Naegleria fowleri: protein structures to facilitate drug discovery for the deadly, pathogenic free-living amoeba

Kayleigh Barrett[1,8], Logan Tillery[1,8], Jenna Goldstein[†9], Jared W. Lassner[†9], Bram Osterhout[†9], Nathan L. Tran[†9], Lily Xu[†9], Ryan M. Young[†9], Justin Craig[1,8], Ian Chun[1,8], David M. Dranow[3,8], Jan Abendroth[3,8], Silvia L. Delker[3,8], Douglas R. Davies[3,8], Stephen J. Mayclin[3,8], Brandy Calhoun[3,8], Madison J. Bolejack[3,8], Bart Staker[2,8], Sandhya Subramanian[2,8], Isabelle Phan[2,8], Donald D. Lorimer[3,8], Peter J. Myler[2,8,10,11,12], Thomas E. Edwards[3,8], Dennis E. Kyle[4], Christopher A. Rice[4 Δ], James C. Morris[5], James W. Leahy[6], Roman Manetsch[7], Lynn K. Barrett[1,8], Craig L. Smith[9], Wesley C. Van Voorhis[1,8,12]

[1]Department of Medicine, Division Allergy and Infectious Disease, Center for Emerging and Re-emerging Infectious Disease (CERID) University of Washington, Seattle, WA, USA

[2]Seattle Children's Research Institute, Seattle, WA, USA

[3]UCB Pharma, Bainbridge Island, WA, USA

[4]Center for Tropical and Emerging Global Diseases, University of Georgia, Athens, GA, USA

[5]Eukaryotic Pathogens Innovation Center, Department of Genetics and Biochemistry, Clemson University, Clemson, SC, USA

[6]Department of Chemistry, University of South Florida, Tampa, FL, USA

[7]Department of Chemistry and Chemical Biology and Department of Pharmaceutical Sciences, Northeastern University, Boston, MA, USA

[8]Seattle Structural Genomics Center for Infectious Diseases, Seattle, WA, USA

[9]Department of Biology, Washington University, St. Louis, MO, USA

[10]Department of Pediatrics, University of Washington, Seattle, WA USA

[11]Department of Biomedical Informatics & Medical Education, University of Washington, Seattle, WA, USA

[12]Department of Global Health (Pathobiology), University of Washington, Seattle, WA, USA

*Corresponding author

E-mail: wesley@u.washington.edu

[†] These authors contributed equally to this work

[Δ] Current address: Department of Pharmaceutical and Biomedical Sciences, College of Pharmacy, University of Georgia, Athens, Georgia, USA.

Short title; *Naegleria fowleri*: protein structures to facilitate drug discovery

ABSTRACT: *Naegleria fowleri* is a pathogenic, thermophilic, free-living amoeba which causes primary amebic meningoencephalitis (PAM). Penetrating the olfactory mucosa, the brain-eating amoeba travels along the olfactory nerves, burrowing through the cribriform plate to its destination: the brain's frontal lobes. The amoeba thrives in warm, freshwater environments, with peak infection rates in the summer months and has a mortality rate of approximately 97%. A major contributor to the pathogen's high mortality is the lack of sensitivity of *N. fowleri* to current drug therapies, even in the face of combination-drug therapy. To enable rational drug discovery and design efforts we have pursued protein production and crystallography-based structure determination efforts for likely drug targets from *N. fowleri*. *N. fowleri* genes were selected if they had homology to drug targets listed in Drug Bank or were nominated by primary investigators engaged in *N. fowleri* research. In 2017, 178 *N. fowleri* protein targets were queued to the Seattle Structural Genomics Center of Infectious Disease (SSGCID) pipeline, and to date 89 soluble recombinant proteins and 19 unique target structures have been produced. Many of the new protein structures are potential drug targets and contain structural differences compared to their human homologs, which could allow for the development of pathogen-specific inhibitors. Five of the structures were analyzed in more detail, and four of five show promise that selective inhibitors of the active site could be found. The 19 solved crystal structures build a foundation for future work in combating this devastating disease by encouraging further investigation to stimulate drug discovery for this neglected pathogen.

1. INTRODUCTION

In Australia in 1965, Fowler and Carter reported the first case of *Naegleria fowleri* infection, commonly referred to as the "brain-eating amoeba", which is the only known species of the *Naegleria* genus associated with human disease (1). The free-living amoeba is found in soil and freshwater on all seven continents, but is mainly found in warmer regions flourishing in freshwater and soils with higher temperatures up to 115°F (46°C) (2).

Infection rates of the amoeba increase during summer months causing the disease, primary amebic meningoencephalitis (PAM) (2,3). PAM is a fatal CNS disease that displays severe meningitis and cranial pressure caused by inflammation of the brain (3,4). The National Institute of Allergy and Infectious Diseases (NIAID) has classified *N. fowleri* as a category B priority pathogen, the second highest class of priority organisms/biological agents. Category B pathogens typically have high mortality rates, are easily disseminated, may cause public panic and social disruption, and require special action for public health preparedness (2,5).

Between the years 1962 and 2016, 145 cases were reported by the CDC within the USA, and only 4 patients (3%) survived *N. fowleri* infection and subsequent PAM (6). This 54-year reporting tally might suggest only several PAM cases occur in the USA a year. However, a recent review found that hundreds of undiagnosed cases of fatal "meningitis and encephalitis" were reported in the summers between the years 1999 and 2010 in persons aged 2-22 years in the Southeast USA (7). Cases lacking diagnostic brain biopsy or spinal fluid analysis of PAM may account for a portion of the undiagnosed, inconclusive cases (7). Thus, the incidence of *N. fowleri* infections in the USA is probably much higher than several a year arguing, that *N.*

fowleri poses a much higher public health threat than the documented cases would suggest (7). Rising global temperatures has also led to claims of increased risk of infections due to spread of suitable water conditions for *N. fowleri* (3,8).

The global burden of *N. fowleri* PAM cases is likely greater than anticipated, yet undefined. Aga Khan University in Karachi, Pakistan has reported a rise in the number of PAM cases. Aga Khan Hospital serves only a small fraction of the Karachi population (~23 million total) and reports around 20 cases/year of PAM. Other hospitals in Karachi did not report a single case during the same time, likely due to the lack of awareness, autopsies, and microscopy (9). Only about 10 cases of *N. fowleri* PAM have been reported in Africa, though the numbers are likely impacted by a reporting bias, as the resources for diagnostic brain biopsy are rarely present, and autopsies are almost never performed (10).

The combination of a devastating mortality rate, a warming climate, and a rapid-onset infection emphasize why *N. fowleri* should not stay a neglected organism. Treatment is extremely limited and not well defined, using a combination of known antifungals, antibiotics and microbicides (11). Amphotericin B is the current drug of choice, in high dosages, in combination with other repurposed drugs such as rifampin, miltefosine, and fluconazole (9,12). Recent studies suggest that posaconazole is more efficacious than fluconazole in vitro and in animal models of PAM (12). Miltefosine was used in combination to successfully treat two patients, one reported and one unreported in the literature. Miltefosine in combination is not always helpful, in that a patient was treated with miltefosine and suffered permanent brain damage and another had a fatal outcome (11,13). The multi-drug therapy is associated with

severe adverse effects and requires higher than normal dosages to penetrate the blood-brain barrier and to reach the CNS (11,14). New development of rapid-onset, brain permeable, efficient, and safe drugs is urgently needed.

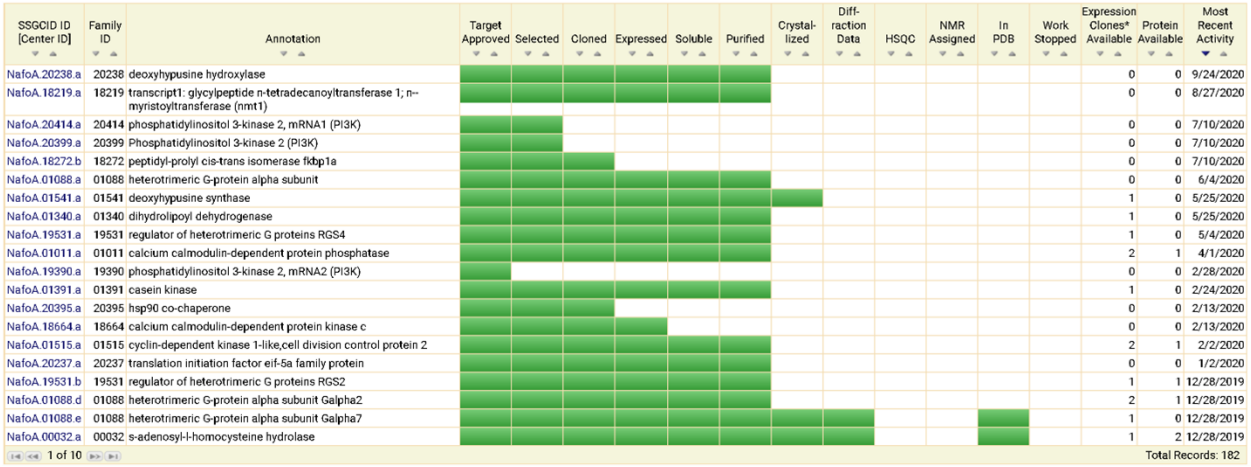
Given the lack of understanding about causes of drug resistance of *N. fowleri* and the urgent need for new drugs, we investigated the proteome of *N. fowleri* for likely drug targets attempting to enable further drug discovery efforts by producing material for characterization of the proteins. This work is a first step towards the discovery of drugs specifically designed against *N. fowleri*.

2. RESULTS

2.1 The *N. fowleri* proteome contains hundreds of potential drug targets.

Potential drug targets were selected by sequence homology to DrugBank protein targets (15). Additional targets were requested by the amoeba research community, leading to a total of 178 *N. fowleri* targets entering the Seattle Structural Genomics for Infectious Disease (SSGCID) structure determination pipeline. The SSGCID is a National Institutes for Allergy and Infectious Disease (NIAID) supported preclinical service for external investigators (www.SSGCID.org) (16). All targets were filtered according to the standard SSGCID target selection protocol and criteria (13): eliminating proteins with over 750 amino acids, 10 or more cysteines, or 95% sequence identity with 70% coverage to proteins already in the PDB, targets claimed or worked on by other scientific groups, and targets with transmembrane domains (except where a soluble domain could be expressed separately) (18). Target criteria resulted in selection of 178 proteins which entered the SSGCID production pipeline. These proteins, homologous to other known

drug targets, consisted of metabolic enzymes, protein synthetases, kinases, and others. Real-time updates to target status progress can be viewed at the SSGCID website. **Figure 1** shows a view of current status. Additionally, a detailed table of the *Naegleria* protein crystallography statistics is available in Supplemental Information.



SSGCID ID [Center ID]	Family ID	Annotation	Target Approved	Selected	Cloned	Expressed	Soluble	Purified	Crystal-lized	Diff-raction Data	HSQC	NMR Assigned	In PDB	Work Stopped	Expression Clones* Available	Protein Available	Most Recent Activity
NafoA.20238.a	20238	deoxyhypusine hydroxylase	■	■	■	■	■	■							0	0	9/24/2020
NafoA.18219.a	18219	transcript1: glycylpeptide n-tetradecanoyltransferase 1; n-myristoyltransferase (nmt1)	■	■	■	■	■	■							0	0	8/27/2020
NafoA.20414.a	20414	phosphatidylinositol 3-kinase 2, mRNA1 (PI3K)	■	■											0	0	7/10/2020
NafoA.20399.a	20399	Phosphatidylinositol 3-kinase 2 (PI3K)	■	■											0	0	7/10/2020
NafoA.18272.b	18272	peptidyl-prolyl cis-trans isomerase fkbp1a	■	■	■										0	0	7/10/2020
NafoA.01088.a	01088	heterotrimeric G-protein alpha subunit	■	■	■	■	■	■							0	0	6/4/2020
NafoA.01541.a	01541	deoxyhypusine synthase	■	■	■	■	■	■	■						1	0	5/25/2020
NafoA.01340.a	01340	dihydrolipoyl dehydrogenase	■	■	■	■	■	■							1	0	5/25/2020
NafoA.19531.a	19531	regulator of heterotrimeric G proteins RGS4	■	■	■	■	■	■							1	0	5/4/2020
NafoA.01011.a	01011	calcium calmodulin-dependent protein phosphatase	■	■	■	■	■	■							2	1	4/1/2020
NafoA.19390.a	19390	phosphatidylinositol 3-kinase 2, mRNA2 (PI3K)	■												0	0	2/28/2020
NafoA.01391.a	01391	casein kinase	■	■	■	■	■	■							1	0	2/24/2020
NafoA.20395.a	20395	hsp90 co-chaperone	■	■	■										0	0	2/13/2020
NafoA.18664.a	18664	calcium calmodulin-dependent protein kinase c	■	■	■	■									0	0	2/13/2020
NafoA.01515.a	01515	cyclin-dependent kinase 1-like,cell division control protein 2	■	■	■	■	■	■							2	1	2/2/2020
NafoA.20237.a	20237	translation initiation factor eif-5a family protein	■	■	■	■	■	■							0	0	1/2/2020
NafoA.19531.b	19531	regulator of heterotrimeric G proteins RGS2	■	■	■	■	■	■							1	1	12/28/2019
NafoA.01088.d	01088	heterotrimeric G-protein alpha subunit Galpha2	■	■	■	■	■	■							2	1	12/28/2019
NafoA.01088.e	01088	heterotrimeric G-protein alpha subunit Galpha7	■	■	■	■	■	■	■	■			■		1	0	12/28/2019
NafoA.00032.a	00032	s-adenosyl-l-homocysteine hydrolase	■	■	■	■	■	■	■	■			■		1	2	12/28/2019

1 of 10 Total Records: 182

Figure 1. Status of *Naegleria* targets in the SSGCID pipeline as of 10/1/2020. This view is publicly available at www.ssgcid.org through the Targets tab, sorting by genus. Targets are displayed by SSGCID identifier and annotation with current status shown by green bars, sorted here by most recent activity reported to the database. The number of expression clones is shown and indicates different forms of the target that are available to external users and can be ordered. Protein availability also indicates samples of protein available to the external community and represents at least one vial of concentrated protein of approximately 100 microliters ranging from 10-50 mg/mL of >95% purity material.

2.2 One third of targets attempted produced soluble protein.

The open reading frames of each target were obtained from AmoebaDB.org. Progression of the targets through the SSGCID protein production pipeline is shown in **Table 1**. Of the 178 NIAID approved targets, 177 were selected for cloning. One protein target was eliminated due to redundancy given its 100% identity match and 84% coverage to another protein target already in the SSGCID pipeline. Of the 177 targets attempted in PCR amplification using *N. fowleri* cDNA,

133 were successfully amplified and cloned into SSGCID expression vectors (75%) (17). In small-scale expression screening, 82 of the 133 successfully cloned targets (61.6%) demonstrated soluble expression with a N-terminal His6-tag vector (17). Of these 82 soluble proteins, 64 proteins were purified to >95% purity with yields ranging from 1.1 mg to 348 mg. These proteins are available, under request, at SSGCID.org.

Table 1. Target status of *N. fowleri* proteins within SSGCID pipeline.

Target approved	**178**
Selected	177
Cloned	133
Expressed	98
Soluble	82
Purified	64
Crystallized	29
Diffraction	27
Native diffraction data	19
In PDB (apo)	19

2.3 *N. fowleri* proteins resulted in a 11% structure determination success rate.

Following purification of high-purity preparations of protein, the targets were submitted for structure determination by X-ray crystallography. Of the 64 proteins produced, 26 crystallized (29%), and 20 diffracted, 19 of which met the SSGCID resolution quality criteria and were submitted to the Protein Data Bank (PDB). **Table 2** lists the 19 targets deposited in PDB that will be reported in this paper, including five structures with a unique ligand bound, for a total of 23 PDB deposits. Overall, this resulted in a structure determination success rate of 11%, which is comparatively higher than usual structural genomic pipeline rates reported by us or other structural genomic groups (19).

Table 2. Listing of *Naegleria fowleri* structures deposited by SSGCID in the PDB.

SSGCID ID	Annotation	AmoebaDB	PDB ID	Metabolic pathway
NafoA. 00005.a	Malate dehydrogenase 2 (mitochondrial)	BF0021050	6UM4	TCA cycle
NafoA. 01242.a	dUTP pyrophosphatase	NF0068730	5VJY, 6MJK	DNA replication
NafoA.00085.a	GDP-L-fucose synthetase	NF0109030	6AQY, 6AQZ	Glycolysis
NafoA.00085.b	GDP-mannose dehydratase	NF0016040	5UZH	Glycolysis
NafoA.00855.a	Glyceraldehyde-3-phosphate dehydrogenase	NF0055660	6NLX	Glycolysis
NafoA.19900.a	Glucokinase	NF0035880	6DA0	Glycolysis
NafoA.01088.e	Heterotrimeric G-protein alpha subunit Galpha7	NF0037180	6NE6	GPCR signaling
NafoA.00438.a	Nucleoside diphosphate kinase	NF0036070	5U2I	Purine metabolism
NafoA.18272.a	Peptidylprolyl isomerase	NF0084240_2	6B4P, 6MKE	Post-translational protein modification
NafoA.01013.a	Phosphoglycerate mutase	NF0048460	5VVE	Glycolysis
NafoA.18681.a	Prolyl-tRNA synthetase	NF0061610	6NAB, 6UYH	Amino acid response
NafoA.01523.a	Protein arginine n-methyltransferase	NF0022020	6CU3, 6CU5	Post-translational protein modification
NafoA.19251.a	Polyubiquitin with 3 ub domains	NF0029300_2	5VIX	Ubiquitination
NafoA.01385.a	Rab GDP dissociation inhibitor alpha	NF0049320	6C87	Signaling pathways
NafoA.00927.a	Ras-related c3 botulinum toxin substrate 1 isoform x2	NF0029600_1	5VCU	Signaling pathways
NafoA.00032.a	S-adenosyl-L-homocysteine hydrolase	NF0035640	5V96	Methylation metabolism
NafoA.01238.a	Serine-tRNA synthetase	NF0119870	6BLJ	Protein synthesis
NafoA.00964.a	Trafficking protein particle complex subunit 3	NF0066220	6AQ3	Protein trafficking
NafoA.00601.c	Ubiquitin-conjugating enzyme e2	NF0031160_1	5V0R, 6MJ9	Ubiquitination

2.4 Comparative analysis of *N. fowleri* and *Homo sapiens* enzyme active sites suggest some targets have promise for selective inhibition.

We have already published that glucokinase inhibitors can be obtained that are selective for *N. fowleri* vs. the human homolog, supporting glucokinase as a target for *N. fowleri* therapeutics (20). We analyzed the other structures determined, comparing the *N. fowleri* structure to human homolog structures, in order to determine opportunities for selective design of chemical inhibitors. Comparison of the *N. fowleri* determined structure to human structures available in the PDB was done by superimposition of the coordinate files (**Table 3**). With the exception of a pair of 96% identical *N. fowleri* and human ubiquitin-conjugating enzymes e2, all of the *N. fowleri* enzymes differed from human homologs by more than 38% (**Table 3**). We wanted to focus on the known ligand binding sites, to search for potential differences for inhibitors. A PDB search revealed that five of the 19 *N. fowleri* structures determined also had human

173 homolog structures determined which contained a known inhibitor of the human protein (**Table**
174 **3)**. We then manually inspected and compared the binding sites of these five proteins, described
175 below for each protein. Despite sequence similarities of the active sites, there were four cases
176 where a case for active site specificity could be made, supporting these proteins as targets for
177 therapeutics for *N. fowleri*.
178

179 **Table 3. Comparative analysis of *N. fowleri* and human structures deposited the PDB.**

Annotation	PDB ID	# Amino acids	Human PDB ID	Coverage (#aa)	% Coverage	% Identity of coverage	Inhibitor or ligand in human structure
malate dehydrogenase (mitochondrial) (MDH2)	6UM4	436	4WLF	221	51%	26%	-
dUTP pyrophosphatase	6MJK, 5VJY	147	2HQU_A	136	93%	63%	-
GDP-L-fucose synthetase	6AQY, 6AQZ	333	4E5Y_A	311	93%	52%	-
GDP-mannose-dehydratase	5UZH	380	6GPJ_A	343	90%	68%	-
glyceraldehyde-3-phosphate dehydrogenase	6NLX	333	6YNF_A	324	97%	65%	-
heterotrimeric G-protein alpha subunit Galpha7	6NE6	321	6K41_A	312	97%	41%	-
nucleoside diphosphate kinase	5U2I	151	1JXV_A	151	100%	63%	-
peptidylprolyl isomerase	6MKE, 6B4P	119	4DRO_A	113	95%	53%	FK506-AN
phosphoglycerate mutase	5VVE	250	5Y65_C	250	100%	61%	KH2
prolyl-tRNA synthetase	6NAB, 6UYH	535	5VAD_A	526	98%	54%	91Y
protein arginine n-methyltransferase	6CU3, 6CU5	328	6NT2_A	310	95%	53%	GSK3368715
polyubiquitin with 3 ub domains	5VIX	230	5H07_A	227	99%	96%	-
rab GDP dissociation inhibitor alpha	6C87	444	n/a				-
ras-related c3 botulinum toxin substrate 1 isoform x2	5VCU	200	1I4D_D	191	96%	55%	-
S-adenosyl-L-homocysteine hydrolase	5V96	472	1LI4_A	427	90%	57%	Neplanocin
Serine-tRNA ligase	6BLJ	477	4L87_A	458	96%	52%	-
trafficking protein particle complex subunit 3	6AQ3	187	2CFH_A	176	94%	47%	-
ubiquitin-conjugating enzyme e2	6MJ9, 5V0R	161	4ONM_A	126	78%	48%	-

180

181 ***N. fowleri* S-adenosyl-L-homocysteine hydrolase (*Nf*SAHH)** catalyzes the breakdown of
182 S-adenosyl-homocysteine (SAH) into adenosine and homocysteine. SAH is a byproduct of S-

adenosyl-L-methionine as a methyltransferase; the transfer of a methyl group to its respective cellular substrates such as DNA or rRNA, produces SAH (21). SAH hydrolases play a central role in methylation reactions required for growth and gene regulation, and inhibitors of SAH hydrolase are expected to be antimicrobial drugs, especially for eukaryotic parasites (21). Ribavirin is structurally similar to adenosine and has been proved to produce a time-dependent inactivation of human (*Hs*) SAHH and *Trypanosoma cruzi* (*Tc*) SAHH (22).

The *Nf*SAHH asymmetric unit contains a homo-tetramer (**Figure 2**). Although each chain contains an active site, structural analysis indicates that two chains must be present for the hydrolysis reaction to occur successfully. Each chain consists of three domains: a substrate-binding, a cofactor-binding, and a C-terminus domain (23). When substrates are not bound, the substrate-binding domain is located on the exterior, far from the meeting point of all four subunits of the asymmetric unit (24). The C-terminus domain is involved in both cofactor binding and protein oligomerization (23). In addition to the three main constituents, the structure contains two hinge regions that connect the substrate-binding and cofactor-binding domains. When substrates bind, the hinge region changes conformation, closing the cleft between the substrate-binding domain and the cofactor-binding domain of the respective chain (24). In the structure of *Nf*SAHH, all subunits exhibit a closed conformation.

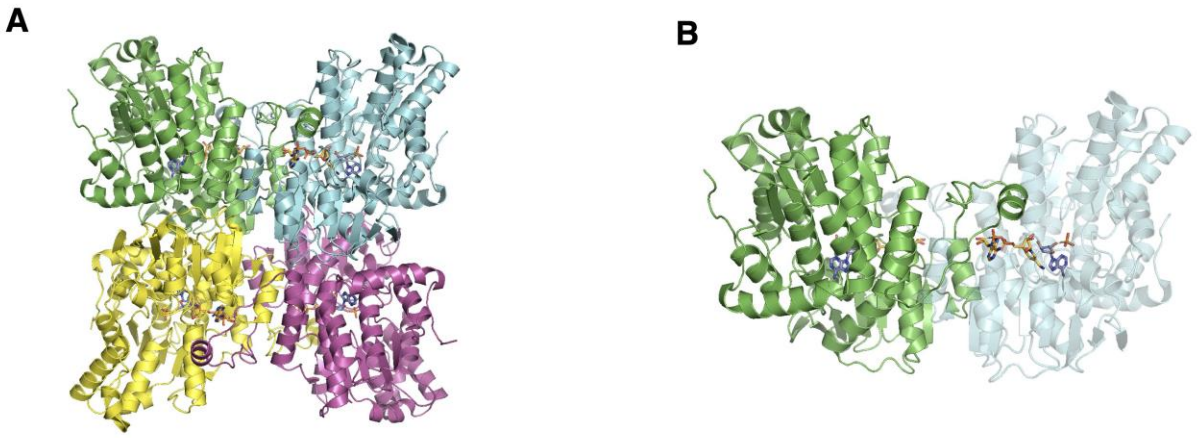


Figure 2. Crystal structure of *Nf*SAHH (PDB: 5V96) solved at a 2 Å resolution. (**A**) The asymmetric unit of SAHH. Individual polypeptide chains are colored green (chain A), pink (chain B), yellow (chain C), and cyan (chain D). (**B**) The biological unit is a homotetramer with a 2-fold axis of symmetry. Each of the four chains has its own active site containing one NAD^+ molecule (yellow), one adenosine molecule (purple), and a phosphate (orange).

SAHH is one of the most highly conserved proteins among species, with many of the same amino acids binding the same substrates across homologs. *Nf*SAHH is 62% identical to the human homolog. In the NAD binding region, conserved Lys and Tyr bind via hydrogen bonds to oxygens of NAD in both *Nf*SAHH and *Hs*SAHH (24) (**Figure 3**). Residues involved with binding adenosine (ADO) are also highly conserved, with Gly341-His342-Phe333 being completely conserved (24). To regulate the entrance of substrates into/out of the active site, there is a highly conserved His-Phe sequence within the cofactor-binding domain. This works as a molecular gate that, when the protein is in open conformation, allows access to the substrate-pocket (23).

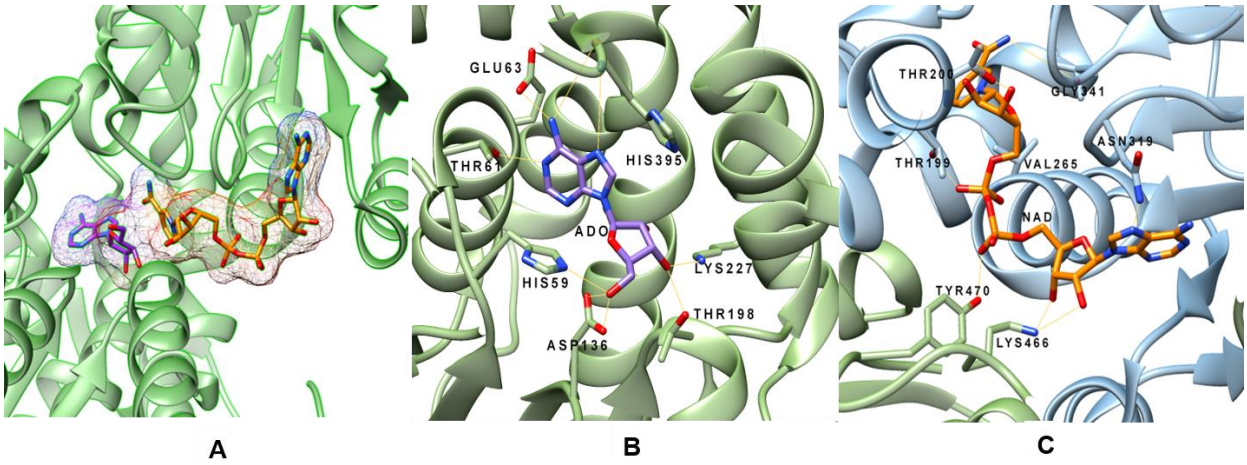


Figure 3. Overview of *Nf*SAHH active site binding. (**A**) Nicotinamide adenine dinucleotide (NAD/orange sticks) and adenosine (ADO/purple sticks) both fit well within the structure of SAHH. (**B**) Hydrogen-bond network denoted by yellow dashed lines around ADO involving chain A. (**C**) Hydrogen-bond network involving NAD with chain A (blue) and chain B (green).

A search of the PDB for *Hs*SAHH structures found multiple inhibitor bound human structures including 1LI4 (neplanocin) and 5W49 (oxadiazole compound). A comparison of the *Nf*SAHH structure to the neplanocin bound human structure revealed a highly conserved conformation of the protein. In the neplanocin bound structure, the two domains of the monomer are similar to the *Nf*SAHH structure. However, in the oxadiazole bound structure, two domains of the monomer are in a more open conformation, where the C-terminal and N-terminal domains have opened up relative to each other in a hinge-opening motion. The oxadiazole compound stretches across the interface and is surrounded by 11 residues within 4 Å. Of the 11 residues coordinating the inhibitor oxadiazole compound of the 5W49 structure, 10 are identical between the *Naegleria* and human SAHH. Only one change of M351T relative to the human enzyme is present, suggesting a highly conserved inhibitor binding site.

The crystal structure of *Nf*SAHH (**Figure 4**) contains the adenosine substrate and NAD cofactors bound to the active site to guide structure-activity relationships that could help to optimize adenosine analog compounds. The sequence differences that line the access channel at the dimer interface allow a rational approach to selectively inhibit the otherwise highly conserved active site (25). Amoeba SAHHs have an additional helix insertion that in *Nf*SAHH forms a hydrophobic groove accessible from the adenosine binding site (**Figure 4 B, C**). Specificity could be achieved by designing compounds that simultaneously target this hydrophobic pocket and the active site (**Figure 4**). Thus, we feel a reasonable case can be made that structural differences, close to the active site, would allow development of specific *Ng*SAHH inhibitors supporting development of a therapeutic.

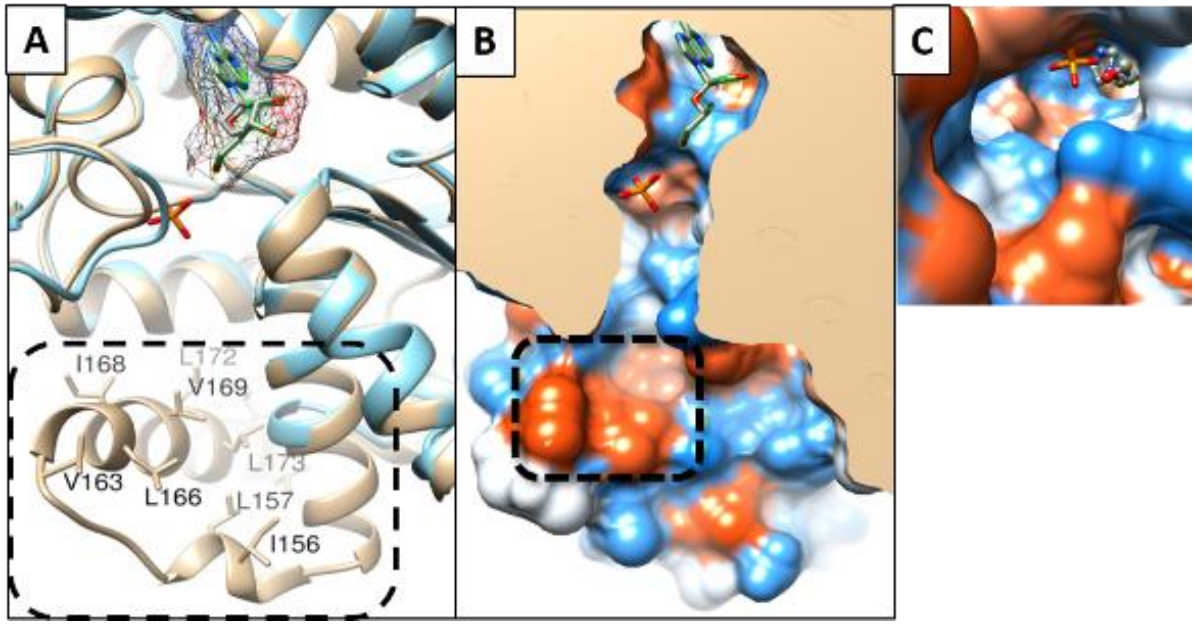


Figure 4. SAHH active site analysis. (**A**) Adenosine-bound *Nf*SAHH (PDB: 5V96, tan) vs. *Hs*SAHH bound to adenosine analogue (PDB: 1A7A, blue); box highlights *Nf*-specific insertion with labelled hydrophobic residues. (**B**) Slice through *Nf*SAHH, with surface colored on Kite-Doolittle (blue-red) hydrophobicity scale; dashed box

indicates hydrophobic groove formed by Val163, Leu166, Val169 and Leu173 at opening of deep-seated adenosine pocket. (**C**) Same view as (**B**) tilted 90° shows opening from hydrophobic groove to adenosine pocket.

***N. fowleri* phosphoglycerate mutase (*Nf*PGM)**, a glycolysis enzyme, catalyzes the isomerization of 3-phosphoglycerate and 2-phosphoglycerate during glycolysis and gluconeogenesis and is regarded as a key enzyme in most organism's central metabolism (26). There are two distinct forms of PGMs, differentiated by their need of 2, 3-bisphosphoglycerate as a cofactor. PGM in mammals require the cofactor whereas PGM present in nematodes and bacteria do not (27). The *Nf*PGM is likely the cofactor-dependent PGM type. The crystal structure of *Nf*PGM (PDB: 5VVE) was solved at a resolution of 1.7 Å and consists of 250 amino acid residues (~30 kDa).

The *Hs*PGM and *Nf* PGM structures share 61% identity. Residues surrounding the binding pocket for citrate acid are all conserved, with the exception of a conservative change from a Thr30 (*Nf*) to Ser30 (*Hs*). A comparison of the *Nf*PGM structure to the homologous human enzyme *Hs*PGM (PDB: 5Y65) shows a conformational opening of the substrate binding site to accommodate the KH2 ligand. However, the residues surrounding the inhibitor molecule and supporting the movement of the peptide are identical between the two enzymes. It is likely that with this high homologous identity that *Nf*PGM is not a strong candidate for selective active site inhibitor design.

***N. fowleri* protein arginine N-methyltransferase (*Nf*PRMT1)** methylates the nitrogen atoms found on guanidinium side chains of arginine residues within proteins. The methylation of nucleotide bases is a well-known mechanism of importance that influences DNA, nucleosomes, and transcription functionalities (28). The enzyme is highly conserved across eukaryotes. Faulty

regulation or deviating expression of PRMTs is associated with various diseases including inflammatory, virus-related, pulmonary, and carcinogenesis (29). Overexpression of PRMTs has been observed in multiple forms and types of cancer, including PRMT1v1 overexpression in colon cancer (30) and large increases of PRMT1v2 in breast cancer (31). Inhibitor discovery and testing using PRMTs in cancer has been frequently employed (29). *Nf*PRMT1 was compared to the drug bound structure of *Hs*PRMT1 (PDB: 6NT2). The protein binds ligands at a dimer interface closing around two inhibitor molecules, one on each monomer. A large ligand binding loop is disordered in the *Nf*PRMT1 structure, presumably becoming ordered and visible in the crystal structure in the presence of inhibitor in the human structure. Due to the large binding surface for peptide substrates, PRMTs typically are promiscuous in nature with a wide range of binding substrates (29). Comparison of over 40 PRMT-inhibitor complexes revealed 5 distinct binding mechanisms at multiple sites including active site and allosteric binding pockets (32). Isozyme specific peptide mimics have been identified which preferentially bind *Hs*PRMT1 vs. *Hs*PRMT5 enzyme. A similar approach could be considered for selective *Nf*PRMT inhibitor development (33,34). There is still a need to improve both the affinity and selectivity of these micromolar, sub-micromolar potent PRMT inhibitors as well as to better understand the enzyme's biological and disease processes in greater scope (35).

Despite high sequence identity in the ligand binding pocket, there are distinct differences in side chain orientation between the structures. These residues may change conformation upon binding inhibitor. A number of distinct features of *Nf*PRMT1 exist which can be exploited for potential structure-based approach to developing selective allosteric inhibitors against the *Nf* enzyme. A methionine is present in the *Nf*PRMT1 structure adjacent to the adenine moiety of the

S-adenosyl homocysteine (SAH) which differs significantly from all nine-known human PRMTs. The substrate binding region is lined by residues variant between *Nf* and all nine-known human PRMTs; for example, though the *Nf*PRMT1 pocket is similar to the allosteric inhibition pocket of *Hs*PRMT3, there are two tyrosine substitutions lining the pocket (36). Additionally, N-terminal residues which interact with inhibitors of *Hs*PRMT1 are largely not present or have limited interactions in *Nf*PRMT1 (37). Thus, inhibitors that selectively target *Nf*PRMT1 vs. the 9 *Hs*PRMTs are envisioned due to structural differences near the ligand binding sites.

***N. fowleri* peptidylprolyl isomerase (*Nf*PPI)** is a member of a superfamily of proteins comprised of three structurally distinct main families: cyclophilins, FK506 binding proteins (FKBPs), and parvulins. Based on structural and sequence alignment, the *N. fowleri* structure falls in the FKBP family, a group of enzymes inhibited by compounds such as FK506 and rapamycin (38). PPIs assist protein folding and influence protein denaturation kinetics by catalyzing the cis/trans isomerization of peptide bonds preceding prolyl residues (39). The enzymes participate in a diverse array of processes ranging from signal transduction to gene regulation and have been found to have close interaction with heat shock 90 proteins (40). PPI inhibitors are an emerging class of drugs for many therapeutic areas including infectious diseases and many potent small molecule inhibitors have been derived for each of the members of the superfamily. However, selective inhibitor design has been difficult due to the shallow, broad, solvent-exposed active sites and their conservation between homologs and protein families (41).

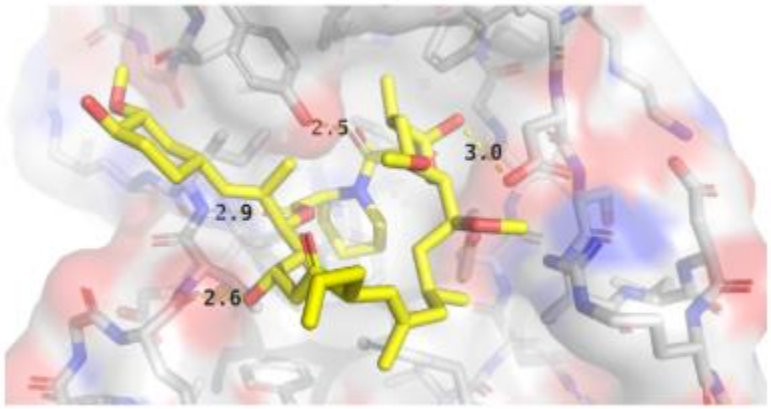


Figure 5. Interactions between *Nf*PPI and FK506. FK506 (yellow-carbon stick model) sits inside a mostly hydrophobic binding pocket of *NfPPI* (white-carbon stick and surface model) consisting of Tyr38, Phe58, Trp71, and Phe111 on the distal side of the substrate. Hydrogen-bonding interactions exist between FK506 and the side chains of Asp49 and Tyr94, and the backbones of Glu66 and Ile68.

The interior of the binding pocket of *Nf*PPI is mostly hydrophobic (**Figure 5**). Only four putative hydrogen-bonding interactions are observed between the enzyme and substrate. All residues involved in polar interactions in the *Nf*PPI are also present in the human homolog *Hs*FKBP51 (PDB: 1KT0), but the regions occupied by two hydrophilic residues in *Hs*FKBP51(Ser118 and Lys121) are instead occupied by hydrophobic residues, (Ile and Leu, respectively) (40). Another difference found in the conformation of this loop region is the insertion of an additional residue after Gly95 of *Nf*PPI. These changes in structure and sequence may lead to selective inhibition and thus establish PPIs as a selective drug target for *Naegleria*.

***N. fowleri* Prolyl-tRNA synthetase (*Nf*ProRS).** Aminoacyl-tRNA synthetases (ARSs) are globally essential enzymes among all living species. Their roles in protein translation and biosynthesis have been heavily researched and understood as attractive therapeutic targets. Recently, evidence of their propensity for adding new sequences or domains during ARS

evolution hints at broader functions and complexity outside of translation (42). Protein translation as a drug target has been validated for anti-infective compounds for a wide array of microbes (43). The natural product known as febrifugine, a quinazolinone alkaloid, and its analogues have shown antiparasitic activity in targeting ProRS. Halofuginone, a halogenated derivative of febrifugine, has shown promising potency though a lack of specificity, in that it inhibits both the parasite and human ProRS (43).

The structure of *Nf*ProRS folds into a α2 homodimer (**Figure 6A**) with each subunit containing three domains characteristic of Class II ARSs: the catalytic domain, the anticodon binding domain, and the editing domain (**Figure 6B**). The *Nf*ProRS catalytic domain features the three motifs which are exclusively conserved between class II ARSs for both sequence and structure-function (**Figure 6C**). Motif 1 is located at the interface of the dimer and is hypothesized to facilitate communication between the active sites of the two subunits (44). Motif 2 consists of β-strands connected by a variable loop which makes critical contacts with the acceptor stem of $tRNA^{Pro}$ and thus plays an important role in proper tRNA recognition (45). Motif 3 is made up of entirely hydrophobic residues and comprises an integral part of the aminoacylation active site.

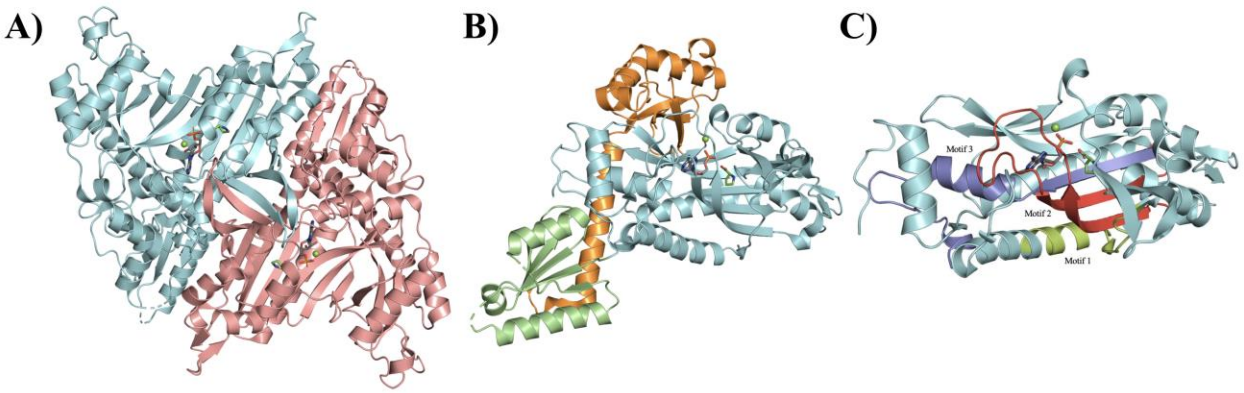
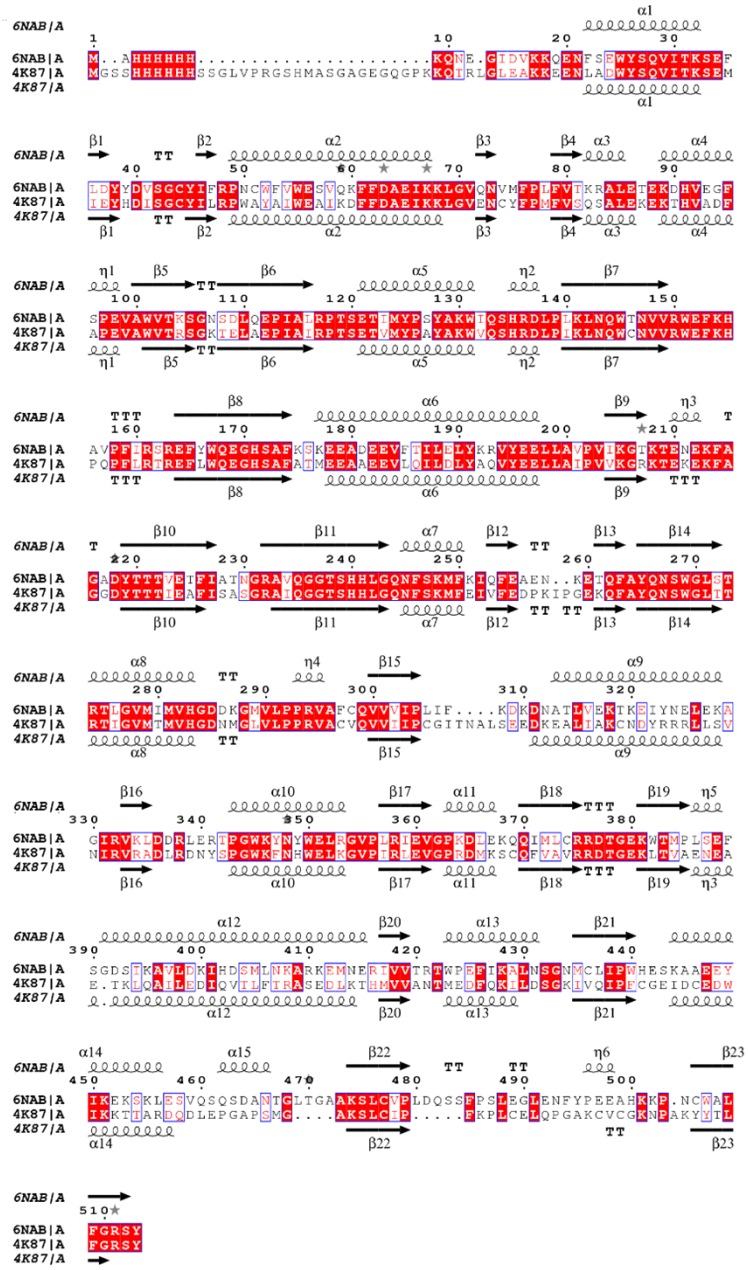


Figure 6. Structure of the *Nf*ProRS. (**A**) Both the biological and asymmetric unit of the structure are homodimeric. Individual polypeptide chains are shown in cyan and salmon. AMP, proline, and magnesium ligand molecules are also shown in yellow, purple, and green; respectively. (**B**) The three structure-function domains of *Nf*ProRS. The catalytic, anticodon binding, and editing domains are colored cyan, green, and orange; respectively. (**C**) The three highly conserved sequence motifs that characterize class II ARSs. Motif 1, colored lime, comprises the dimer interface. Motif 2, colored red, forms part of the acceptor stem. Motif 3, colored blue, is involved in forming the activated prolyl-adenylate.

Alignment of *Nf*ProRS bound to AMP and proline ligands (PDB: 6NAB) with apo *Hs*ProRS (PDB: 4K87) exhibits no significant structural changes between the apo and ligand forms of the ARS. The eukaryotic and archaeal origins of these ProRS make them suitable comparisons for the reason mentioned earlier: their strict conservation in all three structural domains. Both the proline and AMP bound *Nf*ProRS (PDB: 6NAB) and the halofuginone and AMP-PNP bound *Nf*ProRS (PDB: 6UYH) structures have been solved. The proline and AMP *Nf*ProRS (6NAB) shares structural homology with the halofuginone liganded ProRS (6UYH) and halofuginone binding induces a conformational change of residues 80-87 of the *N. fowleri* enzyme. In the proline bound 6NAB structure, residues 80-88 form a two-turn alpha helix (α4 in **Figure 7**). However, the halofuginone compound displaces Phe87 and disrupts the short helical structure. Residues making up this helix (EKDHVEGFS) are disordered in the 6UYH coordinate set. The equivalent region of the human ProRS (PDB: 4K87) is structurally homologous to the proline bound *N. fowleri* in the absence of halofuginone binding. The equivalent helix in 4K87, residues 90-98 (EKTHVADFA), includes non-conservative amino acid substitution adjacent to the crucial phenylalanine which must be displaced for halofuginone to bind the human enzyme, including E85-G86 which are A-D in the human sequence. Exploiting differences in the mobility of this non-conserved loop adjacent to the active site of *Nf*ProRS and *Hs*ProRS could enable selective

379 targeting. In addition, allosteric inhibitors that take advantage of sequence differences
380 throughout the *Nf*ProRS might be found by screening, as was the case for *P. falciparum* ProRS
381 (*Pf*ProRS) (46). Thus, ProRS may be a reasonable drug target for *N. fowleri* drug development.
382



383

384 **Figure 7. Alignment of *Nf*ProRS against its human homolog.** Red background denotes residue conservation
385 between the two structures, black text with white background signifies differences and red text encompassed in a

blue box demonstrates differences in residues that are not categorized as significant and the residues belong to the same grouping. Secondary structure annotations signaling helices and sheet are reflective of both structures respectively. Other structures used in alignment*: H. sapiens* (PDB: 4K87_A).

3. CONCLUSION

This manuscript reports 19 new protein structures from *N. fowleri* that are potential targets for structure-based drug discovery. Eighteen of 19 possess a >38% difference in AA alignment in comparison to the human homologs, suggesting selective inhibitors may be found by screening campaigns. In this paper we analyzed five of the *N. fowleri* enzymes that have ligands that define the active sites and compared them to human homologs. Though all are somewhat homologous at the active site, differences in four of the five *N. fowleri* enzymes analyzed support the hypothesis that selective active site inhibitors could be developed as therapeutics.

There are therapeutic opportunities, as well for some of the other 14 unexamined proteins as well. For example, the *Nf* serine tRNA synthetase (*Nf*SerRS) structure (PDB: 6BLJ). SerRS is required for charging tRNAs with serine critical for protein synthesis and thus is an essential gene. An insertion of four residues (391-395) adjacent to the substrate and tRNA binding sites creates a pocket with differential sequence identity to *Hs*SerRS and provides a foothold for the design of selective inhibitors blocking tRNA charging.

Even if selective active site inhibitors cannot be identified, high-throughput screening of compound libraries can still reveal selective inhibitors, as was found for *Plasmodium falciparum* ProRS compared with human *Hs*ProRS (46). In this case, two allosteric inhibitors were found to bury themselves into a lobe of the *Pf*ProRS enzyme, distant from the active site, and inhibit the

activity of the *Pf*ProRS enzyme, but not *Hs*ProRS. Selective high throughput screening of a eukaryotic enzyme including counter screening against the homologous human enzyme, can also identify selective inhibitors as has been shown by us in the case of *Plasmodium* N-myristoyltransferase (32). Focusing on essential genes and drug targets of other eukaryotes and producing a pool of potential drug target structures, SSGCID has created a foundation on which to build structure-based drug discovery. The relatively quick successful progress through the pipeline has catalyzed a consortium of investigators interested in addressing *N. fowleri* drug discovery.

4. MATERIALS AND METHODS

4.1 Bioinformatics

The complete genome and transcriptome is available on the EupathDB BRC website (www.amoebadb.org) (2). The complete ORFs and annotated predicted proteome from *Naegleria fowleri* strain ATCC30863 was downloaded from AmoebaDB release 24. Analysis of the ORFs indicated that 39% were missing a start codon and 12% were missing a stop codon. The sequence authors, the Wittwer group at the Spiez Laboratory, confirmed that the 40% of transcripts without an AUG start codon were most likely due to the ORF finder they used, which searches for the longest ORFs in the RNA assembly, but has no start codon finding function. To address this issue, we applied a conservative strategy to select high quality sequences from the draft genome. A sequence homology search using BlastP against DrugBank v.4.3 targets (4,212 sequences) (15) and potential drug targets in the SSGCID pipeline (9,783 sequences) was performed. Sequences with at least 50% amino acid sequence identity over 70% coverage were selected for further filtering. Manual inspection indicated that half the potential targets without a

start codon appeared to be significantly truncated when compared to the *Naegleria gruberi* and other closely related Eukaryota orthologues. Therefore, additional filters were applied to remove likely truncated sequences: (1) targets without a start or stop codon were discarded, (2) remaining candidates were blasted against the *Naegleria gruberi* proteome and sequences with over 10% length difference to their *Naegleria gruberi* orthologues were discarded, and (3) shorter variants with 100% match to a longer ORF transcript were discarded. In the end, 178 ORFs with a start and stop codon were identified, nominated, and approved by the SSGCID target selection board and NIAID to attempt structure determination.

4.2 High-throughput Protein Expression and Purification

All proteins discussed were PCR-amplified using cDNA as a template. RNA template of *Naegleria fowleri* ATCC30215 was provided by Dr. Christopher Rice (University of Georgia, Athens) through RNA extraction and cDNA synthesis using previously published methodology in *Acanthamoeba* (47). PCR, cloning, screening, sequencing, expression screening, large-scale expression and purification of proteins were performed as described in previous SSGCID publications (17,48). All described constructs were cloned into a ligation-independent cloning (LIC) pET-14b derived, N-terminal His tag expression vector, pBG1861. Targets were expressed using chemically competent *E.coli* BL21(DE3)R3 Rosetta cells and grown in large-scale quantities in an auto-induction media (49). All purifications were performed on an ÄKTAexplorer (GE) using automated IMAC and SEC programs in adherence to prior established procedures (17).

4.3 Crystallization and Structure Determination

Crystal trials, diffraction, and structure solution were performed as previously published (16). Protein was diluted to 20 mg/mL and single crystals were obtained through vapor diffusion in sitting drops directly. The screens and conditions that yielded the crystals are listed in **Supplementary Table** 1. The screens that were used to find the crystallization conditions were typically JCSG+ (Rigaku Reagents), MCSG1 (Microlytic/Anatrace), Morpheus (Molecular Dimensions), in some cases supplemented by ProPlex (Molecular Dimensions) and JCSG Top96 (Rigaku Reagents). All data was integrated and scaled with *XDS* and *XSCALE* (50). Structures were solved by molecular replacement with *MOLREP* (51-53), *as implemented in MoRDa*. The structures were refined in iterative cycles of reciprocal space refinement in *Phenix* and real space refinement in *Coot* (54,55). The quality of all structures was continuously checked using *MolProbity* (56) as implemented in Phenix. Structural comparisons for analysis among homologues was done using DALI Protein Structure Comparison Server.

5. ACKNOWLEDGMENTS

This work was funded by National Institute of Allergy and Infectious Diseases contract numbers HHSN272201700059C and HHSN272201200025C. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (Grant 085P1000817). We thank the SSGCID cloning and protein production groups at the Seattle Children's Research Institute and at the University of Washington and Allison Pires for assistance with the crystallography table. We would like to thank Erin Egan, Veda Gadiraju,

Caroline Francis, Becca Marks, and Matthew Howard who worked with Craig L. Smith on analyses of many of the *N. fowleri* reported structures, but their analyses were not reported in this paper.

6. REFERENCES

1. Marciano-Cabral F, Cabral GA. The immune response to Naegleria fowleri amebae and pathogenesis of infection. FEMS Immunol Med Microbiol. 2007 Nov;51(2):243–59.

2. Grace E, Asbill S, Virga K. Naegleria fowleri: pathogenesis, diagnosis, and treatment options. Antimicrob Agents Chemother. 2015 Nov;59(11):6677–81.

3. Yoder JS, Eddy BA, Visvesvara GS, Capewell L, Beach MJ. The epidemiology of primary amoebic meningoencephalitis in the USA, 1962–2008. Epidemiol Infect. 2010 Jul;138(07):968–75.

4. Visvesvara GS, Moura H, Schuster FL. Pathogenic and opportunistic free-living amoebae: Acanthamoeba spp., Balamuthia mandrillaris, Naegleria fowleri and Sappinia diploidea. FEMS Immunol Med Microbiol. 2007 Jun;50(1):1–26.

5. NIAID Emerging Infectious Diseases/ Pathogens | NIH: National Institute of Allergy and Infectious Diseases [Internet]. [cited 2020 Oct 12]. Available from: http://www.niaid.nih.gov/research/emerging-infectious-diseases-pathogens

6. Vugia DJ. Notes from the Field: Fatal Naegleria fowleri Meningoencephalitis After Swimming in Hot Spring Water — California, 2018. MMWR Morb Mortal Wkly Rep. 2019;68.

7. Matanock A, Mehal JM, Liu L, Blau DM, Cope JR. Estimation of undiagnosed Naegleria fowleri primary amebic meningoencephalitis, United States1. Emerg Infect Dis. 2018 Jan;24(1):162–4.

8. El-Sayed A, Kamel M. Climatic changes and their role in emergence and re-emergence of diseases. Environ Sci Pollut Res. 2020 Jun;27(18):22336–52.

9. Siddiqui R, Khan NA. Primary amoebic meningoencephalitis caused by Naegleria fowleri: an old enemy presenting new challenges. PLoS Negl Trop Dis. 2014 Aug;8(8):e3017.

10. Chomba M, Mucheleng'anga LA, Fwoloshi S, Ngulube J, Mutengo MM. A case report: primary amoebic meningoencephalitis in a young Zambian adult. BMC Infect Dis. 2017 Dec;17(1).

11. Debnath A, Nelson AT, Silva-Olivares A, Shibayama M, Siegel D, McKerrow JH. In Vitro Efficacy of Ebselen and BAY 11-7082 Against Naegleria fowleri. Front Microbiol. 2018;9:414.

12. Colon BL, Rice CA, Guy RK, Kyle DE. Phenotypic screens reveal posaconazole as a rapidly acting amebicidal combination partner for treatment of primary amoebic meningoencephalitis. J Infect Dis. 2019 Mar 15;219(7):1095–103.

13. Cope JR, Conrad DA, Cohen N, Cotilla M, DaSilva A, Jackson J, et al. Use of the novel therapeutic agent miltefosine for the treatment of primary mmebic meningoencephalitis: report of one fatal and one surviving case. Clin Infect Dis Off Publ Infect Dis Soc Am. 2016 Mar 15;62(6):774–6.

14. Bellini NK, Santos TM, da Silva MTA, Thiemann OH. The therapeutic strategies against Naegleria fowleri. Exp Parasitol. 2018 Apr;187:1–11.

15. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, et al. DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res. 2014 Jan;42:D1091-1097.

16. Myler PJ, Stacy R, Stewart L, Staker BL, Van Voorhis WC, Varani G, et al. The Seattle Structural Genomics Center for Infectious Disease (SSGCID). Infect Disord Drug Targets. 2009 Nov;9(5):493–506.

17. Bryan CM, Bhandari J, Napuli AJ, Leibly DJ, Choi R, Kelley A, et al. High-throughput protein production and purification at the Seattle Structural Genomics Center for Infectious Disease. Acta Crystallograph Sect F Struct Biol Cryst Commun. 2011 Sep 1;67(Pt 9):1010–4.

18. Baugh L, Gallagher LA, Patrapuvich R, Clifton MC, Gardberg AS, Edwards TE, et al. Combining functional and structural genomics to sample the essential Burkholderia structome. PloS One. 2013;8(1):e53851.

19. Stacy R, Begley DW, Phan I, Staker BL, Van Voorhis WC, Varani G, et al. Structural genomics of infectious disease drug targets: the SSGCID. Acta Crystallograph Sect F Struct Biol Cryst Commun. 2011 Sep 1;67(Pt 9):979–84.

20. Milanes JE, Suryadi J, Abendroth J, Van Voorhis WC, Barrett KF, Dranow DM, et al. Enzymatic and structural characterization of the Naegleria fowleri glucokinase. Antimicrob Agents Chemother. 2019 Apr 25;63(5).

21. Tehlivets O, Malanovic N, Visram M, Pavkov-Keller T, Keller W. S-adenosyl-L-homocysteine hydrolase and methylation disorders: yeast as a model system. Biochim Biophys Acta. 2013 Jan;1832(1):204–15.

22. Cai S, Li Q-S, Borchardt RT, Kuczera K, Schowen RL. The antiviral drug ribavirin is a selective inhibitor of S-adenosyl-L-homocysteine hydrolase from Trypanosoma cruzi. Bioorg Med Chem. 2007 Dec 1;15(23):7281–7.

23. Czyrko J, Sliwiak J, Imiolczyk B, Gdaniec Z, Jaskolski M, Brzezinski K. Metal-cation regulation of enzyme dynamics is a key factor influencing the activity of S-adenosyl-L-homocysteine hydrolase from Pseudomonas aeruginosa. Sci Rep. 2018 Jul 27;8(1):11334.

24. Kusakabe Y, Ishihara M, Umeda T, Kuroda D, Nakanishi M, Kitade Y, et al. Structural insights into the reaction mechanism of S-adenosyl-L-homocysteine hydrolase. Sci Rep. 2015 Dec;5(1).

25. Reddy MCM, Kuppan G, Shetty ND, Owen JL, Ioerger TR, Sacchettini JC. Crystal structures of Mycobacterium tuberculosis S-adenosyl-L-homocysteine hydrolase in ternary complex with substrate and inhibitors. Protein Sci Publ Protein Soc. 2008 Dec;17(12):2134–44.

26. Oost J, Huynen MA, Verhees CH. Molecular characterization of phosphoglycerate mutase in archaea. FEMS Microbiol Lett. 2002 Jun;212(1):111–20.

27. Sainas S, Dosio F, Boschi D, Lolli ML. Targeting human onchocerciasis: recent advances beyond ivermectin. Annual Reports in Medicinal Chemistry. Elsevier. 2018.1–38.

28. Krause CD, Yang Z-H, Kim Y-S, Lee J-H, Cook JR, Pestka S. Protein arginine methyltransferases: Evolution and assessment of their pharmacological and therapeutic potential. Pharmacol Ther. 2007 Jan;113(1):50–87.

29. Hu H, Qian K, Ho M-C, Zheng YG. Small molecule inhibitors of protein arginine methyltransferases. Expert Opin Investig Drugs. 2016;25(3):335–58.

30. Mathioudaki K, Papadokostopoulou A, Scorilas A, Xynopoulos D, Agnanti N, Talieri M. The PRMT1 gene expression pattern in colon cancer. Br J Cancer. 2008 Dec 9;99(12):2094–9.

31. Baldwin RM, Morettin A, Paris G, Goulet I, Côté J. Alternatively spliced protein arginine methyltransferase 1 isoform PRMT1v2 promotes the survival and invasiveness of breast cancer cells. Cell Cycle. 2012 Dec 15;11(24):4597–612.

32. Tewary SK, Zheng YG, Ho M-C. Protein arginine methyltransferases: insights into the enzyme structure and mechanism at the atomic level. Cell Mol Life Sci CMLS. 2019 Aug;76(15):2917–32.

33. Mann SA, DeMart MK, May B, Causey CP, Knuckley B. Histone H4-based peptoids are inhibitors of protein arginine methyltransferase 1 (PRMT1). Biochem J. 2020 Aug 28;477(16):2971–80.

34. Mann SA, Salsburg A, Causey CP, Knuckley B. The development and characterization of a chemical probe targeting PRMT1 over PRMT5. Bioorg Med Chem. 2019 01;27(1):224–9.

35. Fulton MD, Brown T, Zheng YG. Mechanisms and Inhibitors of Histone Arginine Methylation. Chem Rec. 2018 Dec;18(12):1792–807.

36. Kaniskan HÜ, Eram MS, Zhao K, Szewczyk MM, Yang X, Schmidt K, et al. Discovery of potent and selective allosteric inhibitors of protein arginine methyltransferase 3 (PRMT3). J Med Chem. 2018 Feb 8;61(3):1204–17.

37. Fedoriw A, Rajapurkar SR, O'Brien S, Gerhart SV, Mitchell LH, Adams ND, et al. Anti-tumor Activity of the Type I PRMT Inhibitor, GSK3368715, Synergizes with PRMT5 Inhibition through MTAP Loss. Cancer Cell. 2019 08;36(1):100-114.e25.

38. Stocki P, Sawicki M, Mays CE, Hong SJ, Chapman DC, Westaway D, et al. Inhibition of the FKBP family of peptidyl prolyl isomerases induces abortive translocation and degradation of the cellular prion protein. Mol Biol Cell. 2016 Mar 1;27(5):757–67.

39. Ünal CM, Steinert M. Microbial peptidyl-prolyl cis/trans isomerases (PPIases): virulence factors and potential alternative drug targets. Microbiol Mol Biol Rev MMBR. 2014 Sep;78(3):544–71.

40. Sinars CR, Cheung-Flynn J, Rimerman RA, Scammell JG, Smith DF, Clardy J. Structure of the large FK506-binding protein FKBP51, an Hsp90-binding protein and a component of steroid receptor complexes. Proc Natl Acad Sci. 2003 Feb 4;100(3):868–73.

595 41. Dunyak BM, Gestwicki JE. Peptidyl-Proline Isomerases (PPIases): Targets for natural
596 products and natural product-inspired compounds. J Med Chem. 2016 10;59(21):9622–44.

597 42. Lee E-Y, Kim S, Kim MH. Aminoacyl-tRNA synthetases, therapeutic targets for infectious
598 diseases. Biochem Pharmacol. 2018 Aug;154:424–34.

599 43. Pham JS, Dawson KL, Jackson KE, Lim EE, Pasaje CFA, Turner KEC, et al. Aminoacyl-
600 tRNA synthetases as drug targets in eukaryotic parasites. Int J Parasitol Drugs Drug
601 Resist. 2014 Apr;4(1):1–13.

602 44. Pang YLJ, Poruri K, Martinis SA. tRNA synthetase: tRNA aminoacylation and beyond.
603 Wiley Interdiscip Rev RNA. 2014 Aug;5(4):461–80.

604 45. Burke B, Yang F, Chen F, Stehlin C, Chan B, Musier-Forsyth K. Evolutionary coadaptation
605 of the motif 2--acceptor stem interaction in the class II prolyl-tRNA synthetase system.
606 Biochemistry. 2000 Dec 19;39(50):15540–7.

607 46. Hewitt SN, Dranow DM, Horst BG, Abendroth JA, Forte B, Hallyburton I, et al. Biochemical
608 and structural characterization of selective allosteric inhibitors of the Plasmodium
609 falciparum drug target, prolyl-tRNA-synthetase. ACS Infect Dis. 2017 Jan 13;3(1):34–44.

610 47. Rice CA, Campbell SJ, Bisson C, Owen HJ, Sedelnikova SE, Baker PJ, et al. Structural
611 and functional studies of histidine biosynthesis in Acanthamoeba spp. demonstrates a
612 novel molecular arrangement and target for antimicrobials. PloS One.
613 2018;13(7):e0198827.

614 48. Choi R, Kelley A, Leibly D, Nakazawa Hewitt S, Napuli A, Van Voorhis W. Immobilized
615 metal-affinity chromatography protein-recovery screening is predictive of crystallographic
616 structure success. Acta Crystallograph Sect F Struct Biol Cryst Commun. 2011 Sep
617 1;67(9):998–1005.

618 49. Studier FW. Protein production by auto-induction in high-density shaking cultures. Protein
619 Expr Purif. 2005 May;41(1):207–34.

620 50. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010 Feb;66(Pt 2):125–32.

621 51. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta Crystallogr D Biol
622 Crystallogr. 2010 Jan;66(Pt 1):22–5.

623 52. Lebedev AA, Vagin AA, Murshudov GN. Model preparation in *MOLREP* and examples of
624 model improvement using X-ray data. Acta Crystallogr D Biol Crystallogr. 2008 Jan
625 1;64(1):33–9.

626 53. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al.
627 *REFMAC* 5 for the refinement of macromolecular crystal structures. Acta Crystallogr D Biol
628 Crystallogr. 2011 Apr 1;67(4):355–67.

629 54. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D
630 Biol Crystallogr. 2004 Dec;60(Pt 12 Pt 1):2126–32.

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

633 56. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al.
634 MolProbity: all-atom structure validation for macromolecular crystallography. Acta
635 Crystallogr D Biol Crystallogr. 2010 Jan;66(Pt 1):12–21.

636