Downloaded from http://portlandpress.com/biochemi/article-pdf/doi/10.1042/BCJ20210041/920687/bcj-2021-0041.pdf by UK user on 29 September 202

Crystallographic approach to fragment-based hit discovery against *Schistosoma mansoni* purine nucleoside phosphorylase

Muhammad Faheem^{1,2,3,7}, Napoleão Fonseca Valadares¹, José Brandão-Neto³, Domenico Bellini^{3,4}, Patrick Collins³, Nicholas M. Pearce^{3,5}, Louise Bird³, Juliana Roberta Torini⁶, Raymond Owens³, Humberto Muniz Pereira⁶, Frank Von Delft^{3,5,8}, João Alexandre Ribeiro Gonçalves Barbosa^{1,2}

1- Laboratório de Biofísica Molecular, Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, 70910-900, Brazil. 2-Programa de Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, SGAN 916 Módulo B Avenida W5 - Asa Norte, Brasília, DF, Brazil. 3- Harwell Science and Innovation Campus, Fermi Ave, Didcot, OX11 0DE, UK. 4- MRC LMB, Francis Crick Avenue, Cambridge, CB2 0QH, UK. 5-Structural Genomics Consortium, Old Road Campus Research Build, Roosevelt Dr, Oxford, OX3 7DQ, UK. 6-Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP, 13563-120, Brazil. 7- Department of Biological Sciences, National University of Medical Sciences, The Mall, Abid Majeed Rd, Rawalpindi, Punjab, 46000, Pakistan. 8- Department of Biochemistry, University of Johannesburg, Auckland Park 2006, South Africa.

1. Muhammad Faheem

(faheem08@live.com)

2. Napoleão Fonseca Valadares

(napoleaofv@gmail.com)

3. José Brandão-Neto

(jose.brandao-neto@diamond.ac.uk)

4. Domenico Bellini

(dbellini@mrc-lmb.cam.ac.uk)

5. Patrick Collins

(pcollins@exscientia.co.uk)

6. Nicholas M. Pearce

(n.m.pearce@uu.nl)

7. Louise Bird

(lbird@exscientia.co.uk)

8. Juliana Roberta Torini

(jutorini@gmail.com)

9. Raymond Owens

(ray.owens@rc-harwell.ac.uk)

10. Humberto Muniz Pereira

(hmuniz.pereira@gmail.com)

11. Frank Von Delft

(frank.von-delft@diamond.ac.uk)

12. João Alexandre Ribeiro Gonçalves Barbosa

(joaobarbosa@unb.br)

Corresponding Author:

João Alexandre Ribeiro Gonçalves Barbosa

(joaobarbosa@unb.br)

Contact: +55-61-991010205

ABSTRACT

Several *Schistosoma* species cause Schistosomiasis, an endemic disease in 78 countries that is ranked second amongst the parasitic diseases in terms of its socioeconomic impact and human health importance. The drug recommended for treatment by the WHO is praziquantel (PZQ), but there are concerns associated with PZQ, such as the lack of information about its exact mechanism of action, its high price, its effectiveness – which is limited to the parasite's adult form – and reports of resistance. The parasites lack the *de novo* purine pathway, rendering them dependent on the purine salvage pathway or host purine bases for nucleotide synthesis. Thus, the *Schistosoma* purine salvage pathway is an attractive target for the development of necessary and selective new drugs. In this study, the purine nucleotide phosphorylase II (PNP2), a new isoform of PNP1, was submitted to a high-throughput fragment-based hit discovery using a crystallographic screening strategy. PNP2 was crystallized and crystals were soaked with 827 fragments, a subset of the Maybridge 1000 library. X-ray diffraction data was collected and structures were solved. Out of 827-screened fragments we have obtained a total of 19 fragments that show binding to PNP2. 14 of these fragments bind to the active site of PNP2, while five were observed in three other sites. Here we present the first fragment screening against PNP2.

INTRODUCTION

Schistosoma mansoni is a parasitic trematode that causes the common intravascular infection schistosomiasis (1). Amongst the parasitic diseases, schistosomiasis ranks second in terms of social and economic impact, and public health importance (2). According to the WHO, schistosomiasis has been reported in 78 countries. In 2013 more than 40 million people were treated for schistosomiasis, 261 million required preventive treatment, and nearly 700 million are at risk of infection (1,3). This burden of infection makes schistosomiasis a major health problem, particularly in developing countries.

The *Schistosoma* genome project was established in 1992 to improve the understanding of *Schistosoma* biology, with a focus on the characterization of new genes, the discovery and development of new drug targets and vaccines, and the determination of mechanisms of drug resistance and antigenic variation that enable evasion of the host's immune system (4). The *S. mansoni* 363 megabases (MB) nuclear genome was published in 2009, and a new 364.5 MB version was made available in 2012. This genomic information revealed a total of 10,852 genes encoding over 11,000 proteins, 45 % of which remain without known or predicted function (5–7). The breakthrough in the availability of genomic data allows new opportunities for innovation in the control of schistosomiasis. These data offer a new pipeline for the identification of novel drug targets and vaccine candidates through a system-wide perspective (6,8,9).

Schistosoma lacks the capacity for *de novo* synthesis of purine nucleosides, and is dependent exclusively on the salvage pathway for their purine requirements(10–12). Accordingly, *S. mansoni* acquire purine nucleosides from the host via a purine salvage pathway; this brings attention to the enzymes of the *S. mansoni* salvage pathway as potential drug targets for novel chemotherapy. In the past, the use of purine and purine nucleoside analogues have been successfully exploited as drug targets against several other parasites (13–15). It has been shown

that inhibition of a salvage pathway enzyme, Hypoxanthine–guanine phosphoribosyltransferase (HGPRTase), via siRNA decreased the viability of *Schistosoma* (16). Therefore, the *Schistosoma* purine salvage pathway is a suitable target for the development of novel compounds for the combatting of schistosomiasis.

Purine nucleoside phosphorylase 1 (PNP1) is a member of the purine salvage pathway of *S. mansoni*, and we reported its crystal structure in 2003 (10,17). PNP1 is involved in the conversion of the purine salvage pathway intermediate inosine to hypoxanthine (10). PNP1 catalyses the cleavage of the glycosidic bond in purine ribo- and deoxyribonucleosides, in the presence of inorganic phosphate as a second substrate generating a purine base and ribose-1-phosphate. PNP1 also has been sudied for discovery of inhibitors (17–19).

Purine nucleoside phosphorylase 2 (PNP2) (PDB ID 5CXQ) is a new isoform of PNP, which was identified in the *S. mansoni* genome-wide Open Reading Frame search. The new isoform, which presents a high degree of conservation to PNP1, has three substitutions in the active site, making its catalytic site very different from the PNP1 (20). PNP2 is highly expressed in the cercariae, where PNP1 has its lowest expression level(6,7), which makes it a suitable drug target.

In the past fifteen years fragment-based drug design has been established as an effective alternative to high throughput screening for the identification of hit compounds in drug discovery. Fragment based screening and optimization have attained reliable success in numerous drug discovery project with one approved drug in the market and several other compounds in clinical trials(21,22). This approach allows the study of the interactions of very simple molecules (fragments) with protein targets, providing useful information for drug design. Structural methods allow us to rapidly and effectively explore the complementarity between a protein active site and drug-like molecules via the use of fragments (23). Recent advancements in the synchrotron facilities with collection of greater amounts of high-quality

diffraction data, and automation in the data analysis has increased the utility of fragment screening (23,24).

PNP-based drug design is an apposite target as it has been explored for other diseases (25), and there are no reports of a high throughput fragment screening for *S. mansoni*. The reproducibility of PNP2 crystals is highly consistent and this makes it a tempting target for fragment screening. Here, we report an extensive crystal optimization and dimethyl sulfoxide (DMSO) tolerance test for PNP2 crystals, and a subsequent fragment-based screening of 827 compounds (a subset of the Maybridge fragment library). This resulted in 19 new PNP2 crystal structures with bound fragments. The majority of the fragments were observed in the active site, including a fragment that explores a previously unidentified pocket closer to α 6- and α 8-helices. Furthermore, fragments were observed in three other binding sites. Our findings reveal a great deal of atomic-resolution structural information regarding the interaction of fragments with PNP2, and a validated methodology to improve crystal quality for fragment screening campaigns.

MATERIAL AND METHODS

Protein Expression and Purification

The PNP2 gene (Smp_179110) was identified in the *S. mansoni* genome (5). The PNP2 open reading frame (ORF) gene was synthesized with codon optimization by GenScript and cloned into pOPINS3C (26) using the In-Fusion method (Clontech, Mountain View, CA, United States). Recombinant pOPINS3C-PNP2 vector was transformed into *Escherichia coli* (*E. Coli*) OmniMaxII cells, positive transformants were selected using the chromogenic substrate X-gal and by colony PCR. Positive white colonies were harvested and cultured in Power Prime BrothTM medium (Molecular Dimensions, UK) supplemented with 50 µg/mL carbenicilin and 30 µg/mL chloramphenicol in a shaker at 37 °C and 220 rpm. The plasmid mini-preps were

performed on the Bio-Robot 8000, according to the manufacturer's protocols. The purified recombinant plasmid (pOPINS3C-PNP2) was transformed in E. coli Lemo21 (DE3). A single colony of transformed cells was cultured overnight in 50 mL power prime broth media supplemented with 50 µg/mL carbenicillin and 35 µg/mL chloramphenicol at 37 °C and 220 Biochemical Journal. This is an Accepted Manuscript. You are encouraged to use the Version of Record that, when published, will replace this version. The most up-to-date-version is available at https://doi.org/10.1042/BC.20210041 rpm. 25 mL of the overnight cultured cells were added to 1 L of sterile power prime broth media supplemented with 50 µg/mL carbenicillin and 35 µg/mL chloramphenicol and incubated in a shaker at 37 °C and 220 rpm to an OD₆₀₀ of 0.6. The cells were transferred to room temperature for 30 min, induced with a final concentration of 1 mM IPTG and incubated overnight at 20 °C and 220 rpm. Cells were harvested by centrifugation at 5000 g for 15 min at 4 °C and re-suspended in lysis buffer (50 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 10 mM imidazole). Cells were disrupted with a constant cell disruption system under pressure at 25 KPSI (CONSTANT SYSTEMS Ltd, UK) and centrifuged at 34,000 g for 20 min at 4 °C. The supernatant was collected and applied to a 5 mL Ni Sepharose column (GE) previously equilibrated with 6 column volumes of lysis buffer. The column was washed with 6 column volumes of lysis buffer and the protein was eluted in a gradient purification length of 6-column volume with elution buffer (50 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 500 mM imidazole). The purified protein was mixed with human rhinovirus (HRV) 3C protease to cleave the his-tag. Cleavage was performed in 14 kDa dialysis membrane in dialysis buffer (50 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 22 mM imidazole, 1 mM tris(2-carboxyethyl)phosphine (TCEP)) with continuous stirring overnight at 4 °C. Cleaved protein was re-applied to a Ni Sepharose column (GE) equilibrated with 6 column volumes of the same overnight dialysis buffer. The column flow through was collected which contains his-tag free protein by observing the UV₂₈₀ chromatogram. The HRV 3C protease has a His-tag and was also bound to the affinity column and separated from the purified protein. The purified protein was concentrated and injected on a Superdex 75 16/600 GL (GE) column previously equilibrated with 2 column volumes of purification buffer (20 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 5 % glycerol, 5 mM β mercaptoethanol). Fractions from all the purification steps and overnight cleavage were visualized on Novex Pre-Cast SDS PAGE gels (Life Technologies).

Crystallization Screening

The purified protein was concentrated to 4.6 mg/mL in purificatication buffer (20 mM NaH₂PO₄ pH 7.5, 300 mM NaCl, 5 % glycerol, 5 mM β-mercaptoethanol). Crystallization experiments were setup in Greiner Crystal-Quick sitting drop plates (Hampton Research) with a Cartesian robot (Digilab Microsystem, England) using the crystallization kits SALT-RX, INDEX HT, PEG/Ion from Hampton Research and PACT premier, Morpheus, JCSG-Plus from Molecular Dimensions. The sitting drop volume setup was 100 nL of protein with 100 nL of reservoir solution. Plates were incubated at 20 °C and crystals appeared in all the screens in different conditions after 2 days. Crystals were selected from each kit on the basis of size, mounted and stored in liquid nitrogen for data collection. X-ray diffraction data were collected at beamline I04-1 of the Diamond Light Source (DLS) (UK) and autoprocessed. The diffraction resolution values were obtained via the autoprocessing pipeline at DLS (27–29), which employed the PNP2 structure (PDB ID 5CXQ) as a template in the molecular replacement procedures.

DMSO tolerance test

Crystals were reproduced in 3 well crystallization Swissci plates using a Cartesian robot for INDEX HT condition G7 (0.2 M C₂H₇NO₂, 0.1 M BIS-TRIS pH 6.5, 25% w/v PEG 3,350) and SALT-RX condition F6 (2.5 M (NH₄)₂SO₄, 0.1 M Tris pH 8.5), each from a different space group. Crystals from both space groups were soaked for 1 and 4 hours with 5 %, 10 %, 20 % and 40 % DMSO dispensed using an echo-acoustic liquid dispenser (Labcyte 550, USA) at room temperature. Control crystals were not submitted to the soaking procedure. Crystals from

Crystallization optimization

Manual optimization was performed by varying protein and PEG concentration for INDEX HT condition G7 (0.2 M C₂H₇NO₂, 0.1 M BIS-TRIS pH 6.5, 25 % w/v PEG 3,350). Optimization was performed in Swissci plates. The protein to reservoir solution volume was varied 1:1, 1:2 and 2:1 as mentioned above in different PEG 3350 concentration (18 %, 20 %, 22 %, 25 %, 27 % and 30 %).

Fragment library screening and data analysis

The purified protein was subjected to the fragment screening procedure (**Figure 1**). Crystals were reproduced in Swissci plates in a 1:1 protein to reservoir solution concentration as mentioned previously. The protein concentration employed was 4.6 mg/mL and the crystallization condition contained 0.2 M C₂H₇NO₂, 0.1 M Bis-Tris pH 6.5, 30 % w/v PEG 3350. Crystals were soaked using a total of 827 fragments derived from Maybridge fragment library and an ECHO liquid handler operated with TeXRank (30). The selected Maybridge fragments met the criteria of purity, molecular weight, removal of any inappropriate functionality, drug likeness and reproducibility. These fragments were also screened for its Soaking was performed in 30 % DMSO and 60 mM fragments in the final volume of the crystals drop and incubated for one hour at room temperature. Following incubation, soaked trystals were mounted and stored in liquid nitrogen. X-ray diffraction data was collected at beamline 104-1 of DLS and autoprocessed. Dimple and REFMAC were used for initial structure refinement (31) and AceDRG for generation of ligand coordinates and restraints (32).

used for fragment fitting and subsequent model (re-)building (34). Subsequent refinements were performed with PHENIX (35) and analyzed with coot. The final models were submitted to Protein Data Bank (PDB) and assigned with the codes listed in **Table 3**.

Figure 1. Flow chart for the optimized high-throughput fragment screening methodology adopted for PNP2 (Lab Xchem, I04-1, DLS, UK).

RESULTS

Crystal Selection

The protein PNP2 has crystallized in 33 different conditions out of 6 crystallization kits (**Table 1**). The extensive range of crystallization conditions provided a broad range of options of crystal optimization for fragment screening. 9 different conditions were selected for data collection based on the size of the crystals (0.50 μ m on average), and X-rays diffraction data was collected (**Table 1**). The resolution of the collected diffraction data was in the range of 1.34 Å to 2.51 Å. The crystals belonged to two space groups, 6 conditions were assigned to the space group P 2₁ 3 and 3 conditions were indexed as F4₁ 3 2.

Table 1. Crystallization conditions and collected X-rays diffraction resolution.

DMSO tolerance test

The fragments in the Maybridge library are provided dissolved in DMSO. Crystals were soaked in 4 different concentrations of DMSO for two time intervals to check the effect of DMSO on the integrity of the crystals. We selected two different space groups, which were P 2_1 3 for INDEX HT condition G7, and F 4_1 3 2 for SALT-RX condition F6 with resolutions of 1.34 Å and 2.18 Å, respectively (**Table 1**). Conditions leading to different space groups with highest resolution were selected to evaluate the DMSO effect on each type of crystal lattice. Crystals were reproduced in Swiss-CI plates for both conditions, soaked with increasing concentration of DMSO, and incubated either for 1 hour or 4 hours at room temperature. Two repeats were performed for each concentration of DMSO and each time interval (**Table 2**). Controls were kept with no DMSO to check diffraction consistency, evaluated basically by resolution limit. The diffraction pattern for the condition G7 with space group P 2₁ 3 has shown tolerance to all the concentrations of the DMSO irrespective of the time interval (**Table 2**). Crystals had started to disassemble and started to become more rounded than the original cubic shape at the 4-hour interval but the diffraction resolution was not affected. In fact, the diffraction resolution was more dependent on the size of the crystals irrespective of the DMSO concentration and time interval, as shown in Table 2 where 40 % DMSO with 100 μ m diffracted with 1.65 Å, while 15 μ m gave 3.95 Å diffraction (**Table 2**). In contrast, for condition F6 with space group F 4₁ 3 2 the crystals did not show any tolerance to DMSO even at the lowest 5 % DMSO concentration (**Table 2**) and were dissolved in all the DMSO concentrations in both time intervals. Therefore, no diffraction data available for DMSO soaked crystals with space group F 4₁ 3 2.

Table 2. Diffraction resolution for crystals in space group P 2_1 3 soaked in different DMSO concentrations and time intervals showing changes with respect to their crystal size.

Crystal size optimization

In the DMSO tolerance test, we have observed that P 2_1 3 space group crystals endured DMSO at all the concentrations tested. It was also observed that the diffraction resolution of the DMSO soaked crystals was size dependent. We have manually optimized PEG 3350 concentration for INDEX HT condition G7 (0.2 M C₂H₇NO₂, 0.1 M Bis-Tris pH 6.5, 25 % w/v PEG 3350) along with the concentration of protein in the crystallization drop. Crystals with size ranging from 50

 μ m - 120 μ m were obtained in 30 % PEG 3350 concentration. Protein concentration did not affect the crystal size. Crystals were only obtained in crystallization drops with a protein-toreservoir ratio of 1:1, while precipitation was observed in 1:2 and 2:1 ratios. The highest– resolution diffraction data was collected from crystals produced in PEG 3350 30 % concentration, which consistently produced larger crystals (50 μ m - 105 μ m) (**Table 3**). The resolution of the diffraction data was improved, in the range of 1.50 Å – 2.50 Å.

Table 3. Effects of different PEG 3350 concentrations on the crystals size.

Fragment library soaking and hit detection

It was observed that larger crystals generally diffracted to higher resolution. In the fragmentsoaking condition containing 30 % DMSO and 60 mM of the fragment, soaking was performed using 827 fragments; diffraction data was subsequently collected from the resulting crystals and autoprocessed, with a resolution range of 1.60 - 2.50 Å. Fragment hits were identified with PanDDA using the default settings and visually analyzed in COOT [33,34]. A total of 19 fragments were identified bound to PNP2 (**Table 4**).

Table 4. PDB IDs, chemical names and structures of each fragment observed bound to PNP2.

 The number of DMSO molecules built and resolution of each structure are also listed.

Fragment binding sites

Of the 19 observed binders, a total of 14 fragments were found in the active site, and the other five fragments were observed in three additional sites (**Figure 2**).

Three of the five non-active site binders were observed in one site, located between the helices α 5 and α 9, corresponding to the PDB IDs 6B3L, 6B71 and 6BI1. Met275 adopts a rotamer that exposes a predominantly hydrophobic pocket delimited by Ile179, Phe185, Leu272, Met275 and Trp276. Only two polar residues are nearby this site, Asn183 and Glu279, and of these Asn183 was observed to hydrogen bond the amine nitrogen of a fragment (PDB ID 6B3L) (**Figure 3**). The overall architecture of the site is conserved in PNP1, with the exception of Met275 (Ser275 in PNP1) and Glu279 (Lys 279 in PNP1). Additionally, Ile179 is mutated to Val179 in PNP1, but this change does not impact the region where the fragments were observed to bind in this site.

Figure 3. Fragment found in PDB ID 6B3L was observed at the surface of the first site of PNP2. The hydrophobic pocket that accommodates the fragment 5 is formed by residues Phe185, Leu272 and Met275, while a single hydrogen bond is observed between the amine of fragment (blue sphere) and the side-chain oxygen (red sphere) of Asn183.

The second site is also observed between the helices $\alpha 5$ and $\alpha 9$, but closer to the C-terminal of PNP2, and is formed by Pro285, Leu171, Leu175, Lys178 and Arg283 (PDB ID 6B7I). A salt bridge is observed between an acid group of the fragment and Arg283, while Lys178 adopts a different rotamer in relation to the other structures in order to hydrogen bond the acid group and a hydroxyl of the fragment (**Figure 4**). Although this site is shallow, good complementarity and the satisfaction of the hydrogen bonding potential in both the protein and the fragment

Figure 4. Fragment in PDB ID 6B7I was observed in the second site of PNP2. The acidic group of the fragment forms salt bridges with both Lys178 and Arg283. The side chains of non-polar interacting residues are depicted as spheres, while Lys178 and Arg283 are shown as sticks.

The third site is formed by the 15-residue long loop connecting the β -strands β 2 and β 3, and is delimited by Ile53, Lys60, Thr61 and Thr62 (PDB ID 6B56). These residues adopt slightly different positions when compared to the other PNP2 crystal structures discussed here. This region is very different in PNP2 and PNP1, as the loop comprising residues 60 to 62 adopts a different conformation in the crystal structures of each enzyme. A series of polar contacts is observed involving the acid group of the fragment, which hydrogen bonds both the side chain hydroxyl and the main chain nitrogen atom of Thr61, and also interacts with Lys60 and two tightly bound water molecules (**Figure 5**). Lys60 also interacts with an aromatic nitrogen atom of the fragment.

Figure 5. Fragment observed in PDB ID 6B56 is bound to the third site of PNP2. The acid group of the fragment hydrogen bonds both the side chain and main chain of Thr61, and is also close to the amine nitrogen of the side chain of Lys60. The side chains of the residues interacting with the fragment are labeled and represented as spheres, while water molecules are depicted as smaller red spheres.

The active site of PNP2 has been described in the literature (20).

Nitrogenous bases occupy a well-defined pocket, closer to helices $\alpha 6$ and $\alpha 8$, and are observed hydrogen bonding to residues Glu203, Asn245 and Ser247. Pentose sugar moieties are closer to the helix $\alpha 9$, and their carbon 3 hydroxyl group interacts with Tyr90, while the hydroxyl group of the carbon 5 is observed to hydrogen-bond Tyr202 and His259.

In the structures presented here, 14 fragments were observed bound to the active site, a buried pocket delimited by the main chain of Leu118, Ala119 and Gly120, and the side chains of Tyr202, Glu203, Met221, Ala244, Asn245 and Ser247. The interactions of the fragments in this site is coherent with the pattern observed in the previously obtained co-crystals with adenine, cytidine, hypoxanthine, cytosine and tubercidin bound (20). For example, all fragments in this site position a ring moiety in the same region occupied by adenine and most of them are hydrogen bonded to Glu203 (**Figure 6**). The 12 smallest fragments are positioned in the same place where nitrogenous bases like adenine and cytosine are only involved in polar contacts with the residues Glu203, Asn245 and Ser247, and some of them also interact with a water molecule. Five fragments are observed to hydrogen bond to Asn197, an interaction not observed for nitrogenous bases. When compared to PNP1, all residues in this region are conserved, with the exception of L118A and A244T.

Figure 6. PNP2 active site bound to adenine and three representative fragments.

(A) Adenine (PDB ID 5TBS). (B) Fragment in PDB ID 6BJ7 (C) Fragment in PDB ID 6AXA.
(D) Fragment in PDB ID 6BI9. These planar fragments observed in the active site lie in the same plane as adenine. Several fragments interact with both Asn245 and Glu203, mimicking the interaction of adenine (B). Other fragments interact with Asn197 (C and D).

Interestingly, only one fragment, PDB ID 6BJ6, occupies the sugar-binding site, and is observed to hydrogen bond to Ser35, Asn117, and the carboxyl oxygen of Leu118, in addition to two water molecules (**Figure 7**). In comparison to the other structures presented here, Ser35 and His259 are slightly displaced, and Asn117 adopts a different rotamer. The binding of the sulfinyl oxygen atom of fragment 18 to Asn117 and to the main chain of Leu118 causes a small perturbation in the β -strand β 5.

Figure 7. Fragment observed in PDB IF 6BJ6 bound in the active site of PNP2. The fragment shows hydrogen bonds to residues Ser35 and Asn117. The side chains of residues Met221, His259 and Val262 delimit this pocket, and all labeled residues are conserved between PNP1 and PNP2.

One of the hits of our fragment screening reveals several novel structural features: fragment 1 (PDB ID 6AWE) explores a previously unidentified cavity, adjacent to the nitrogenous base pocket, formed in the vicinity of helices $\alpha 6$ and $\alpha 8$. In PNP2, the residues between Ile248 and Asn253 form the helix $\alpha 8$, but in all deposited PNP1 structures these residues do not adopt a helical conformation. Several residues in this region are not conserved between PNP1 and PNP2, among these Ile257 (Pro257 in PNP1), which appears to have a prominent role in the binding of fragment 1, as it closes the pocket together with the conserved Tyr202. The conserved active site residues Tyr202 and Glu203 in the loop preceding helix $\alpha 6$, together with Asn197, delimit one side of the pocket, while the other side is formed by Asn245, Ser247, Asp250, Val255 and Ile257. Fragment 1 presents only one polar contact in this region, a hydrogen bond between its pyridine nitrogen and the side chain of Asn245. On its other extremity the nitrile group of fragment 1 acts as a strong hydrogen bond acceptor by interacting with Asn197 (**Figure 8**). Another favorable hydrogen bond is observed between the side chain

of Glu203 and the secondary amine of fragment 1, which is a good hydrogen bond donor since it is linked to a benzene ring substituted with two electron withdrawing groups, a fluorine atom and the nitrile group. The analysis of the interactions of this fragment provides useful insights into the development of new inhibitors, as it is well buried into the nitrogenous-base binding site, but at the same time performs an additional interaction with Asn197 and places eight atoms, including a six-membered aromatic ring, in a previously uncharacterized pocket. Since this pocket is not conserved between PNP1 and PNP2, its exploration appears to be very promising for the development of potent and selective inhibitors for PNP2.

Figure 8. Fragment in PDB ID 6AWE in the active site of PNP2. Fragment sits over β -strands β 5 and β 8 in a way similar to adenine, but positions a pyridine in a previously uncharacterized pocket, where its nitrogen hydrogen bonds the side chain of Asn245. Fragment also interacts with two key residues, its nitrile group hydrogen bonds Asn197 and its secondary amine group is bound to Glu203. Asp250 and Ile257 also delimit this pocket.

DISCUSSION

Schistosoma mansoni is a common parasitic infection in the tropical and subtropical region; it is a chronic and advanced disease that causes abdominal pain, diarrhoea, blood in stools, liver cirrhosis, portal hypertension and premature death (36). Praziquantel and oxamniquine are the only drugs used to treat schistosomiasis in Africa and America. Coelho *et al* have reported a natural strain of *S. mansoni* partially resistant to oxamniquine (Mansi, Pfizer)(37). Praziquantel, a pyrazinosioquinoline derivative with low toxicity, has proved to be effective against *S. mansoni* and other species of the genus *Schistosoma* that parasitize humans. The exact mechanism of the praziquantel is still not elucidated (38). Praziquantel is a WHO-approved drug, but a major limitation to schistosomiasis control has been the limited availability of

praziquantel. According to 2013 reports, only 13.1 % of people requiring treatment were reached (3). Therefore, there is an emerging need for more effective drugs for schistosomiasis. There are no reports of a structure-based fragment screening for S. mansoni, and this approach is effective in the identification of hits which can be developed into lead compounds. Here we report the first high-throughput fragment screening process for S. mansoni. The fragment screening strategy that we describe is the robust technology available at Diamond Light Source, UK, which demands a conveniently short amount of time (typically less than 30 days) to screen a large library of fragments against a protein (Figure 2). The methodology employs several critical processing techniques that facilitate the otherwise prohibitive experimental and analytical workload. The most important of these is the use of the ECHO liquid handler, targeted with TeXRank, to perform the soaking procedure; this greatly simplifies sample handling, reducing the time required to perform the experiment, but also increasing the accuracy of the soaking. The autoprocessing pipeline for data collection and structure solution at DLS I04-1, including the identification of fragment hits with PanDDA(33), is another backbone of the fragment-screening methodology. The processing time of data collection and analysis is reduced to days rather than months; in our case, we were able to identify structurebound fragments for all 827 crystallographic datasets in two days.

The screening process is dependent on the tolerance of the protein crystals to DMSO, as all the fragments are dissolved in DMSO. Our crystals showed tolerance to all the different concentrations of DMSO that were tested. Moreover, we observed that the resolution of diffraction was dependent on the crystal size: small crystals have low diffraction resolution compared to larger crystals. Accordingly, larger crystals were obtained by using a higher concentration of PEG 3350, which facilitated the screening process. Thus, in our experience, high DMSO tolerance and large crystals were instrumental to the success of fragment screening using the high-throughput screening pipeline at I04-1, DLS.

As a result, we report 19 novel fragments for PNP2 as binding hits that may be developed into drug leads against S. mansoni. We have identified not only fragments that bind in the competitive site of the protein (14 fragments) but also three additional non-competitive sites (5 fragments) at the surface of PNP2. In the active site of PNP2, we have discovered a binding pocket that was previously not described (6AWE) for PNP2. In addition to engaging the adenine site, fragment 1 positions its pyridine group in this novel site, indicating that this nonconserved pocket might be occupied by relatively large groups. Such new information may be valuable for designing more selective and effective inhibitors for the protein. Two fragments observed in the protein (6AWE and 6BIF) are poised fragments, which can be utilized for straightforward expansion through synthesis and target the active site. Amongst the observed fragments in the active site, only fragment 18 (6BJ6) occupies the binding site of sugar moiety and phosphate as observed for cytidine (5TBT). Thus, fragments 1 and 18 are amongst the most important fragments that can block the whole active site. These fragments bring rich structural information that can be explored for further drug design. It remains to be studied if the noncompetitive sites that were identified via this screening present any effect on the trimerization of PNPs and its activity.

Now identified, these fragment hits can be utilized in a follow up chemical synthesis to design new potentially-active leads against *S. mansoni*. Additionally, the wealth of structural data presented here can be put to use in similarity searches for already available and potentially inexpensive compounds that can be readily tested in further crystallization assays. The new drug target PNP2, which is an intermediate in the *S. mansoni* purine salvage pathway, is an appealing target as it can be inhibited in both genders and at all the stages of the parasite's lifecycle. *Schistosoma* infection is not only restricted to snails and humans but also other animals (39). Our screened fragments may also be utilized in identifying molecules for drug repurposing against *Schistosoma* in humans and animals.

Data Availability Statement

The atomic coordinates and related experimental data have been deposited at the Protein Data Bank (PDB) and released with the following IDs: 6B56, 6B7I, 6B3L, 6B71, 6B11, 6AWE, 6AXA, 6B2L, 6B37, 6B4Q, 6B4T, 6B6K, 6BB7, 6BFV, 6BHB, 6BI9, 6BIF, 6BJ6, 6BJ7. For direct access to the files use: www.rcsb.org/structure/ID. For example, for structure 6B56, the link is <u>https://www.rcsb.org/structure/6B56</u>.

Acknowledgements

We would like to acknowledge FAPDF (193.000.482/2011), CNPq (309054/2014-1 and 564007/2010-2), CAPES and Diamond Light Source for providing financial support to perform this study.

Conflict of interest

There is no conflict of interest amongst the authors.

PDB ID codes - the atomic coordinates and experimental data have been released.

6B56, 6B7I, 6B3L, 6B71, 6BI1, 6AWE, 6AXA, 6B2L, 6B37, 6B4Q, 6B4T, 6B6K, 6BB7, 6BFV, 6BHB, 6BI9, 6BIF, 6BJ6, 6BJ7

References

1. Gray DJ, Ross AG, Li Y-S, McManus DP. Diagnosis and management of schistosomiasis. BMJ. 2011 May 17;342:d2651.

2. Xiao S, Tanner M, N'Goran EK, Utzinger J, Chollet J, Bergquist R, et al. Recent investigations of artemether, a novel agent for the prevention of schistosomiasis japonica, mansoni and haematobia. Acta Trop. 2002 May;82(2):175–81.

3. W H O. Schistosomiasis [Internet]. World Health Organization. 2020 [cited 2020 Jun 29]. Available from: http://www.who.int/mediacentre/factsheets/fs115/en/.

4. Franco GR, Valadão AF, Azevedo V, Rabelo EM. The Schistosoma gene discovery program: state of the art. Int J Parasitol. 2000 Apr 10;30(4):453–63.

5. Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA, et al. A systematically improved high quality genome and transcriptome of the human blood fluke Schistosoma mansoni. PLoS Negl Trop Dis. 2012 Jan 10;6(1):e1455.

6. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, et al. The genome of the blood fluke Schistosoma mansoni. Nature. 2009 Jul 16;460(7253):352–8.

7. Zerlotini A, Aguiar ERGR, Yu F, Xu H, Li Y, Young ND, et al. SchistoDB: an updated genome resource for the three key schistosomes of humans. Nucleic Acids Res. 2013 Jan;41(Database issue):D728-31.

8. Han Z-G, Brindley PJ, Wang S-Y, Chen Z. Schistosoma genomics: new perspectives on schistosome biology and host-parasite interaction. Annu Rev Genomics Hum Genet. 2009;10:211–40.

9. DeMarco R, Verjovski-Almeida S. Schistosomes--proteomics studies for potential novel vaccines and drug targets. Drug Discov Today. 2009 May;14(9–10):472–8.

10. Pereira HM, Cleasby A, Pena S SDJ, Franco G GR, Garratt RC. Cloning, expression and preliminary crystallographic studies of the potential drug target purine nucleoside phosphorylase from Schistosoma mansoni. Acta Crystallogr Sect D, Biol Crystallogr. 2003 Jun;59(Pt 6):1096–9.

11. Dovey HF, McKerrow JH, Wang CC. Purine salvage in Schistosoma mansoni schistosomules. Mol Biochem Parasitol. 1984 Apr;11:157–67.

12. Senft AW, Miech RP, Brown PR, Senft DG. Purine metabolism in Schistosoma mansoni. Int J Parasitol. 1972 Jun;2(2):249–60.

13. Carson DA, Chang KP. Phosphorylation and anti-leishmanial activity of formycin B. Biochem Biophys Res Commun. 1981 Jun 16;100(3):1377–83.

14. Nelson DJ, LaFon SW, Tuttle JV, Miller WH, Miller RL, Krenitsky TA, et al. Allopurinol ribonucleoside as an antileishmanial agent. Biological effects, metabolism, and enzymatic phosphorylation. J Biol Chem. 1979 Nov 25;254(22):11544–9.

15. Berg M, Van der Veken P, Goeminne A, Haemers A, Augustyns K. Inhibitors of the purine salvage pathway: A valuable approach for antiprotozoal chemotherapy? Curr Med Chem. 2010;17(23):2456–81.

16. Pereira TC, Pascoal VDB, Marchesini RB, Maia IG, de Almeida Magalhães L, Zanotti-Magalhães EM, et al. Schistosoma mansoni: evaluation of an RNAi-based treatment targeting HGPRTase gene. | Semantic Scholar. undefined. 2008;

17. Pereira HD, Franco GR, Cleasby A, Garratt RC. Structures for the potential drug target purine nucleoside phosphorylase from Schistosoma mansoni causal agent of schistosomiasis. J Mol Biol. 2005 Oct 28;353(3):584–99.

18. Pereira HM, Berdini V, Ferri MR, Cleasby A, Garratt RC. Crystal structure of Schistosoma purine nucleoside phosphorylase complexed with a novel monocyclic inhibitor. Acta Trop. 2010 May;114(2):97–102.

19. Castilho MS, Postigo MP, Pereira HM, Oliva G, Andricopulo AD. Structural basis for selective inhibition of purine nucleoside phosphorylase from Schistosoma mansoni: kinetic

and structural studies. Bioorg Med Chem. 2010 Feb 15;18(4):1421-7.

20. Torini JR, Romanello L, Batista FAH, Serrão VHB, Faheem M, Zeraik AE, et al. The molecular structure of Schistosoma mansoni PNP isoform 2 provides insights into the nucleoside selectivity of PNPs. PLoS One. 2018 Sep 7;13(9):e0203532.

21. Kumar A, Voet A, Zhang KYJ. Fragment based drug design: from experimental to computational approaches. Curr Med Chem. 2012;19(30):5128–47.

22. Baker M. Fragment-based lead discovery grows up. Nat Rev Drug Discov. 2013 Jan;12(1):5–7.

23. Jhoti H, Cleasby A, Verdonk M, Williams G. Fragment-based screening using X-ray crystallography and NMR spectroscopy. Curr Opin Chem Biol. 2007 Oct;11(5):485–93.

24. Tickle I, Sharff A, Vinkovic M, Yon J, Jhoti H. High-throughput protein crystallography and drug discovery. Chem Soc Rev. 2004 Oct 20;33(8):558–65.

25. Koellner G, Luić M, Shugar D, Saenger W, Bzowska A. Crystal structure of calf spleen purine nucleoside phosphorylase in a complex with hypoxanthine at 2.15 A resolution. J Mol Biol. 1997 Jan 17;265(2):202–16.

26. Bird LE. High throughput construction and small scale expression screening of multitag vectors in Escherichia coli. Methods. 2011 Sep;55(1):29–37.

27. Winter G, McAuley KE. Automated data collection for macromolecular crystallography. Methods. 2011 Sep;55(1):81–93.

28. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr Sect D, Biol Crystallogr. 2011 Apr;67(Pt 4):235–42.

29. Winter G, Waterman DG, Parkhurst JM, Brewster AS, Gildea RJ, Gerstel M, et al. DIALS: implementation and evaluation of a new integration package. Acta Crystallogr D Struct Biol. 2018 Feb 1;74(Pt 2):85–97.

30. Ng JT, Dekker C, Kroemer M, Osborne M, von Delft F. Using textons to rank crystallization droplets by the likely presence of crystals. Acta Crystallogr Sect D, Biol Crystallogr. 2014 Oct;70(Pt 10):2702–18.

31. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr Sect D, Biol Crystallogr. 2011 Apr;67(Pt 4):355–67.

32. Long F, Nicholls RA, Emsley P, Graéulis S, Merkys A, Vaitkus A, et al. Validation and extraction of molecular-geometry information from small-molecule databases. Acta Crystallogr D Struct Biol. 2017 Feb 1;73(Pt 2):103–11.

33. Pearce NM, Krojer T, Bradley AR, Collins P, Nowak RP, Talon R, et al. A multicrystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. Nat Commun. 2017 Apr 24;8:15123.

34. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr Sect D, Biol Crystallogr. 2010 Apr;66(Pt 4):486–501.

35. Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr Sect D, Biol Crystallogr. 2010 Feb;66(Pt 2):213–21.

36. Danso-Appiah A, Olliaro PL, Donegan S, Sinclair D, Utzinger J. Drugs for treating Schistosoma mansoni infection. Cochrane Database Syst Rev. 2013 Feb 28;(2):CD000528.

37. Coelho PM, Lima e Silva FC, Nogueira-Machado JA. Resistance to oxamniquine of a Schistosoma mansoni strain isolated from patient submitted to repeated treatments. Rev Inst Med Trop Sao Paulo. 1997 Apr;39(2):101–6.

38. Doenhoff MJ, Cioli D, Utzinger J. Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. Curr Opin Infect Dis. 2008 Dec;21(6):659–67.

39. Webster BL, Southgate VR, Littlewood DTJ. A revision of the interrelationships of

Schistosoma including the recently described Schistosoma guineensis. Int J Parasitol. 2006 Jul;36(8):947–55.



Figure 1. Flow chart for the optimized high-throughput fragment screening methodology

adopted for PNP2 (Lab Xchem, I04-1, DLS, UK).



Figure 2: Three-dimensional structure of PNP2 with observed fragments binding sites. Red circles represent fragment binding sites of PNP2 along with PDB ID's in which fragments were observed. The active site had the maximum number of 14 fragments bound.



Figure 3. Fragment found in PDB ID 6B3L was observed at the surface of the first site of PNP2. The hydrophobic pocket that accommodates the fragment 5 is formed by residues Phe185, Leu272 and Met275, while a single hydrogen bond is observed between the amine of fragment (blue sphere) and the side-chain oxygen (red sphere) of Asn183.



Figure 4. Fragment in PDB ID 6B7I was observed in the second site of PNP2. The acidic group of the fragment forms salt bridges with both Lys178 and Arg283. The side chains of non-polar interacting residues are depicted as spheres, while Lys178 and Arg283 are shown as sticks.



Figure 5. Fragment observed in PDB ID 6B56 is bound to the third site of PNP2. The acid group of the fragment hydrogen bonds both the side chain and main chain of Thr61, and is also close to the amine nitrogen of the side chain of Lys60. The side chains of the residues interacting with the fragment are labeled and represented as spheres, while water molecules are depicted as smaller red spheres.



Figure 6. PNP2 active site bound to adenine and three representative fragments.

(A) Adenine (PDB ID 5TBS). (B) Fragment in PDB ID 6BJ7 (C) Fragment in PDB ID 6AXA.
(D) Fragment in PDB ID 6BI9. These planar fragments observed in the active site lie in the same plane as adenine. Several fragments interact with both Asn245 and Glu203, mimicking the interaction of adenine (B). Other fragments interact with Asn197 (C and D).



Figure 7. Fragment observed in PDB IF 6BJ6 bound in the active site of PNP2. The fragment shows hydrogen bonds to residues Ser35 and Asn117. The side chains of residues Met221, His259 and Val262 delimit this pocket, and all labeled residues are conserved between PNP1 and PNP2.



Figure 8. Fragment in PDB ID 6AWE in the active site of PNP2. Fragment 1 sits over β -strands β 5 and β 8 in a way similar to adenine, but positions a pyridine in a previously uncharacterized pocket, where its nitrogen hydrogen bonds the side chain of Asn245. Fragment 1 also interacts with two key residues, its nitrile group hydrogen bonds Asn197 and its secondary amine group is bound to Glu203. Asp250 and Ile257 also delimit this pocket.

Crystallization kits	Crystallization conditions	Crystals diffracted	Space groups	Diffraction resolution (Å)
SALT-RX	E6, F5, F6	F6	F 41 3 2	2.18
		E6	F 41 3 2	2.39
INDEX HT	D6, D11, F9, F11, G6, G7, G10	G7	P 2 ₁ 3	1.34
		F11	P 21 3	2.13
		C5	F 41 3 2	2.26
PEG/Ion	A4, A5, B5, B11, B12, C6, C11	B12	P 2 ₁ 3	2.01
		C11	P 2 ₁ 3	2.12
PACT premier	B3, C3, D3	C3	P 2 ₁ 3	2.15
Morpheus	A1, B1, D1, E1	-	-	-
JCSG-Plus	A8, B7, B9, C1, C3, D4, H3, H8, H10	H10	P 2 ₁ 3	2.51

 Table 1. Crystallization conditions and collected X-rays diffraction resolution.

DMSO concentration (%)	Soaking time (h)	Crystals size (μm)	Diffraction resolution (Å)
Control	1	30	2.84
Control	1	25	2.59
Control	4	120	1.80
Control	4	25	2.90
5	1	25	2.28
5	1	50	2.28
5	4	30	2.9
	4	-	No diffraction
10	1	50	1.91
	1	60	1.69
10	4	-	No diffraction
	4	100	2.17
20	1	30	4.05
	1	50	2.07
20	4	50	2.02
	4	80	2.11
40	1	50	4.07
	1	100	1.65
40	4	60	1.94
	4	15	3.95

Table 2. Diffraction resolution for crystals in space group P 2_1 3 soaked in different DMSO concentrations and time intervals showing changes with respect to their crystal size.

PEG 3350	Crystals	Crystal size	
concentration		(µm)	
(%)			
18	× × _{101microns}	27	
20	×-×101microns	36	
22	↔ ×→×101microns	46	
25	× ×101microns	64	
27	× × 101microns	72	
30	×-×101microns	100	

Table 3. Effects of different PEG 3350 concentrations on the crystals size.

Table 4. PDB IDs, chemical names and structures of each fragment observed bound to PNP2.

The number of DMSO molecules built and resolution of each structure are also listed.

List of screened fragments found in the structures								
PDB ID's	Chemical Name	Structures	# of DMSO	Resolution (Å)				
6AWE	2-fluoro-6-(2-pyridin-2- ylethylamino)benzonitril e	F N N	8	1.65				
6AXA	(3S)-5-fluoro-3-hydroxy- 1,3-dihydroindol-2-one		8	1.59				
6B2L	Piperidin-2-imine		7	1.56				
6B37	1,3-benzothiazol-2-ol	HO	6	1.50				
6B3L	2,3-dihydro-1H-inden-2- amine	H ₂ N	6	1.52				
6B4Q	Pyridin-4-ol	OH	6	1.60				





