYS168, PHE171, and SER207 (Figure 7C).

When RUBi014 was docked to *Tb*DHFR it formed hydrogen bonds with the NADPH cofactor and VAL33, PRO91, ILE160, and TYR166 (Figure 8C). It also formed π-π interactions with PHE58, MET55, and PHE94 (Figure 8C). RUBi014 bound *Tb*DHFR active site with a docking energy of -8.7 kcal/mol. As was seen in the *Tb*DHFR-RUBi004 complex, the residues that RUBi014 interacted with in *Tb*DHFR were similar to those of known *Tb*DHFR inhibitors such as pyrimethamine and WR99210 [38]

Molecular dynamics: The RMSF calculations revealed that the binding of the caused significant changes in the protein structure (Figure S2P). The protein backbone showed notable conformational change, while the NADPH cofactor and ligand showed only slight changes (Figure S2P). We observed a significant increase in the flexibility of the substrate binding loop and the small α6 helix (Figure S5N). We also observed increases in the flexibility of residues ASP161, MET163, TYR174, ASN175, and PRO204-ARG229 (Figure S5N). The *Tb*PTR1-RUBi014 complex showed very strong alteration of the motion of the substrate binding loop, α6 helix, and C-terminal residues HIS267-ALA268 as shown by the RMSF analysis and PCA analysis (Figure S5N and Figure 12). It appears to make the active site widen and extend (Figure 12) when compared to the apo protein (Figure 9). It appears RUBi014 interacts primarily with the CYS60-TYR174 loop as shown by the per residue energy decomposition shown in Figure 15B-iii. RUBi014 forms favorable energy contributions with MET163, PHE171, and TYR174 (Figure 15C). RUBi014 formed an average of 4 hydrogen bonds with the protein residues ASP161, ASN175, PRO204 and GLY205 during the MD simulation (Figure S4N and Figure 7C).

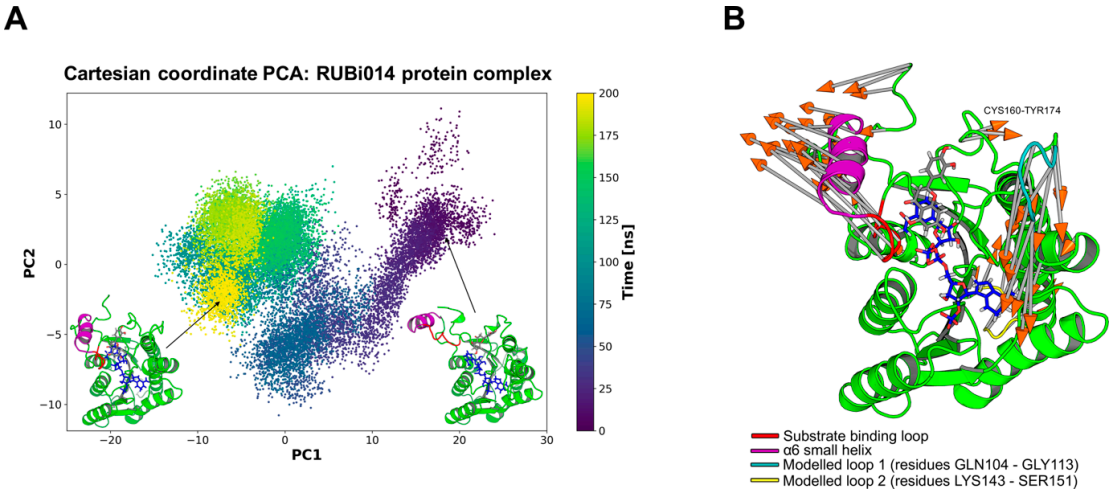


Figure 12. Principal component analysis of the RUBi014 - protein complex. The motions of the complex during 200ns of all atom MD simulation shown along the first and second principal components (PC1 and PC2). The substrate binding loop (residues SER207-GLU215) is colored red, α6 alpha helix (residues GLY214 - VAL225) is colored magenta, modelled missing residues loop 1 colored cyan, modelled missing loop residues loop2 colored yellow, NADPH cofactor colored blue, and RUBi014 colored gray. PC1 explained 66% of the variance while PC2 explained 12%. A)

Projection of the protein-ligand complex dynamics along the PC1 and PC2. B) The differential motions described by PC1 and PC2 are shown by light gray arrows with orange tips.

Binding free energy: RUBi014 bound to the protein stably throughout the MD simulation with a free binding energy of -56.399 ± 11.696 kJ/mol (Table 2). A per residue decomposition showed that residues PHE97, MET163, PHE171, and TYR174 contributed -11. 369 kJ/mol, -5.340 kJ/mol, -3.000 kJ/mol, and -4.34 kJ/mol respectively (Figure 15A and Figure 15B-iii). ASP161 gave an unfavorable energy contribution of 10.130 kJ/mol (Figure 15A and Figure 15B-iii). ASP161, MET163, TYR174 all showed increases in flexibility in their RMSF (Figure S5N), and CYS160 that is covalently bonded to ASP161 showed an increase in *average BC* (Figure S7C)

Dynamic residue network: Comparative DRN analysis between apo protein and *Tb*PTR1-RUBi014 complex revealed significant differences especially in the substrate binding loop and the small α6 helix (Figure S7C). A general trend between three metrics (*Average BC*, L^{-1} and $RMSF^{-1}$) was again in agreement (Figure S7C; Table S4. In comparison to the apo protein, residues VAL211-LYS218 showed increased *Average L* (Figure S7C). The observed increases in *Average L* correlated well with residue fluctuations (Figure S7C and Table S4) [41]. The *Average L* of the substrate binding loop and the small α6 helix did not change very much during the simulation as shown by in Figure S7C.

In the *Tb*PTR1-RUBi014 complex, residues CYS160, GLY205, PRO210 and SER233 showed increases in average BC. According to our MSA results, CYS160 and PRO210 are conserved only among *Trypanosoma* while SER233 is conserved only among the trypanosomatids (Figure 2). Residue GLY205 is, on the other hand, conserved among all the PTR1 orthologues (Figure 2).

BC related residue interaction analysis: As seen in the *Tb*PTR1-RUBi004 complex, the residues that showed increases in average BC were involved in ligand interaction. Residue CYS160 was covalently bound to ASP161 and had vdW interactions with PRO204 and (Figure S8-4-A). Residue PRO205 formed a hydrogen bond with RUBi014 and was covalently bonded to residue PRO204 (Figure S8-4-B). Residue PRO210 formed vdW interactions with LEU208 and VAL211 (Figure S8-4-C). Lastly, residue SER233 had vdW interactions with SER207 (Figure S8-4-D). *Tb*PTR1 residues ASP161, PRO204, and GLY205 were shown to be important in RUBi014 binding (Figure 7C) while residues SER207, LEU208, and VAL211 are located in the substrate binding loop (Figure 2).

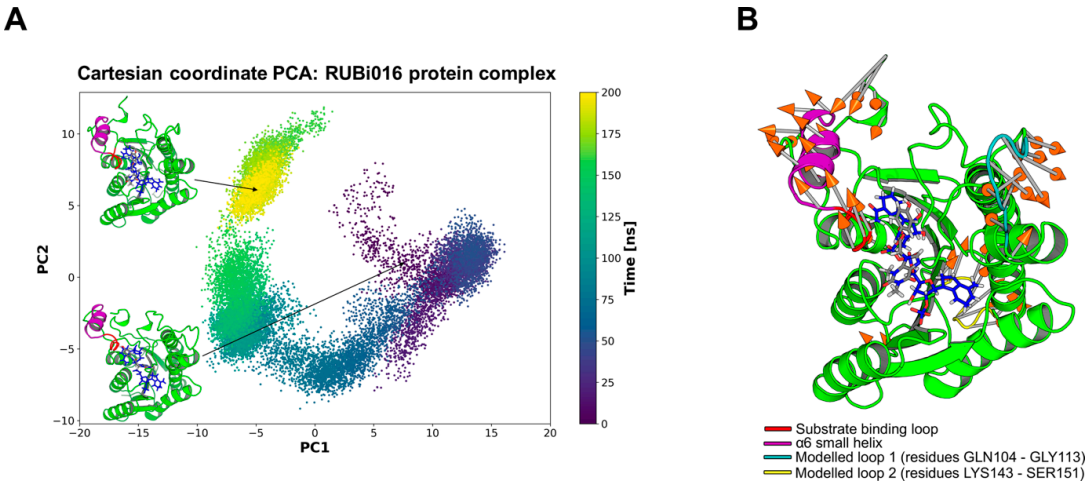
2.4.5. Compound RUBi016 analysis

Ligand docking: Analysis of the *Tb*PTR1-RUBi016 complex showed that compound formed a π-alkyl interaction with the NADPH nicotinamide ring and TYR98. It formed hydrogen bonds with ALA96 (RUBi016 N2 acceptor; ALA96 HA donor), and LEU208 (RUBi016 H14 donor; LEU208 O acceptor) (Figure 7D). RUBi016 formed vdW interactions with ARG14, SER95, PHE97, LEU209, and PRO210 (Figure 7D). In the *Tb*DHFR-RUBi016 complex, the compound formed hydrogen bonds with the NADPH cofactor, TYR166 and VAL32 along with π-π interactions with PHE58 and ILE160 (Figure 8D). As seen in the other compounds this binding mode is similar to that of known *Tb*DHFR inhibitors [38].

Molecular dynamics: The RMSD calculations showed that the *Tb*PTR1-RUBi016 complex was very stable with minimal conformational changes (Figure S2S). The complex showed a linear trajectory with the protein backbone, the NADPH cofactor, and RUBi016 were all very stable (Figure S2S). We observed a slight decrease in the Rg due to the binding of the ligand (Figure S3P). We observed a significant increase in the flexibility of the substrate binding loop and the small α6 helix because of the ligand binding (Figure S5P and Figure S7D). Like in the *Tb*PTR1-RUBi014 complex it appears the binding of the RUBi016 ligand destabilizes the active site (Figure S5P). Furthermore, we observed decreased flexibility of residues SER96 and TYR98 while residues PRO167 and LEU208 – GLY228 showed increased flexibility (Figure S5P and Figure S7D). RMSF analysis and PCA analysis showed that RUBi016 had a similar alteration in the motion of the substrate binding loop, α6 helix, and C-terminal residues HIS267-ALA268 as observed in the *Tb*PTR1-RUBi014 complex (Figure S5P and Figure 13). RUBi016 had favorable binding energy interactions with the substrate binding loop

546 residue PRO210 (Figure 15D). On average the ligand formed three hydrogen bonds with the protein
547 during the MD simulation and these were with ALA96, LEU208, and the NADPH cofactor (Figure
548 S4P and Figure 7D).

549



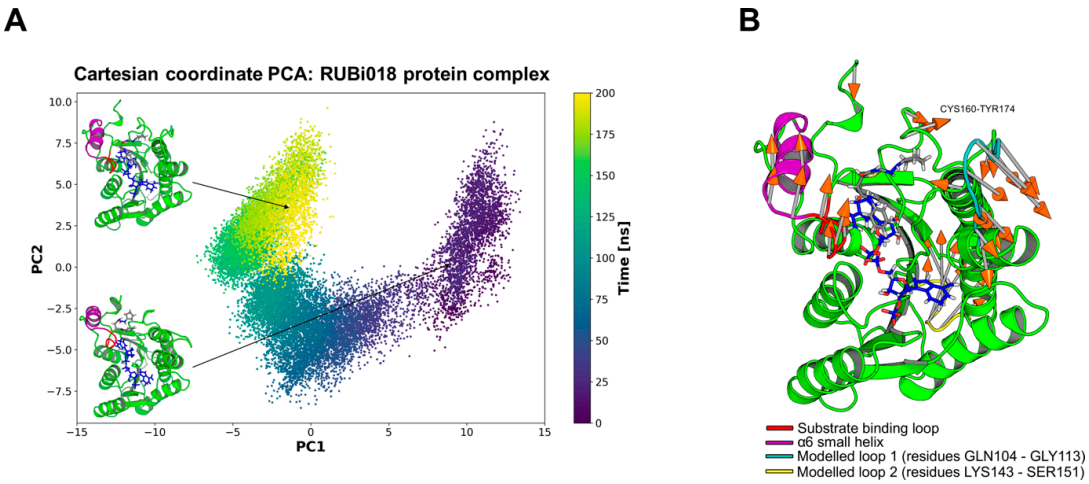
550 **Figure 13.** Principal component analysis of the RUBi016 - protein complex. The motions of the
551 complex during 200ns of all atom MD simulation shown along the first and second principal
552 components (PC1 and PC2). The substrate binding loop (residues SER207-GLU215) is colored red, α6
553 alpha helix (residues GLY214 - VAL225) is colored magenta, modelled missing residues loop 1
554 colored cyan, modelled missing loop residues loop2 colored yellow, NADPH cofactor colored blue,
555 and RUBi016 colored gray. PC1 explained 57% of the variance while PC2 explained 20%. A)
556 Projection of the protein-ligand complex dynamics along the PC1 and PC2. B) The differential
557 motions described by PC1 and PC2 are shown by light gray arrows with orange tips.

558 ***Binding free energy:*** RUBi016 bound to the protein stably throughout the MD simulation with a
559 free binding energy of -23.353 ± 10.361 kJ/mol (Table 2). A per residue decomposition showed that
560 residues ALA96, PHE97, TYR98, and PRO210 contributed -4.202 kJ/mol, -7.992 kJ/mol, -5.497 kJ/mol,
561 and -4.590 kJ/mol respectively. ARG14 and GLU122 had unfavorable energy contributions of 3.230
562 kJ/mol and 8.980 kJ/mol respectively (Figure 15A and Figure 15B-iv). ALA96, PHE97 and TYR 98 all
563 showed reduced flexibility (Figure S5P). This coincided with a reduction in the centrality of SER95
564 (Figure S7D). PRO210 showed increased flexibility and centrality (Figure S5P and Figure S7D).
565 ***Dynamic residue network***: Comparison of the apo protein to the *Tb*PTR1-RUBi016 complex
566 using *Average L* indicated spatial displacement of the substrate binding loop and the small α6 helix
567 (Figure S7D). In comparison to the apo protein, residues VAL211-LYS218 showed increased *Average*
568 *L* (Figure S7D). The observed increases in *Average L* correlated well with residue fluctuations (Figure
569 S7D; Table S4) [41]. The substrate binding loop and the small α6 helix *Average L* showed minimal
570 changes during the simulation as shown by in Figure S7D. SER95 and PHE97 became less central
571 while CYS160, GLY205, PRO210, and SER264 became more central in the *Tb*PTR1-RUBi016 complex
572 (Figure S7D; Table S4). Multiple sequence analysis showed that residue SER95 was conserved
573 among the trypanosomatids only while SER264 was conserved only in *Tb*PTR1 and *Lm*PTR1 (Figure
574 2).
575 ***BC related residue interaction analysis***: SER95 formed vdW interactions with the RUBi016
576 ligand, ALA96, and the NADPH cofactor. Residue CYS160 formed vdW interactions with PRO204
577 and alkyl interactions with ALA203 (Figure S8-5-B). Residue GLY205 formed vdW interactions with
578 the NADPH cofactor, ALA203, PRO204, VAL206 and SER264 (Figure S8-5-C). Residue PRO210
579 formed vdW interactions with the ligand while SER264 formed vdW interactions with GLY205 and
580 VAL206 (Figure S8-5-D).
581

582 2.4.6. Compound RUBi018 analysis

583 ***Ligand docking:*** Analysis of the *Tb*PTR1-RUBi018 complex showed that compound formed
584 T-shaped π-π interactions with PHE97 and PHE171. RUBi018 also formed alkyl and π-alkly
585 interactions with MET163, CYS168, and HIS267 (Figure 7E). RUBi018 formed vdW interactions with
586 ASP161, VAL164, TYR174, GLY205, VAL206, PRO210, MET213, TRP221, LEU263, and the NADPH
587 cofactor (Figure 7E). In the *Tb*DHFR-RUBi016 complex, the compound formed hydrogen bonds with
588 the NADPH cofactor and SER89 along with π-π interactions with PHE58 in a similar binding mode
589 to known *Tb*DHFR inhibitors [38] (Figure 8E).
590 ***Molecular dynamics:*** The RMSD calculations showed that the *Tb*PTR1-RUBi018 complex was
591 stable. The protein backbone and NADPH cofactor were stable while the ligand showed some
592 instability (Figure S2U). We did not observe any major changes in the Rg due to the binding of the
593 ligand (Figure S3S). We observed an increase in the flexibility of the substrate binding loop (PRO210
594 – MET213) because of the ligand binding (Figure S5S). Furthermore, we observed decreased
595 flexibility of residues PHE97 and MET163 while PRO167, CYS168, and PHE171 showed increased
596 flexibility (Figure S5S and Figure S7E). RMSF analysis and PCA analysis showed that RUBi018 had
597 the largest motions in the substrate binding loop, the α6 helix, CYS160-TY174 loop region,
598 C-terminal residues ALA268, the modelled loop 1, and modelled loop 2 (residues LYS143-SER151)
599 (Figure 14). RUBi018 had favorable binding contributions from α6 helix residue TRP221, loop
600 residues MET163, VAL164, PHE171, and C-terminal residue HIS267.
601

602



603 **Figure 14.** Principal component analysis of the RUBi018 - protein complex. The motions of the
604 complex during 200ns of all atom MD simulation shown along the first and second principal
605 components (PC1 and PC2). The substrate binding loop (residues SER207-GLU215) is colored red, α6
606 alpha helix (residues GLY214 - VAL225) is colored magenta, modelled missing residues loop 1
607 colored cyan, modelled missing loop residues loop2 colored yellow, NADPH cofactor colored blue,
608 and RUBi018 colored gray. PC1 explained 44% of the variance while PC2 explained 21%. A)
609 Projection of the protein-ligand complex dynamics along the PC1 and PC2. B) The differential
610 motions described by PC1 and PC2 are shown by light gray arrows with orange tips.

611 ***Binding free energy:*** RUBi018 bound to the protein stably throughout the MD simulation with a
612 free binding energy of -87.914 ± 12.026 kJ/mol (Table 2). A per residue decomposition showed that
613 residues MET163, VAL164, PHE171, TRP221, and HIS267 contributed -6.747 kJ/mol, -3.210 kJ/mol,
614 -5.715 kJ/mol, -8.716 kJ/mol, and -3.560 kJ/mol respectively (Figure 15A and Figure 15B-v). ASP161
615 had an unfavorable energy contribution of 4.850 kJ/mol (Figure 15A and Figure 15B-v). ASP 165 that
616 is covalently bonded to VAL164 showed increased centrality (Figure S7E).

Dynamic residue network: Comparison of the apo protein to the *Tb*PTR1-RUBi018 complex using *Average L* indicated spatial displacement of the substrate binding loop and the small α6 helix (Figure S7E). The substrate binding loop and the small α6 helix *Average L* showed minimal change during the simulation as shown by in Figure S7E. In comparison to the apo protein, residues VAL211-GLU215 showed increased *Average L* (Figure S7E). The observed increases in *Average L* correlated well with residue fluctuations (Figure S7E; Table S4) [41]. GLY16, ASP165, VAL206, LEU208, PRO210, and ALA232 showed increases in *Average BC* (Figure S7E). Multiple sequence analysis showed that residues GLY16 was conserved among all the PTR1 orthologues (Figure 2). Residue VAL206 was only present in *Tb*PTR1 while the other PTR1 orthologues had a leucine instead (Figure 2). Residue ASP165 and ALA232 were conserved only among the trypanosoma. Furthermore, GLY16 formed vdW interactions with the NADPH cofactor and ASN93 (Figure S8-6-A).

BC related residue interaction analysis: ASP165 had alkyl interactions with the RUBi018 ligand and formed vdW interactions with MET163 (Figure S8-6-B). Residue VAL206 formed vdW interactions with both the RUBi018 ligand and the NADPH cofactor (Figure S8-6-C). Residue LEU208 had a hydrogen bond with the NADPH cofactor and formed vdW interactions with PRO210 (Figure S8-6-D). Residue PRO210 formed vdW interactions with the RUBi018 ligand and PHE97 (Figure S8-6-E). Lastly, residue ALA232 formed a hydrogen bond with VAL206 and SER207 (Figure S8-6-F).

Overall, our *in silico* rational based drug discovery approach was able to identify five compounds that showed anti-trypanosomal *in vitro* activity. The per residue energy contribution offers an interesting insight into the study compounds (Figure 15). Many of the protein-ligand complexes had residues contributing to binding that were of catalytic importance and involved in the binding of known *Tb*PTR1 inhibitors (Figure 15) [16, 19, 26, 27]. Residues LYS13, ARG14, PHE97, MET163, TYR174, and substrate binding residues LEU209-TRP221 appear to enhance ligand interaction as they contribute the most favorably energetically to ligand binding (Figure 15A). RUBi004 and RUBi007 showed NADPH cofactor binding residues LYS13 and ARG14 contributing favorably to binding (Figure 15B-i and Figure 15B-ii). Active site residue ASP161 generally had poor interactions with the ligands where it gave unfavorable energy contribution to ligand binding (Figure 15). ASP161 forms hydrogen bonds with MET163 and TYR174 and is important in proton transfer to the substrate during catalysis [34]. Notably in RUBi016 where there wasn't any interaction with residues MET163 or TYR174 as shown in the docking analysis, ASP161 did not give an unfavorable energy contribution (Figure 7D and Figure 15B-iv). This work provides insight into important *Tb*PTR1 protein-ligand interactions that can be used in rational based drug design to characterize potential inhibitors with the end goal of designing and optimizing HAT anti-folate drugs.

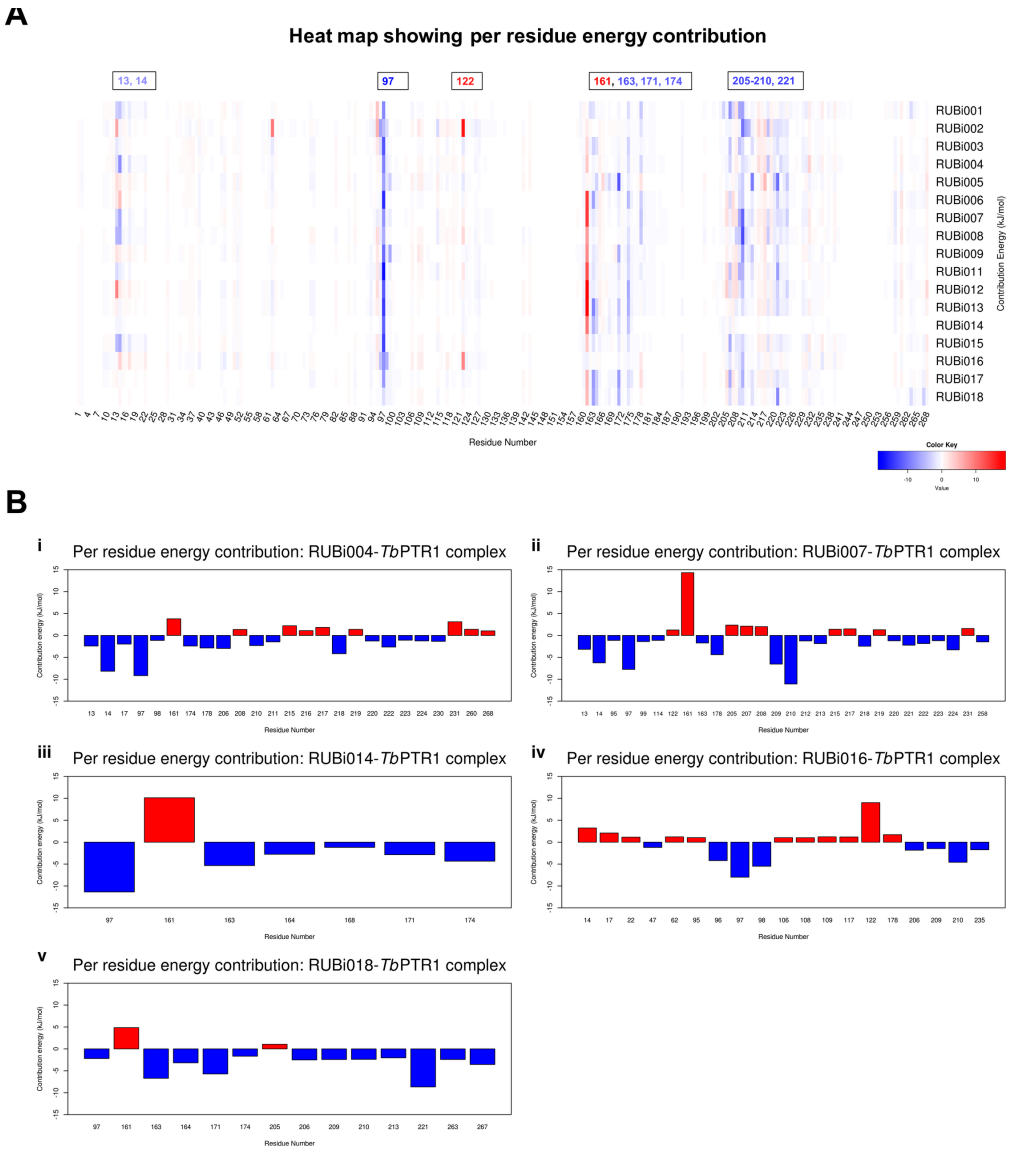


Figure 15. A heat map and histogram showing the per residue energy contributions to binding (energy in kJ/mol). The compounds bound in similar conformations to known pterin, folates and PTR1 inhibitors as is shown by the common contributing residues to binding. A) A heat map showing the per residue energy contributions for all the protein-ligand systems B) A histogram showing the main residues contributing energetically to binding in i) *Tb*PTR1-RUBi004, ii) *Tb*PTR1-RUBi007, iii) *Tb*PTR1-RUBi014, iv) *Tb*PTR1-RUBi016, and v) *Tb*PTR1-RUBi018 complexes.

3. Discussion

In this study, structure based molecular docking was used to screen 5742 selected compounds against trypanosomatid PTR1s and human homolog (*Hs*DHSR4) to identify potential hits for HAT. Eighteen compounds showed good selectivity for trypanosomatid PTR1s, and only compound RUBi006 bound to the *Hs*DHSR4 active site but with a weaker binding energy than the trypanosomatids. MD simulations, DRN calculations, and MMPBSA free energy calculations indicated that all 18 compounds were potentially good hits. Of the 18, 13 commercially available compounds were tested for anti-trypanosomal activity using *in vitro* inhibition assays. Five compounds out of the 13 (RUBi004, RUBi007, RUBi014, RUBi16, and RUBi018) exhibited anti-trypanosomal activity against trypanosomes in culture with IC_{50}s of 4.1 μM, 32.6 μM, 16.3 μM, 31.9 μM, and 12.3 μM respectively with no significant human cell cytotoxicity. When used in

combination with WR99210, RUBi004, RUBi007, RUBi014 and RUBi018 displayed antagonistic effects, while RUBi016 showed an additive effect in the isobologram assay.

When anti-folate drugs that target trypanosomatid DHFR-TS are used, PTR1 is over-expressed allowing for a by-pass mechanism to ensure parasite survival [8, 13]. This is the escape mechanism that has hampered the use of traditional anti-folates against trypanosomatids. PTR1 is an important drug target as demonstrated by gene knock out in *Leishmania* [8] and knock down in *Trypanosoma brucei* [17, 34] studies that show that the enzyme is essential for parasite survival. However *Tb*PTR1 is less susceptible to inhibition than *Tb*DHFR [21, 34, 37]. Given the nature of the interaction between *Tb*PTR1 and *Tb*DHFR, a combination therapy would offer several advantages especially against resistance problems as shown by anti-malarial combination treatment strategies [34, 44].

Our experimental data allow us to draw the following conclusions: Compounds RUBi004, RUB007, RUBi014 and RUBi018 inhibited parasite growth with IC_{50}s of 4.1 μM, 32.6. μM, 16.3 μM, and 12.3 μM when assayed on their own. When used in combination with WR99210 (IC_{50} 0.55 μM), a known TBDHFR inhibitor, each compound showed an antagonistic effect. From our molecular docking studies, we demonstrated that it is reasonable that the compounds RUBi004, RUB007, RUBi014 and RUBi018 can bind both *Tb*PTR1 and *Tb*DHFR with good binding affinities and in binding modes similar to those of traditional folates, pterins and known inhibitors. Our molecular dynamics simulations also show that the ligands bind to the *Tb*PTR1 stably and with acceptable binding energies. In line with these results we theorize that the compounds could be competing for the *Tb*DHFR active site with WR99210 which would result in the observed antagonism. Further, the resulting over-expression of *Tb*PTR1 would result in further reduction in the compound efficacy.

Compound RUBi016 inhibited parasite growth with an IC_{50} 31.9 μM when assayed on its own. From our molecular docking results, it appears that RUBi016 binds *Tb*DHFR with a lower binding affinity than *Tb*PTR1: -7.6 kcal/mol and -8.9 kcal/mol respectively. During MD simulation it bound to *Tb*PTR1 stably and had a good binding energy. Unlike the four compounds, RUBi016 showed an additive effect when used in combination with WR99210. We theorize that RUBi016 may be more selective for *Tb*PTR1, compared to the other compounds which may have a tendency to bind to both PTR1 and DHFR. The other compounds thus display antagonistic effects with WR99210, while RUBi016 is additive. In the case of RUBi016, the addition of WR99210 inhibits *Tb*DHFR further impeding folate reduction resulting in more inhibition of parasite survival hence the observed additive effect. Further, the fact that RUBi016 does not have the lowest IC_{50} does not exclude the possibility that it might be the most selective of the compounds. Interestingly, the ratio of its binding energy for PTR1 and DHFR is the highest (except RUBi018 for PTR1), which means the docking scores would predict it to be the most selective. For example, RUBi004 has the lowest IC_{50} and the lowest binding energy, but it also has a very low binding energy for DHFR which means it is not really selective.

We note that compounds RUBi004 and RUBi014 contain PAINS features [45]. While it is a cause for concern, there are over 60 FDA approved drugs that have PAINS features [45, 46]. The binding patterns of all the five compounds are consistent with those of folates, pterins and known inhibitors. Further analyses such as ligand structure activity relationship (SAR) analysis and protein-ligand co-crystallizations, are required to validate these compounds as potential HAT anti-folate chemotherapeutics [47].

4. Materials and Methods

4.1. Ligand Library Preparation

The small-molecule ligands were obtained from the South African Natural Compounds database (SANCDB) [48] and the ZINC database (ZINC15) [49]. The compounds in the ZINC dataset already conform to the Lipinski rules, and are commercially available drug-like compounds [49, 50]. The ligand library was prepared by filtering 10 639 555 compounds from the ZINC Drugs Now subset [49, 50] for compounds with XlogP ≤ 3, fewer than four rotatable bonds, at least 2 hydrogen donors, a net charge of zero and a molecular weight ≤ 490. The ZINC subset was reduced to 5107

compounds after filtering, while the SANCDB contained 635 compounds. The final ligand library screened comprised of 5742 compounds.

4.2. Preparation of Protein-Ligand Complexes

The crystal structure of *Tb*PTR1 that has a resolution of 1.15 Å was retrieved from RCSB Protein Data Bank (PDB:2X9N) [26]. Multiple sequence analysis, using MUSCLE [51], was carried out to analyze *Tb*PTR1 (Uniprot: O76290) and the homologues sequences including *T. cruzi* (Uniprot: O44029), *L. major* (Uniprot: Q01782, PDB: 1E92), and *H. sapiens* Dehydrogenase/reductase SDR family member 4 (DHRS4) (Uniprot: Q9BTZ2, PDB: 3O4R). The crystal structure of *Tb*DHFR was also retrieved and had a resolution of 2.2 Å (PDB: 3QFX) [38]. Homology modelling was done using in-house Python scripts to fix missing residues in 2X9N (residues GLN104 – GLY113 and LYS143 – SER151) as well as to generate a homotetramer *Tc*PTR1 structure from its PTR2 isoform (Uniprot: Q8I814, PDB: 1MXH). The modelling was done by MODELLER (version 9.19) using the 'automodel' class and included the NADPH cofactor [52, 53]. For both *Tb*PTR1 and *Tc*PTR1, of the 100 models generated, the top models were validated using the ProSA [54] online server (Figure S1). A table gathering a summary of the *Tb*PTR1, *Tc*PTR1, *Lm*PTR1, *Hs*DHSR4, and *Tb*DHFR protein structures is presented in Table S1.

We carried out blind docking of the ligand library against *Tb*PTR1, *Tc*PTR1, *Lm*PTR1 and *Hs*DHSR4 tetrameric protein structures that included their NADPH cofactors using Autodock Vina (version 7.4) [55]. Later the compounds were also blind docked to the *Tb*DHFR (PDB:3QFX) [38] dimeric structure that included its NADPH cofactors using Autodock Vina. The docking parameters used for each of the proteins are summarized in Table S2. Protein-ligand complexes were then evaluated based on if the ligand is located in the active site, as well as based on binding mode, selectivity, docking energy scores, and hydrogen bonding. The docking energies were further evaluated by re-docking the compounds to their protein targets and then using Xscore to give an independent energy score [31].

4.3. Prediction of blood-brain barrier permeability

To prioritize compounds that can cross the blood-brain barrier (BBB), a principal component analysis (PCA) was carried out to identify which compounds occupied the same chemical space as known central nervous system (CNS) permeable drugs [56–58]. The PCA was based on the molecular descriptors of the top binding compounds, Food and Drug Administration (FDA) approved drugs, and FDA approved CNS permeable drugs [49]. The molecular descriptors used included XlogP, number of H-bond donors (HBD), number of H-bond acceptors (HBA), net charge (NC), topological polar surface area (tPSA), molecular weight (MWT), number of rotatable bonds (NRB), polar and apolar desolvation. The first and second principal components were used to create a scatter plot that explained largest percentage of the variance.

4.4. Molecular Dynamics

Eighteen protein-cofactor-ligand complexes were then parametrized using the AMBER03 force field utilizing ACPYPE [59] and GROMACS (5.1.4) [60]. Each protein-ligand complex was solvated using a Simple Point Charge (SPC) water model in a cubic box of 5.07 × 5.18 × 5.16 (nm) with a 2.37 × 2.86 × 3.30 (nm) center. A minimum distance of 1.5 nm was allowed between any protein or ligand atom with the wall. The systems were then neutralized using Na^+ and Cl^- counter ions. The MD systems also included simulating the protein in complex with the NADPH cofactor without the ligand. The MD simulations were performed using GROMACS 5.1.4 [60]. To correct for any structural distortions, the systems were minimized using a steepest descent algorithm using a 100 kJ/mol/nm tolerance value. This was followed by an equilibration using the NPT (constant number of particles, pressure, and temperature) and NVT (constant number of particles, volume, and temperature) ensembles. This was then finally followed by a 200 ns production run at 300 K without any restraints. Trajectories were generated in every 3 fs and saved after every 10 ps. The MD

trajectory analysis included Root Mean Square Deviation (RMSD), Radius of gyration (Rg), Root Mean Square Fluctuation (RMSF), and Principal Component Analysis (PCA) using the GROMACS toolbox, Visual Molecular Dynamics (VMD) [61], and ProDy [62].

Graphs and diagrams were generated using, JChemPaint [63], PyMOL [29] and GRACE software (http://plasma-gate.weizmann.ac.il/Grace/). Protein-ligand complex structures were generated from the equilibrated trajectories at the end of the simulation. These structures were then used to analyze the protein-ligand interactions as well as residue interactions using Discovery Studio [30].

4.5. MM-PBSA Free Energy Calculations

The last 50 ns of the equilibrated MD trajectories were used to perform binding free energy (BFE) calculations of the ligand-protein complexes using the g_mmpbsa package (version 1.6) [64]. The BFE calculation was based on the Molecular Mechanics/Poisson-Boltzman Surface Area (MM-PBSA) method [65, 66]. The BFE of the protein-ligand complexes is calculated using the equations below (in general terms):

$$\Delta G_{binding} = G_{complex} - (G_{protein} + G_{ligand}, \quad (1)$$

$$G_x = \langle E_{MM} \rangle - TS + \langle G_{solvation} \rangle, \quad (2)$$

$$E_{MM} = E_{bonded} + E_{nonbonded} = E_{bonded} + (E_{electrostatic} + E_{vdW}), \quad (3)$$

$$G_{solvation} = G_{polar} + G_{nonpolar}, \quad (4)$$

(1) The binding free energy of the protein-ligand complex in solvent ($\Delta G_{binding}$) where $G_{complex}$ is described as the total energy of the protein-ligand complex. $G_{protein}$ is the isolated free energy of the protein while and G_{ligand} is the isolated free energy of the ligand.

(2) The free energy of either the ligand, protein or protein ligand complex (x). The average mechanical potential in a vacuum is described as $\langle E_{MM} \rangle$. While TS describes the entropic contribution (T is temperature and S is entropy) and $G_{solvation}$ describes the free energy of solvation.

(3) The vacuum molecular mechanics potential energy E_{MM} where E_{bonded} are bonded interactions such as bonds, dihedrals, angles and improper interactions. The non-bonded interactions ($E_{nonbonded}$) are modelled using the Coulomb and Lennard-jones (LJ) potential function. They include: electrostatic interactions ($E_{electrostatic}$) and van der Waals (E_{vdW}) interactions.

(4) The energy required to transfer the protein-ligand solute from a vacuum into a solvent is described as the free energy of solvation ($G_{solvation}$). G_{polar} and $G_{nonpolar}$ describe the electrostatic and non-electrostatic energy contributions respectively. [64]

Furthermore, to determine the energy contribution of each protein residue that binds with the ligand, a free energy decomposition was carried out using g_mmpbsa [64]. This allowed a better understanding of the protein-ligand interactions and helped identify PTR1 binding residues of functional significance.

4.6. Average shortest path (Average L), and Average Betweenness Centrality (Average BC)

We carried out dynamic network analysis on the equilibrated (after 50ns) apo protein and protein-ligand MD trajectories using MD-TASK [39] in order to identify changes in the topological properties of the proteins brought about by the ligand interactions. This was used to glean the impacts of ligand binding on protein dynamics, function, and conformation. A cut off of 6.7 Å was used in the creation of the dynamic residue networks in MD-TASK. The average shortest path (*Average L*) gave the density of shortest paths (*L*) between all node pairs [39]. The average Betweenness centrality (*Average BC*) was used to identify residues in the dynamic network that were important for communication flow. Additionally, by comparing the apo protein and the protein-ligand complexes we were able to use *Average BC* to assess how communication flow across

the dynamic network was altered by ligand binding during the MD simulation. We generated equilibrated structures at the end of the simulations in order to map the interaction networks of any identified important residues using Discovery Studio [30].

4.7. Trypanosoma in vitro Inhibition Assay

Compounds RUBi001, RUBi005, and RUBi015 were purchased from MCULE while RUBi002, RUBi004, RUBi007, RUBi008, RUBi010, RUBi011, RUBi012, RUBi014, RUBi016, and RUBi018 from MolPort (Not all the compounds were commercially available). These compounds were assayed for trypanocidal activity by adding 20 uM of each compound to cultures of *T. b. brucei* (strain Lister 427) in 96-well plates. The parasites were maintained at 37 °C and 5% CO_2 in IMDM medium containing 25 mM HEPES, 10 % fetal bovine serum, 1 mM hypoxanthine, 0.05 mM bathocuproine disulfonic acid, 1.5 mM cysteine, 1.25 mM pyruvic acid, 0.09 mM uracil, 0.09 mM cytosine, 0.16 mM thymidine and 0.014 % 2-mercaptoethanol. Parasites were diluted to 2.4 x 10^4 cells in a volume of 200 µl per well and incubated with the test compounds for 24 hours. Parasite percentage viability was determined using the resazurin method [67]. Twenty µl 0.5 mM resazurin in phosphate-buffered saline was added to each well and incubation continued for a further 24 hours, after which fluorescence (Ex_{560}/Em_{590}) was read in a Spectramax M3 microplate reader (Molecular Devices). Trypanocidal activity of the compounds was reported as the percentage of viable parasites in the compound treated wells when compared to untreated controls (% viability). Pentamidine, an FDA approved trypanocidal drug, was used as the control drug standard [68]. For compounds that produced < 20% viability, IC_{50} values were subsequently determined. The assays were conducted as described above, except that parasites were incubated with 3-fold serial dilutions of the test compounds and IC_{50} values derived from % parasite viability vs. log[compound] dose-response plots by non-linear regression analysis using GraphPad Prism (version 5.02). To assess compound interactions, the compounds were assayed for trypanocidal activity when used in combination with WR99210, a known *Tb*DHFR inhibitor [38]. For combination assays, IC_{50} values were determined for RUBi004, RUBi007, RUBi014, RUBi016 and RUBi018 as well as WR99210 alone using a starting concentration of 100 µM and 20 µM (for the RUBi compounds and WR99210, respectively), and in combinations at ratios of 75:25, 50:50 and 25:75 respectively (thus starting concentrations of 75 µM/5 µM, 50 µM/10 µM and 25 µM/15 µM for the RUBi compounds/WR99210). For isobologram analysis, the fractional inhibitory concentrations (FIC) of the RUBi compounds and WR99210 were calculated by dividing the IC_{50}s obtained for the compounds at the various combination ratios with the IC_{50} obtained for the compounds in the absence of partner drug, and the FIC values plotted against each other (RUBi compound FIC *vs.* WR99210 FIC).

4.8. In vitro Human Cytotoxicity Assay

The compounds assayed for trypanocidal activity were also tested to determine if they caused adverse effects against human cells *in vitro*. For this assay, HeLa (human cervix adenocarcinoma) cells were used. The cells were cultured in DMEM supplemented with 10 % fetal calf serum and antibiotics (penicillin/streptomycin/amphotericin B) at 37 °C in a 5 % CO_2 incubator. Cells were plated at a density of 2 x 10^4 cells/well and after an overnight incubation the compounds were assayed for cytotoxic activity by adding 20 µM of each compound to the 96-well plates, followed by incubation for 48 hrs. Cell viability was determined using the resazurin method [67]. Resazurin (0.5 mM in phosphate-buffered saline; 20 µl/well) was added to the cells and, after a 2-hour incubation, fluorescence was read in a Spectramax M3 plate reader at excitation and emission wavelengths of 560 nm and 590 nm, respectively. Fluorescence readings were converted to percentage cell viability relative to control wells untreated with compounds. Emetine, a drug that induces cell apoptosis, was used as a control [69], and produced an IC_{50} of 0.013 µM.

4.9. Pan-Assay Interference Compounds (PAINS) Assay

Compounds that showed trypanocidal activity were also subjected to the Pan Assay Interference compounds (PAINS) assay using the web server located at http://www.cbligand.org/PAINS/. This was done to identify and flag any compounds that contained PAINS features [45, 70].

5. Conclusions

Dual inhibition of *Tb*PTR1 and *Tb*DHFR is a promising approach to successfully developing safe and effective anti-folate based anti-trypanosomal chemotherapeutics. As shown in this study, computation based approaches are useful in fast and rapid rational drug design. Further, in the discovery of novel TbPTR1 inhibitors, when the compounds are assayed in combination with known DHFR inhibitors, careful interpretation of isobologram assays is required to get the most optimal outcome. When used in combination with WR99210, a known *Tb*DHFR inhibitor, compounds RUBi004, RUBi007, RUBi014, and RUBi018 showed moderate to strong antagonism as demonstrated by the isobologram results, which would indicate that they might be binding to both *Tb*PTR1 and *Tb*DHFR. RUBi016 as shown by its additive effect and molecular docking results appears to selectively bind to TbPTR1. The five compounds assayed showed anti-trypanosomal activity with no significant human cell cytotoxicity i*n vitro*. The merging of these scaffolds could yield to the development of even more potent and selective *Tb*PTR1 inhibitors.

Supplementary Materials: The following are available online. Figure S1: Validation of *T. brucei* and *T. cruzi* PTR1 homology models using z-DOPE score and residue score using ProSA. The structural validation of *Tb*PTR1 is shown in A) and B) while that of *Tc*PTR1 in C) and D). Both models show overall reliable structural conformations, Figure S2: Time dependent Root Mean Square Deviation (RMSD) of the MD systems, Figure S3: The Radius of gyration of the protein-ligand complexes, Figure S4: The time evolution of the number of intermolecular hydrogen bonds formed between the compounds and *Tb*PTR1 protein during simulation, Figure S5: The Root mean square fluctuation (RMSF) of all the protein in the protein-ligand complexes., Figure S6: Energetic contribution of the of *Tb*PTR1 residues in binding of the protein-ligand complexes, Figure S7: Per residue Root mean square fluctuations (RMSF), average Betweenness Centrality (*average BC*) and average shortest path (*L*) of *Tb*PTR1 residues, Figure S8: 2 dimensional (2D) *Tb*PTR1 residue interaction networks, Table S1: The protein structures used in molecular docking, Table S2: Molecular Docking Parameters for Autodock Vina, Table S3: Binding modes of top *Tb*PTR1 docked compounds, and Table S4: The Pearson correlation coefficients for RMSF vs *Average L*, Average BC vs 1/(*Average L*), and Average BC vs 1/(RMSF).

Author Contributions: Conceptualization, Özlem Tastan Bishop; Data curation, Magambo Phillip Kimuda; Formal analysis, Magambo Phillip Kimuda, Dustin Laming, Heinrich C. Hoppe and Özlem Tastan Bishop; Funding acquisition, Özlem Tastan Bishop; Methodology, Magambo Phillip Kimuda, Dustin Laming, Heinrich C. Hoppe and Özlem Tastan Bishop; Supervision, Heinrich C. Hoppe and Özlem Tastan Bishop; Validation, Magambo Phillip Kimuda; Writing – original draft, Magambo Phillip Kimuda; Writing – review & editing, Magambo Phillip Kimuda, Heinrich C. Hoppe and Özlem Tastan Bishop.

Funding: This work was supported by the National Research Foundation (NRF) South Africa (Grant Numbers 105267 and 116F229). The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the funder. The bioassays were supported by funding from the S.A. Medical Research Council. The Molecular Dynamics simulations and MM-PBSA calculations were carried out on the Centre for High Performance Computing (CHPC), Cape Town, South Africa.

Acknowledgments: M.P.K. thanks the TrypanoGEN Consortium for its support, and Prof Enock Matovu and Dr Kevin Lobb for their mentorship.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bozza MT, Martins YC, Carneiro LAM, et al (2013) Trypanosoma brucei: Meet the system. Curr Opin Microbiol 3:78 . doi: 10.1016/j.mib.2014.06.007
2. Franco JR, Simarro PP, Diarra A, Jannin JG (2014) Epidemiology of human African trypanosomiasis. Clin. Epidemiol. 6:257–275
3. Funk S, Nishiura H, Heesterbeek H, et al (2013) Identifying Transmission Cycles at the Human-Animal Interface: The Role of Animal Reservoirs in Maintaining Gambiense Human African Trypanosomiasis. PLoS Comput Biol 9: . doi: 10.1371/journal.pcbi.1002855
4. Simarro PP, Diarra A, Ruiz Postigo JA, et al (2011) The Human African trypanosomiasis control and surveillance programme of the World Health Organization 2000-2009: the way forward. PLoS Negl Trop Dis 5:e1007 . doi: 10.1371/journal.pntd.0001007
5. Simarro PP, Cecchi G, Paone M, et al (2010) The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. Int J Health Geogr 9:57 . doi: 10.1186/1476-072X-9-57
6. Barrett MP, Gilbert IH (2006) Targeting of Toxic Compounds to the Trypanosome's Interior. Adv. Parasitol. 63:125–183
7. Fox JT, Stover PJ (2008) Folate-mediated one-carbon metabolism. Vitam Horm 79:1–44 . doi: 10.1016/S0083-6729(08)00401-9
8. Bello a R, Nare B, Freedman D, et al (1994) PTR1: a reductase mediating salvage of oxidized pteridines and methotrexate resistance in the protozoan parasite Leishmania major. Proc Natl Acad Sci U S A 91:11442–11446 . doi: 10.1073/pnas.91.24.11442
9. Berriman M, Ghedin E, Hertz-Fowler C, et al (2005) The genome of the African trypanosome Trypanosoma brucei. Science (80-) 309:416–422 . doi: 10.1126/science.1112642
10. Gangjee A, Jain HD, Kurup S (2008) Recent advances in classical and non-classical antifolates as antitumor and antiopportunistic infection agents: Part II. Anticancer Agents Med Chem 8:205–31 . doi: 10.2174/187152008783497064
11. Zuccotto F, Martin ACR, Laskowski RA, et al (1998) Dihydrofolate reductase: a potential drug target in trypanosomes and leishmania. J Comput Aided Mol Des 12:241–257 . doi: 10.1023/A:1016085005275
12. Sienkiewicz N, Jarosławski S, Wyllie S, Fairlamb AH (2008) Chemical and genetic validation of dihydrofolate reductase-thymidylate synthase as a drug target in African trypanosomes. Mol Microbiol 69:520–533 . doi: 10.1111/j.1365-2958.2008.06305.x
13. Nare B, Hardy LW, Beverley SM, Ptrs L (1997) The Roles of Pteridine Reductase 1 and Dihydrofolate Reductase-Thymidylate Synthase in Pteridine Metabolism in the Protozoan Parasite Leishmania major *. 272:13883–13891
14. Vickers TJ, Beverley SM (2011) Folate metabolic pathways in Leishmania. 1–19 . doi: 10.1042/BSE051xxxx
15. Robello C, Navarro P, Castanys S, Gamarro F (1997) A pteridine reductase gene ptr1 contiguous to a P-glycoprotein confers resistance to antifolates in Trypanosoma cruzi. Mol Biochem Parasitol 90:525–535 . doi: 10.1016/S0166-6851(97)00207-7
16. Gourley DG, Schüttelkopf AW, Leonard GA, et al (2001) Pteridine reductase mechanism correlates pterin metabolism with drug resistance in trypanosomatid parasites. Nat Struct Biol 8:521–525 . doi: 10.1038/88584
17. Sienkiewicz N, Ong HB, Fairlamb AH (2010) Trypanosoma brucei pteridine reductase 1 is essential for survival in vitro and for virulence in mice. Mol Microbiol 77:658–671 . doi:

947 10.1111/j.1365-2958.2010.07236.x
948 18. Cavazzuti A, Paglietti G, Hunter WN, et al (2008) Discovery of potent pteridine reductase inhibitors to
949 guide antiparasite drug development. Proc Natl Acad Sci U S A 105:1448–1453 . doi:
950 10.1073/pnas.0704384105
951 19. Tulloch LB, Martini VP, Iulek J, et al (2010) Structure-based design of pteridine reductase inhibitors
952 targeting African sleeping sickness and the leishmaniases. J Med Chem 53:221–229 . doi:
953 10.1021/jm901059x
954 20. Mpamhanga CP, Spinks D, Tulloch LB, et al (2009) One scaffold, three binding modes: Novel and
955 selective pteridine reductase 1 inhibitors derived from fragment hits discovered by virtual screening. J
956 Med Chem 52:4454–4465 . doi: 10.1021/jm900414x
957 21. Hardy LW, Matthews W, Nare B, Beverley SM (1997) Biochemical and genetic tests for inhibitors of
958 Leishmania pteridine pathways. Exp Parasitol 87:157–169 . doi: S0014489497942077 [pii]
959 22. Nare B, Luba J, Hardy LW, Beverley S (1997) New approaches to Leishmania chemotherapy: pteridine
960 reductase 1 (PTR1) as a target and modulator of antifolate sensitivity. Parasitology 114 Suppl:S101–S110
961 23. Matovu E, Seebeck T, Enyaru JCK, Kaminsky R (2001) Drug resistance in Trypanosoma brucei spp., the
962 causative agents of sleeping sickness in man and nagana in cattle. Microbes Infect 3:763–770 . doi:
963 10.1016/S1286-4579(01)01432-0
964 24. Babokhov P, Sanyaolu AO, Oyibo W a, et al (2013) A current analysis of chemotherapy strategies for the
965 treatment of human African trypanosomiasis. Pathog Glob Health 107:242–52 . doi:
966 10.1179/2047773213Y.0000000105
967 25. Mesu VKBK, Kalonji WM, Bardonneau C, et al (2017) Oral fexinidazole for late-stage African
968 Trypanosoma brucei gambiense trypanosomiasis: A pivotal multicentre, randomised, non-inferiority
969 trial. Lancet
970 26. Dawson A, Tulloch LB, Barrack KL, Hunter WN (2010) High-resolution structures of Trypanosoma
971 brucei pteridine reductase ligand complexes inform on the placement of new molecular entities in the
972 active site of a potential drug target. Acta Crystallogr Sect D Biol Crystallogr 66:1334–1340 . doi:
973 10.1107/S0907444910040886
974 27. Luba J, Nare B, Liang PH, et al (1998) Leishmania major pteridine reductase 1 belongs to the short chain
975 dehydrogenase family: stereochemical and kinetic evidence. Biochemistry 37:4093–4104 . doi:
976 10.1021/bi972693a
977 28. Schormann N, Pal B, Senkovich O, et al (2005) Crystal structure of Trypanosoma cruzi pteridine
978 reductase 2 in complex with a substrate and an inhibitor. J Struct Biol 152:64–75 . doi:
979 10.1016/j.jsb.2005.07.008
980 29. DeLano WL (2002) The PyMOL Molecular Graphics System. Schrödinger LLC wwwpymolorg Version
981 1.:http://www.pymol.org . doi: citeulike-article-id:240061
982 30. San Diego: Accelrys Software Inc. (2012) Discovery Studio Modeling Environment, Release 3.5. In:
983 Accelrys Softw. Inc.
984 31. Obiol-Pardo C, Rubio-Martinez J (2007) Comparative evaluation of MMPBSA and XSCORE to compute
985 binding free energy in XIAP-peptide complexes. J Chem Inf Model 47:134–142 . doi: 10.1021/ci600412z
986 32. Legros D, Ollivier G, Gastellu-Etchegorry M, et al (2002) Treatment of human African
987 trypanosomiasis--present situation and needs for research and development. Lancet Infect Dis
988 2:437–440 . doi: 10.1016/S1473-3099(02)00321-3
989 33. David CC, Jacobs DJ (2014) Principal component analysis: A method for determining the essential

dynamics of proteins. Methods Mol Biol. doi: 10.1007/978-1-62703-658-0_11

34. Dawson A, Gibellini F, Sienkiewicz N, et al (2006) Structure and reactivity of Trypanosoma brucei pteridine reductase: Inhibition by the archetypal antifolate methotrexate. Mol Microbiol 61:1457–1468 . doi: 10.1111/j.1365-2958.2006.05332.x

35. Sawaya MR, Kraut J (1997) Loop and subdomain movements in the mechanism of Escherichia coli dihydrofolate reductase: Crystallographic evidence. Biochemistry 36:586–603 . doi: 10.1021/bi962337c

36. Schnell JR, Dyson HJ, Wright PE (2004) Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase. Annu Rev Biophys Biomol Struct 33:119–140 . doi: 10.1146/annurev.biophys.33.110502.133613

37. Schüttelkopf AW, Hardy LW, Beverley SM, Hunter WN (2005) Structures of Leishmania major pteridine reductase complexes reveal the active site features important for ligand binding and to guide inhibitor design. J Mol Biol 352:105–116 . doi: 10.1016/j.jmb.2005.06.076

38. Vanichtanankul J, Taweechai S, Yuvaniyama J, et al (2011) Trypanosomal dihydrofolate reductase reveals natural antifolate resistance. ACS Chem Biol 6:905–911 . doi: 10.1021/cb200124r

39. Brown DK, Penkler DL, Sheik Amamuddy O, et al (2017) MD-TASK: a software suite for analyzing molecular dynamics trajectories. Bioinformatics. doi: 10.1093/bioinformatics/btx349

40. Penkler D, Atilgan C, Tastan Bishop O (2018) Allosteric Modulation of Human Hsp90α Conformational Dynamics. J Chem Inf Model. doi: 10.1021/acs.jcim.7b00630

41. Atilgan AR, Akan P, Baysal C (2004) Small-World Communication of Residues and Significance for Protein Dynamics. Biophys J 86:85–91 . doi: 10.1016/S0006-3495(04)74086-2

42. van Baren C, Martino V, Di Leo Lira P, et al (2006) Triterpenic Acids and Flavonoids from Satureja parvifolia. Evaluation of their Antiprotozoal Activity. Zeitschrift fur Naturforsch - Sect C J Biosci 61:189–192 . doi: 10.1515/znc-2006-3-406

43. Tasdemir D, Kaiser M, Brun R, et al (2006) Antitrypanosomal and antileishmanial activities of flavonoids and their analogues: In vitro, in vivo, structure-activity relationship, and quantitative structure-activity relationship studies. Antimicrob Agents Chemother 50:1352–1364 . doi: 10.1128/AAC.50.4.1352-1364.2006

44. Olliaro PL (2003) Antimalarial compounds: from bench to bedside. J Exp Biol. doi: 10.1242/jeb.00653

45. Kenny PW (2017) Comment on the Ecstasy and Agony of Assay Interference Compounds. J. Chem. Inf. Model. 57:2640–2645

46. Kilchmann F, Marcaida MJ, Kotak S, et al (2016) Discovery of a Selective Aurora A Kinase Inhibitor by Virtual Screening. J Med Chem 59:7188–7211 . doi: 10.1021/acs.jmedchem.6b00709

47. Dahlin JL, Walters MA (2016) How to Triage PAINS-Full Research. Assay Drug Dev Technol 14:168–174 . doi: 10.1089/adt.2015.674

48. Hatherley R, Brown DK, Musyoka TM, et al (2015) SANCDB: a South African natural compound database. J Cheminform 7:29 . doi: 10.1186/s13321-015-0080-8

49. Irwin JJ, Shoichet BK (2005) ZINC - A free database of commercially available compounds for virtual screening. J Chem Inf Model 45:177–182 . doi: 10.1021/ci049714+

50. Lipinski CA (2000) Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods 44:235–249 . doi: 10.1016/S1056-8719(00)00107-6

51. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797 . doi: 10.1093/nar/gkh340

52. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol

234:779–815 . doi: 10.1006/jmbi.1993.1626

53. Eswar N, Webb B, Marti-Renom M a, et al (2007) Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci Chapter 2:Unit 2.9 . doi: 10.1002/0471140864.ps0209s50

54. Wiederstein M, Sippl MJ (2007) ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35: . doi: 10.1093/nar/gkm290

55. Trott O, Olson A (2010) AutoDock Vina: inproving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455–461 . doi: 10.1002/jcc.21334.AutoDock

56. Clemons PA, Wilson JA, Dancik V, et al (2011) Quantifying structure and performance diversity for sets of small molecules comprising small-molecule screening collections. Proc Natl Acad Sci 108:6817–6822 . doi: 10.1073/pnas.1015024108

57. Singh N, Guha R, Giulianotti MA, et al (2009) Chemoinformatic analysis of combinatorial libraries, drugs, natural products, and molecular libraries Small Molecule Repository. J Chem Inf Model 49:1010–1024 . doi: 10.1021/ci800426u

58. Doniger S, Hofmann T, Yeh J (2002) Predicting CNS permeability of drug molecules: comparison of neural network and support vector machine algorithms. J Comput Biol 9:849–864 . doi: 10.1089/10665270260518317

59. Sousa da Silva AW, Vranken WF, Wang J, et al (2012) ACPYPE - AnteChamber PYthon Parser interfacE. BMC Res Notes 5:367 . doi: 10.1186/1756-0500-5-367

60. Van Der Spoel D, Lindahl E, Hess B, et al (2005) GROMACS: Fast, flexible, and free. J. Comput. Chem. 26:1701–1718

61. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. J Mol Graph. doi: 10.1016/0263-7855(96)00018-5

62. Bakan A, Meireles LM, Bahar I (2011) ProDy: Protein dynamics inferred from theory and experiments. Bioinformatics. doi: 10.1093/bioinformatics/btr168

63. Krause S, Willighagen E, Steinbeck C (2000) JChemPaint - Using the collaborative forces of the internet to develop a free editor for 2D chemical structures. In: Molecules

64. Kumari R, Kumar R, Lynn A (2014) G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations. J Chem Inf Model 54:1951–1962 . doi: 10.1021/ci500020m

65. Homeyer N, Gohlke H (2012) Free energy calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area method. Mol Inform 31:114–122 . doi: 10.1002/minf.201100135

66. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82 . doi: 10.1021/ci100275a

67. Bowling T, Mercer L, Don R, et al (2012) Application of a resazurin-based high-throughput screening assay for the identification and progression of new treatments for human African trypanosomiasis. Int J Parasitol Drugs Drug Resist 2:262–270 . doi: 10.1016/j.ijpddr.2012.02.002

68. Doua F, Miezan TW, Sanon Singaro JR, et al (1996) The efficacy of pentamidine in the treatment of early-late stage Trypanosoma brucei gambiense trypanosomiasis. Am J Trop Med Hyg 55:586–588 . doi: 10.4269/ajtmh.1996.55.586

69. S. Akinboye E (2011) Biological Activities of Emetine. Open Nat Prod J 4:8–15 . doi: 10.2174/1874848101104010008

70. Baell JB, Holloway GA (2010) New substructure filters for removal of pan assay interference

compounds (PAINS) from screening libraries and for their exclusion in bioassays. J Med Chem 53:2719–2740 . doi: 10.1021/jm901137j

Sample Availability: Samples of the compounds used in this study are available from the authors.