UNIVERSITY OF CALIFORNIA SAN DIEGO

Combating *Schistosoma mansoni* by targeting proteasome and a gut-maintaining gene and the discovery of novel serine protease inhibitors in cyanobacteria

A thesis submitted in satisfaction of the requirements for the degree Master of Science

in

Biology

by

Chenxi Liu

Committee in charge:

Professor Anthony O'Donoghue, Chair
Professor Alisa Huffaker, Co-Chair
Professor Matthew Daugherty

2021

The Thesis of Chenxi Liu is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

BTZ…….. Bortezomib

C20S……. Constitutive human proteasome 20S core particle

CFZ…….. Carfilzomib

DMSO…….. Dimethyl sulfoxide

FNKL-AMC……. Ac-Phe-Gln-Lys-Leu-AMC

FnKRRK…….. MCA-Phe-Nle-Lys-Arg-Arg-Lys(DNP)-OH

Hnf4……. A homologue of hepatocyte nuclear factor 4

I20S…….. Immunoproteasome 20S core particle

LLE-AMC…… z-Leu-Leu-Glu-AMC

LLVY-AMC……. Suc-Leu-Leu-Val-Tyr-AMC

LRR-AMC…… Boc-Leu-Arg-Arg-AMC

MSP-MS……... Multiplex Substrate Profiling by Mass Spectrometry

RNAi……. RNA interference

RT……. Room Temperature

Sm20S…….. Schistosoma proteasome 20S core particle

SPLE-AMC……. Ac-Ser-Pro-Leu-Glu-AMC

ZFR-AMC…… Z-Phe-Arg-AMC

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. Anthony O'Donoghue and Dr. Conor Caffrey's mentorship. They enabled me to grow as a scientist with their immense guidance and support.

I would also like to acknowledge Dr. Zhenze Jiang and Dr. Pavla Fajtova for their help throughout the last two years. Zhenze and Pavla taught me many skills, such as LC-MS/MS and proteasome purification. I could not have become a better scientist without their help.

I would also like to acknowledge my friendship with Lawrence Liu, which has helped me get through hard times.

The RNAi of hnf4 for *Schistosoma mansoni* portion of Chapter 1, Chapter 2, and Chapter 3 is a modified reprint of the material as it appears in the article published in Science. 2020 Sep 25;369(6511):1644-1649 by George Wendt, Lu Zhao, Rui Chen, Chenxi Liu, Anthony J. O'Donoghue, Conor R. Caffrey, Michael L Reese, James J Collins 3rd. The thesis author is a co-author of the article.

The Tutuilamide portion of Chapter 1, Chapter 2, and Chapter 3 is a modified reprint of the material as it appears in the article published in ACS Chem Biol. 2020 Mar 20;15(3):751-757 by Lena Keller, Kirley Marques Canuto, Chenxi Liu, Brian M Suzuki, Jehad Almaliti, Asfandyar Sikandar, C Benjamin Naman, Evgenia Glukhov, Danmeng Luo, Brendan M Duggan, Hendrik Luesch, Jesko Koehnke, Anthony J. O'Donoghue, William H. Gerwick. The thesis author is a co-author of the article.

The development of substrates for *S. mansoni* proteasome portion of Chapter 1, Chapter 2, and Chapter 3 is a full reprint of the material as it appears in the manuscript prepared to be submitted by Zhenze Jiang, Chenxi Liu, Pavla Fajtova, Nelly El-Sakkary, Danielle Skinner,

Lawrence Liu, Ali Syed, Steven C. Wang, Conor R. Caffrey, Anthony J. O'Donoghue. The thesis author is a co-author of the article.

ABSTRACT OF THE THESIS

Combating *Schistosoma mansoni* by targeting proteasome and a gut-maintaining gene and the discovery of novel serine protease inhibitors in cyanobacteria

by

Chenxi Liu

Master of Science in Biology

University of California San Diego, 2021

Professor Anthony O'Donoghue, Chair
Professor Alisa Huffaker, Co-Chair

Project 1: Schistosomiasis is a neglected tropical disease with a single drug available: Praziquantel. Using substrate profiling by mass spectrometry on *S. mansoni* proteasome, we obtained the substrate specificity profile for the β1, β2 and β5 subunits of the proteasome. Novel fluorogenic substrates were designed. The new β5 substrate (FNKL-AMC) proved to be much more potent than the commercial substrate (LLVY-AMC). These substrates demonstrated similar

activity in three *Schistosoma* species, proving that they can detect proteasome activity in more than just *S. mansoni*.

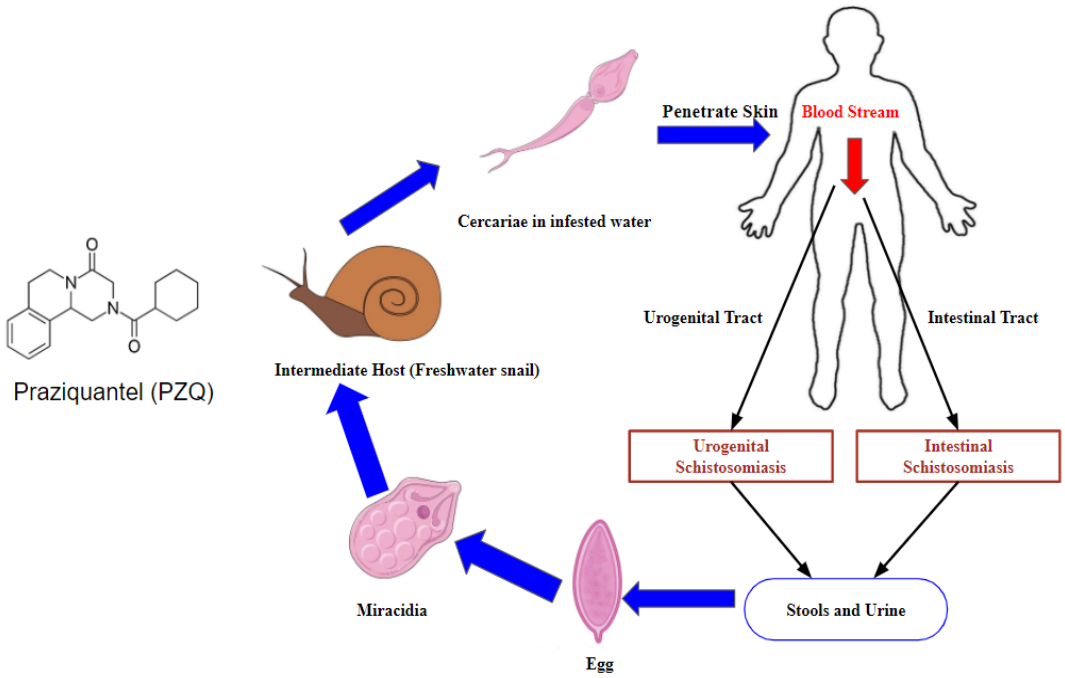
Project 2: Finding novel drug targets within the parasite is needed. Dr. James Collins and his team performed a comprehensive single-cell RNAseq on adult *Schistosoma*. They categorized 43,642 cells and assigned them into 68 distinctive populations. A homologue of hepatocyte nuclear factor 4 (hnf4) was discovered in a lineage of gut cells. The cathepsin activity in worms that received RNAi of hnf4 was measured. It was revealed that cathepsin B activity was decreased 8.2-fold relative to the control group. I showed that hnf4 is essential for the cathepsin-mediated hemoglobin digestion of *S. mansoni*.

Project 3: Marine natural products possess diverse biological activities. IC_{50} values of Tutuilamide A-C and two additional analogues were calculated for pancreatic elastase, fungal proteinase K, trypsin, and chymotrypsin. Tutuilamide A and B demonstrated potent IC_{50} values against pancreatic elastase (1.18 nM and 2.05 nM, respectively) and proteinase K (103.7 nM and 87.6 nM, respectively). These compounds proved to be 2 to 4-fold more potent compared to previously studied elastase inhibitor lyngbyastatin 7.

Introduction

1.1 Schistosomiasis (Bilharzia) and *Schistosoma* species

Schistosomiasis (Bilharzia) is a neglected tropical disease. It affects approximately 200 million people in 74 countries [1]. Among parasitic diseases, schistosomiasis is considered second only to malaria in terms of socioeconomic effects [1]. Several families of the genus *Schistosoma* are responsible for human infections: *Schistosoma japonicum* and *Schistosoma mansoni* infect the liver, whereas *Schistosoma haematobium* infects the bladder and kidneys [2]. *S. mansoni* and *S. haematobium* are found in the Middle East and Africa. *S. japonicum* is mainly located in Asia [3]. Schistosomes undergo several stages during their pathogenic lifecycle (Figure. 1). The eggs of schistosomes are released from feces and urine and are hatched upon reaching freshwater. Ciliated miracidia then infect an appropriate freshwater snail host. After two sporocyst stages in snails, cercariae are released into the water, where they can infect humans [3]. The symptoms of an acute schistosomiasis infection often include fatigue, fever, diarrhea, abdominal pain. Chronic schistosomiasis leads to severe conditions such as liver fibrosis, urinary and intestinal disease [4]. Currently, there is no vaccine for schistosomiasis, and treatment relies on one drug, Praziquantel (PZQ) [5]. Although PZQ is an effective treatment of schistosomiasis in most cases, it lacks efficacy against developing schistosomes, and resistance against PZQ has been reported [6]. There exists an urgent need to develop novel therapeutics and to discover new drug targets within the parasite.



.

Figure 1. **The life cycle of schistosomes and the structure of Praziquantel (PZQ)**. Eggs are released from feces and urine. Miracidia are hatched and proceed to infect freshwater snails. Cercariae are released into the water and infect humans. Praziquantel (PZQ) is a commonly used drug to treat schistosomiasis.

1.2 Proteasome as a potential drug target in *S. mansoni*

The proteasome is a protein complex responsible for protein degradation and turnover. It ubiquitinates proteins targeted for degradation and hydrolyzes them into short peptides. The proteasome consists of the 20S core particle and the 26S regulatory particle. The 20S core particle contains two stacked rings of 7 β subunits between two rings of 7 α subunits [7]. The β subunits are responsible for protein hydrolysis. In eukaryotes such as human and schistosomes, β1, β2, and β5 possess different substrates specificities: "caspase-like," "trypsin-like," and "chymotrypsin-like," respectively [8]. Based on each subunit's profile, several commonly used commercial fluorogenic substrates are used to report their activities: Suc-Leu-Leu-Val-Tyr-AMC (LLVY-AMC) for β5 subunit, z-Leu-Leu-Glu-AMC (LLE-AMC) for β1 subunit, and Boc-Leu-Arg-Arg-AMC (LRR-

AMC) for β2 subunits [9]. Bortezomib (BTZ) and Carfilzomib (CFZ) are common inhibitors of the proteasome. BTZ, originally an anticancer drug, is a boronic acid inhibitor that reversibly binds to the β1 and β5 subunits of the proteasome. CFZ is an anticancer drug that irreversibly inhibits the β5 subunit of the proteasome [10]. Unlike BTZ, CFZ belongs to the epoxyketone family. Researchers in the past have characterized highly potent pathogen-specific proteasome inhibitors, such as the discovery of plant-derived natural proteasome inhibitors for *Mycobacterium tuberculosis* [11]. The proteasome of *Schistosoma mansoni* is essential in its host invasion and degradation of host proteins [12]. It has been shown that BTZ and CFZ both inhibit the activity of Sm20S [12]. The development of novel proteasome inhibitors relies on performing *in-vitro* assays with pathogen proteasome using commercially available fluorogenic substrates: LLVY-AMC, LLE-AMC, and LRR-AMC. Given that there have been no reports on the development of Schistosome-specific fluorogenic substrates, my research group is leveraging mass spectrometry to obtain the substrate specificity profile of Sm20S, thereby designing more efficient substrates the report the proteasome activity.

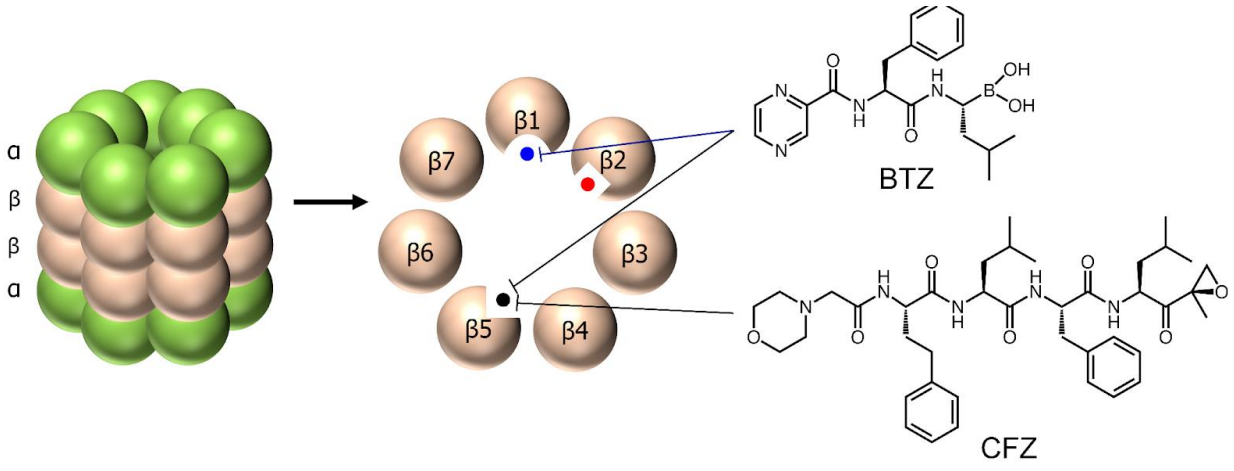


Figure 2. Structure of the schistosome proteasome and Bortezomib (BTZ), Carfilzomib (CFZ). BTZ preferentially inhibits β1 and β5 subunits. CFZ Preferentially inhibits the β5 subunit.

1.3 Multiplex substrate profiling by mass spectrometry

My advisor Dr. Anthony O'Donoghue has previously developed a method to profile the substrate specificity of proteases: multiplex substrate profiling by mass spectrometry (MSP-MS). This method utilizes a rationally designed substrate library consisting of 228 tetradecapeptide synthetic peptides. In the development of this diverse peptide library, all possible neighboring amino acid pairs were included [13]. The library is capable of detecting endopeptidases, carboxypeptidases, and aminopeptidases because of the unmodified carboxyl and amino termini of the substrates [14]. During a typical MSP-MS assay, when the protease of interest is incubated with the 228 tetradecapeptide library, cleavage products of specific time points are preserved and analyzed by LC-MS/MS. The cleavage products are then compared to the original library sequences for profiling the cleavage sites (Figure 3). The efficiency of cleavages is also calculated by label-free quantification (determination of the relative number of peptides in samples). In the past, MSP-MS has been used to determine proteolytic activities of protease of interest, such as a serine carboxypeptidase from *Pseudogymnoascus destructans*, the fungal pathogen that causes White-nose Syndrome in bats [15]. It has also been used to profile the excretory-secretory products of larval migratory schistosomes, adult worms, and eggs [16]. In this thesis, I describe the application of MSP-MS in designing schistosome-specific proteasome substrates. The novel substrates will be compared to "classical" proteasome substrates, and their efficacy will be evaluated in three major Schistosoma species*: S. mansoni*, *S. japonicum*, and *S. haematobium*.

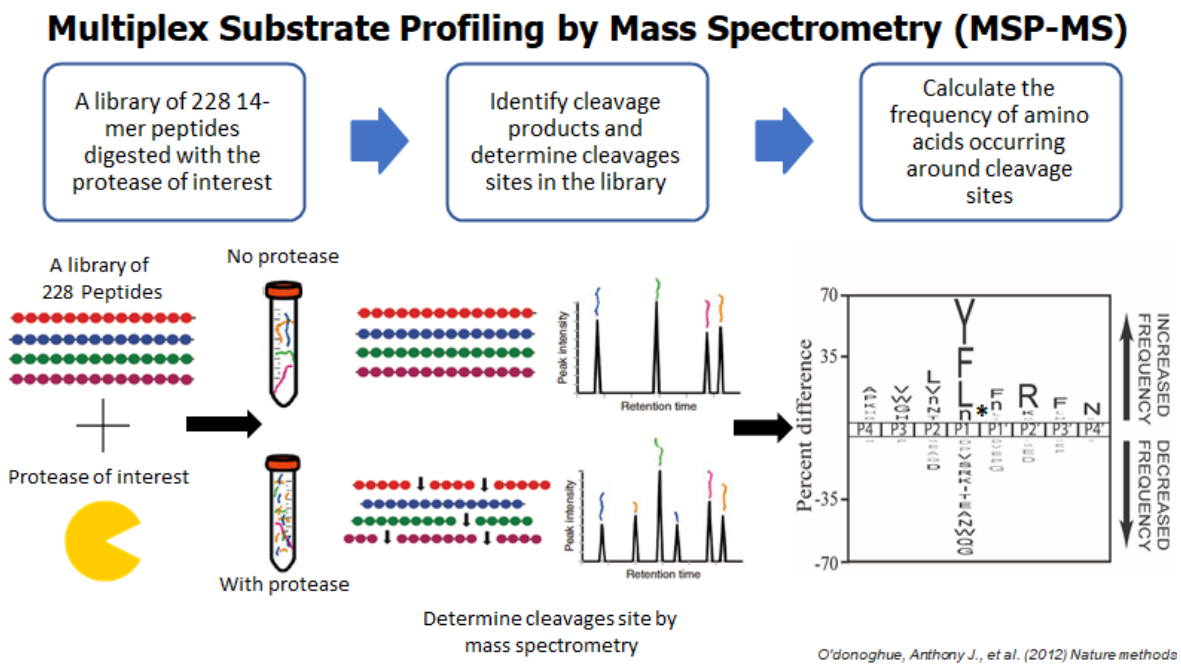


Figure 3. The general workflow of MSP-MS, Multiplex substrate profiling by mass spectrometry. The procedures include incubating the MSP-MS library with the protease of interest, identifying and matching cleavage products by MSP-MS, and generating a logo graph representing cleavage frequency.

1.4 Single-cell RNAseq of *Schistosoma mansoni* cells

Like all other metazoans, schistosomes consist of various tissue types. Therefore, understanding the schistosome's tissues on a molecular level and their roles during infection could yield novel therapeutics. Single-cell RNAseq is a widely used method to profile tissue types in diverse organisms [17]. In a typical single-cell RNAseq study, library construction usually involves dissociation and isolation of single cells, capturing the intracellular mRNA, reverse-transcribing to cDNA. cDNA is then amplified and sequenced. Cellular libraries are then labeled with barcodes and analyzed [18]. Prior to this, no studies have used single-cell RNAseq on pathogenic Schistosomes. Dr. Collin's research at the UT Southwestern Medical Center performed single-cell RNAseq to characterize the pathogenic (adult) stage of the parasite lifecycle. From the data of 43,642 animal cells, they characterized 68 molecularly distinct cell populations that include nearly all tissues described morphologically (Figure 4). Notably, they discovered a lineage of somatic

stem cells that produces and maintains Schistosome's gut [19]. Within this lineage, a homologue of hepatocyte nuclear factor 4 (hnf4) is expressed and is required for gut maintenance, blood feeding, and inducing egg-associated pathology in vivo. This study utilizes single-cell RNAseq to understand Schistosome biology while identifying a potential drug target that can be perturbed to disrupt its pathology in mammalians.

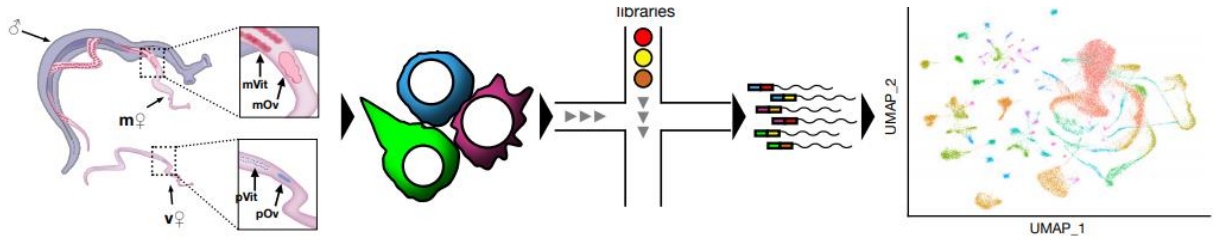


Figure 4. **Schematic of single-cell RNAseq of *S.mansoni***, which includes (from left to right): dissociation of parasites (male paired with a mature female worm (m♀) that possess a mature ovary (mOv) and vitellaria (mVit); unpaired virgin female worms (v♀) possess a primordial ovary (pOv) and vitellaria (pVit); cell-sorting; generate scRNAseq libraries and sequencing; perform clustering.

1.5 Cysteine proteases as digestive enzymes in schistosomes

Cysteine proteases are essential enzymes in *Schistosoma* [20]. Digestion of blood proteins from the host by Schistosome happens in the gut, where an intricate network of cysteine proteases and aspartic proteases form a digestive "network." [21] Cathepsins are diverse proteases that are active at low pH. Specifically, some cathepsins of the cysteine protease family have been shown to mediate the hemoglobin-digestion of Schistosomes, such as the *S. mansoni* cathepsin B1 [21]. After Dr. Collin's group discovered hnf4 in the parasite's gut, we tested whether knocking out hnf4 using RNAi would result in a loss of cysteine protease activity that contributes to hemoglobin digestion. Specifically, the fluorogenic substrate Z-Phe-Arg-AMC can be used to evaluate cysteine protease (cathepsin) activity of lysates from the control worms, and hnf4 (RNAi treated) worms [22]. To distinguish cathepsin B activity from broad cysteine protease activity, the broad cysteine protease inhibitor, E-64, or the cathepsin B-selective inhibitor, CA-074, were used [23]. In addition

to testing the cysteine protease activity, it is also of interest to evaluate the level of aspartic acid protease in lysates. Hence, mca-GKPILFFRL-K(dnp) was to measure aspartic acid protease activity [24]. Pepstatin A, an aspartic acid protease inhibitor, was used to deconvolute this activity [25]. Taken together, these data could show that hnf4 is essential for cathepsin-mediated digestion of hemoglobin in *S. mansoni*.

1.6 **Discovery of novel elastase inhibitors in marine cyanobacteria**

Marine cyanobacteria, also called blue-green algae, is an outstanding natural product source that possesses diverse biological activities, including antifungal, cytotoxic, anti-parasitic, antibacterial, and antiviral activities [26]. In recent years, many cyanobacterial secondary metabolites have been isolated and studied. In this study, using mass spectrometry-based fractionation and molecular networking, a cyanobacterial field collection of *Schizothrix sp*. from American Samoa produced two new cyclic peptides: Tutuilamide A and B [27]. Their structures were established by spectroscopic techniques, including 1D and 2D NMR, HR-MS, and chemical derivatization. These cyclic peptides are characterized by the presence of several unusual residues, including 3-amino-6-hydroxy-2-piperidone (Ahp) and 2-amino-2-butenoic acid (Abu), together with a novel vinyl chloride-containing residue (Figure 5). Cyanobacteria produce a wide range of cyclic depsipeptides that contain an unusual 3-amino-6-hydroxy-2-piperidone (Ahp) moiety, generally referred to as cyanopeptolin-like peptides [28]. The general structure includes an amino acid chain of variable length, of which six amino acid residues form a ring. All cyanopeptolin-like peptides contain the Ahp residue and an ester linkage between the β-OH group of L-threonine and the carboxyl group of the C-terminal amino acid [29]. Additionally, among these Ahp containing peptides, more than 20 compounds have a highly conserved hexadepsipeptide core bearing an

additional 2-amino-2-butenoic acid (Abu) moiety adjacent to the Ahp ring and a highly variable side chain (Table 1). Lyngbyastatin 5-7 were collected in South Florida, Fort Lauderdale/Florida Keys, Summerland Key from *Lyngbya confervoides* [30]. Similarly, Symplostatin 2 was collected in Guam, Pago Bay from *Symploca hydnoides*) [31].

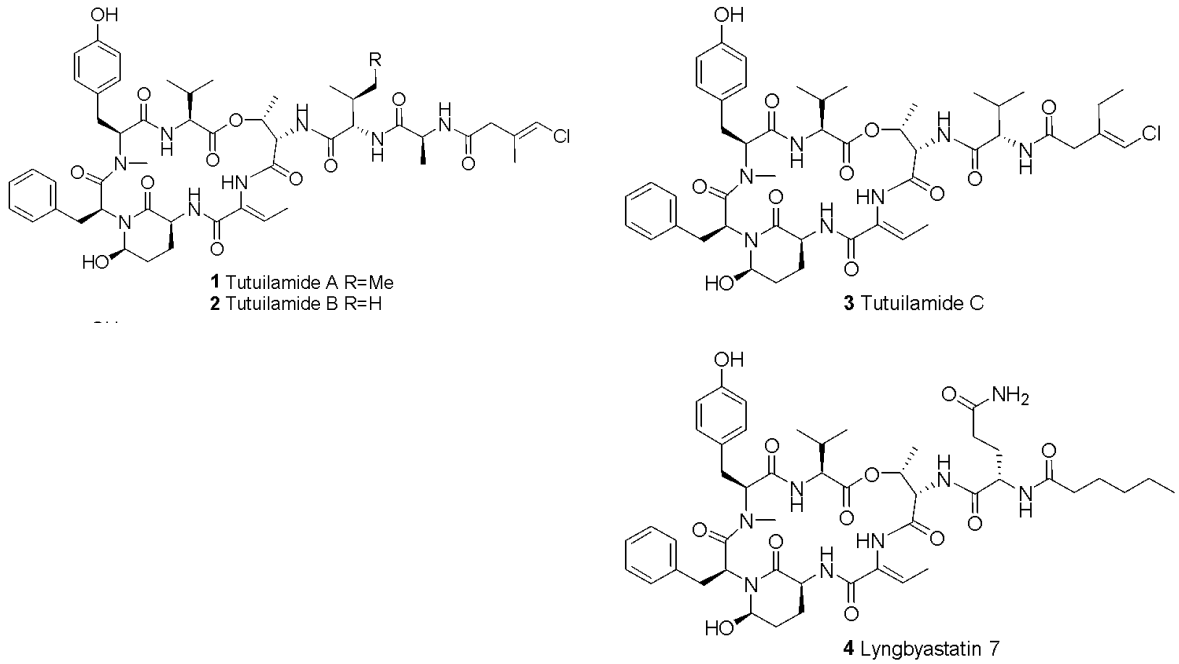


Figure 5. Structures of Tutuilamide A-C and Lyngbyastatin7. 1-3. Structure of Tutuilamide A, B and C. **4.** Structure of Lyngsyastatin 7

Ahp-Abu-containing cyclic hexadepsipeptides commonly inhibit serine proteases with a high level of selectivity for elastase and chymotrypsin inhibition [32]. The binding mode of Tutuilamide A and B to elastase was analyzed by X-ray crystallography, revealing a reversible binding mode that was similar to Lyngbyastatin 7. X-ray crystallography of the elastase complex formed with Lyngbyastatin 7 revealed that these Abu-cyclodepsipeptides act as substrate mimics [33]. The Abu moiety was shown to occupy the S1 substrate-binding pocket and engage in a non-covalent interaction. Additionally, a study of the binding mode of a cyanopeptolin from *Scytonema hofmanni*, scyptolin showed that the Ahp moiety occupies a crucial part of the active site pocket and thereby prevents hydrolysis [34]. It was suggested that the Abu moiety increases the potency of

these elastase inhibitors and that the pendant side chain is responsible for modulating their activity and selectivity [33]. The presence of an additional hydrogen bond with the amino acid backbone of the flexible side chain of Tutuilamide A, compared to Lyngbyastatin 7, facilitates its stabilization in the elastase binding pocket. The new natural products Tutuilamides, as well as two semi-synthetic derivatives, were evaluated for elastase inhibition.

Table 1. **Abu and Ahp-containing peptides isolated from Cyanobacteria.**

Compound name	Year	Collected from	Isolated organism
Dolastatin 13[6]	1989	East Africa, Indian Ocean	*Dolabella auricularia* (sea hare)
Symplostatin 2[7]	1999	Guam, Pago Bay	*Symploca hydnoides*
Somamide A+B[8]	2001	Fiji, Somo Somo	*Lyngbya majuscula / Schizothrix* sp. assemblage
Lyngbyastatin 4[9]	2007	South Florida, Atlantic coast	*Lyngbya confervoides*
Lyngbyastatin 5-7[10]	2007	South Florida, Fort Lauderdale / Florida Keys, Summerland Key	*Lyngbya confervoides*
Lyngbyastatin 8-10[11] / Bouillomide A+B[12]	2009	Guam, Tumon Bay	*Lyngbya semiplena*
Molalssamide[13]	2010	Florida, Molasses Reef	*Dichothrix utahensis*
Stigonemapeptin[14]	2012	Wisconsin, North Nokomis Lake	*Stigonema* sp. (freshwater)
Symplostatin 5-10[15]	2013	Guam, Cetti Bay	red *Symploca* sp.
Kurahamide[16]	2014	Japan, Kuraha	*Lyngbya* sp. assembly

1.7 Goals of this thesis

This thesis is focused on three independent projects that each have a set of goals. For Project 1, the goal was to use MSP-MS to profile the substrate specificity of the proteasome of *Schistosoma mansoni* in order to develop novel fluorogenic substrates. These substrates will then be compared to traditional proteasome substrates and tested in *S. japonicum* and *S. haematobium*. In Project 2, a single-cell RNAseq of Schistosome cells was performed and I will evaluate the effects of gut protease activity on knocking out an analogue of hepatocyte nuclear factor 4. In Project 3: I will describe several novel serine protease inhibitors found in cyanobacteria and their inhibition potency toward elastase and proteinase K.

MATERIALS AND METHODS

2.1 Maintenance of *S. mansoni*

The parasite (NMRI isolate) of *S. mansoni* was maintained by passage through *Biomphalaria glabrata* (NMRI isolate) snails and female LVG Golden Syrian hamsters (infected at 4–6 weeks of age) as intermediate and definitive hosts, respectively [35, 36]. Six weeks post-infection, hamsters were euthanized with an overdose of Fatal-Plus, and the mature, mixed-sex parasites recovered by reverse perfusion of the hepatic portal system [36, 37]. Parasites were washed five times in ice-cold Basch medium 169, and the medium removed prior to freezing the worms at -80 °C [36, 38]. All experiments with mice were performed under a protocol that was approved by the Institutional Animal Care and Use Committee (IACUC) at UCSD. Approval from UCSD-IACUC is granted under two policies: The Animal Welfare Act and Regulations (AWAR) Act and the United States Public Health Service Policy on Humane Care and Use of Laboratory Animals.

2.2 Purification of *S. mansoni* proteasome

S. mansoni worms were homogenized using a pestle connected to a motor in 1.5-ml tubes in ice-cold 100 mM Tris-HCl, 100 µM E-64, pH 7.5. Lysates were centrifuged for 15 min at 14,000 ×*g* and 4 °C, and the supernatant was subjected to two ammonium sulfate precipitation steps at 30 and 60% saturation, respectively, each for 60 min on ice. After centrifugation for 15 min at 14,000 ×g at 4 °C, the supernatant was discarded, and the precipitated proteins were resuspended in ice-cold 100 mM Tris-HCl, 100 µM E-64, pH 7.5. The samples were concentrated using 100 kDa centrifugal filter units (Amicon). A portion of the samples will be saved and assayed as semi-pure *S. mansoni* lysate. Lysates of Schistosoma japonicum and Schistosoma haematobium was produced concurrently using the same method.

The rest of the samples were loaded onto a Superose 6 10/300 gel filtration column under an ÄKTA Pure instrument (GE Healthcare Life Sciences). Proteins were eluted using 50 mM HEPES, 10% glycerol, 0.125 M NaCl, pH 7.5. Fractions of 0.5 ml were collected and assayed for proteasome activity with 25 μM Suc-LLVY-AMC in assay buffer (50 mM Tris, 0.02% SDS, pH 7.5). The assay was performed at 24°C with Ex/Em = 360 nm/460 nm using a Synergy HTX multi-mode reader (BioTek Instruments). Fractions containing proteasome activity were pooled and loaded onto a 5 mL anion exchange HiTrap DEAE FF column (Sigma). Proteins were eluted using 50 mM HEPES, 10% glycerol, pH 7.5, and a linear gradient from 0.125 to 0.6 M NaCl. Fractions (1.5 mL) were collected and assayed with Suc-LLVY-AMC as is described above. Fractions containing proteasome activity were pooled, concentrated using 100 kDa centrifugal filter units (Amicon), and stored in -80°C. The schematic of proteasome purification is presented in Supp. Fig. 1.

2.3 Multiplex substrate profiling of *S. mansoni* proteasome by mass spectrometry

All MSP-MS assays were performed in quadruplicate using a mixture of 228-synthetic tetradecapeptides at a final concentration of 0.5 μM for each peptide. The concentration of *S. mansoni* proteasome used in MSP-MS is determined by turning over LLVY-AMC at 1 RFU/sec. 10 μL of the reaction mixture was removed after 30 and 90 minutes of incubation. Protease activity was quenched by the addition of GuHCl (MP Biomedicals), and samples were immediately stored at -80 °C. Samples were thawed, acidified to pH < 3.0 with 1% formic acid, desalted with C18 spin-columns and dried with using a vacuum centrifuge.

~0.4 μg of peptides were injected into a Q-Exactive Mass Spectrometer (Thermo) equipped with an Ultimate 3000 HPLC. Peptides were separated by reverse-phase chromatography on a C18

column (1.7 um bead size, 75 um x 25 cm, 65°C) at a flow rate of 300 nl/min using a 60-minute linear gradient from 5% to 30% B, with solvent A: 0.1% formic acid in water and solvent B: 0.1% formic acid in acetonitrile. Survey scans were recorded over a 150–2000 m/z range (70,000 resolutions at 200m/z, AGC target 3×106, 100 ms maximum). MS/MS was performed in data-dependent acquisition mode with HCD fragmentation (28 normalized collision energy) on the 12 most intense precursor ions (17,500 resolutions at 200m/z, AGC target 1×105, 50 ms maximum, dynamic exclusion 20 s). Data were processed using PEAKS 8.5 (Bioinformatics Solutions Inc.). MS_2 data were searched against the MSP-MS tetradecapeptide library sequences. A precursor tolerance of 20 ppm and 0.01 Da for MS2 fragments were utilized. No specific protease digestion was specified. Data were filtered to 1% peptide level false discovery rates. Peptides were quantified with label-free quantification, and data were normalized by median and filtered by 0.3 peptide quality. Missing values and zeros are imputed with random normally distributed numbers in the range of the average of the smallest 5% of the data ± SD.

2.4 Proteasome activity and inhibition assays

To compare selected commercial proteasome inhibitors with new substrates, purified S. mansoni proteasome was preincubated with DMSO, 2μM of BTZ, or 2μM of CFZ for 30 minutes in assay buffer: 50 mM Tris-HCl pH 7.5, 0.02% SDS. The final concentration for selected commercial substrates and designed substrates is 10 μM. For the IC50 experiment, dilutions of CFZ and BTZ were set to be 12 points of 3-fold dilution, beginning at the highest concentration of 2 μM. Purified S. mansoni proteasome was preincubated with CFZ, BTZ for 30 minutes in assay buffer: 50 mM Tris-HCl pH 7.5, 0.02% SDS. For the Michaelis–Menten kinetics study with Purified S. mansoni proteasome, the concentrations of substrates were set up as follows: 8 points

of 2-fold dilution with the highest concentration at 100 µM (LLVY-AMC, FnKL-AMC), 8 points of 2-fold dilution with the highest concentration at 500 µM (LLE-AMC, Mca-VDQMDGWK(DNP)-NH2), 8 points of 2-fold dilution with the highest concentration at 500 µM (LRR-AMC, Mca-FnKRRK(DNP)-OH). For all the experiments above, release of fluorescence was recorded in a Synergy HTX multi-mode reader (BioTek Instruments, Winooski, VT) with excitation and emission wavelengths at 320 nm and 400 nm for IQ substrates, and 360 nM and 460 nM for AMC substrates, respectively. The maximum velocity was calculated in RFU/sec over 10 sequential points on the linear part of the progress curve. All data presented as triplicates ± standard deviation.

2.5 Single-cell RNA sequencing and RNA interference of hnf4

FACS-sorted *Schistosoma* cells were centrifuged at 500 g for 10 minutes at 4°C then resuspended in 0.2% BSA in PBS. Libraries were created using a Chromium Controller (10x Genomics) according to manufacturer guidelines and sequenced using a NextSeq 500 (Illumina). Sequencing data were processed and mapped to the *Schistosoma mansoni* genome (v7) using Cell Ranger (10x Genomics). Unfiltered data from Cell Ranger was imported into Seurat (v3.1.1)34, 35 and cells were filtered as follows: Female (nFeature_RNA (> 750), nCount_RNA (1500-20000), Percent Mitochondrial (<3%); Male/Virgin female (nFeature_RNA (> 750), nCount_RNA (1000-20000), Percent Mitochondrial (<3%)). Mitochondrial genes were identified as those with the prefix “Smp_9”. Each of the 9 individual datasets was normalized (NormalizeData) and variable features were identified (FindVariableFeatures, selection.method = “vst”, nfeatures = 2000). From here, integration anchors were identified (FindIntegrationAnchors, dims 1:78), the data was integrated (IntegrateData, dims = 1:78, features.to.integrate = features), and scaled

(ScaleData). We then ran RunPCA, RunUMAP (reduction = "pca", dims = 1:78, n.neighbors = 40), FindNeighbors (reduction = "pca", dims = 1:78), FindClusters (resolution = 5). The number of principal components (78) used for this analysis was defined by JackStraw. Analysis of the resulting single-cell map found that clusters 27 and 50 contained few enriched markers. Therefore we removed the 964 cells present in these clusters and reran the analysis with 78 principal components. From here we generated the final UMAP projection plot with RunUMAP (n.neighbors = 36, min.dist = 0.70, dims = 1:80). Next, we generated clusters (FindClusters, resolution = 5) and manually inspected the unique genes expressed in each of the clusters. In some cases, we found that some of the 85 resulting clusters did not express a core set of unique genes. Therefore, these clusters were merged into a single cluster of cells as follows: Neoblasts (clusters 0,1,2,6,7,37), Neoblast progeny (cluster 4,8), Neuron 1 (clusters 10, 60, 68), Neuron 6 (clusters 24, 26), Parenchyma (clusters 11, 12, 51), flame cells (clusters 14, 41), S1 Cells (clusters 3, 9, 32, 42) and tegument (clusters 36, 63). After merging, we were left with a final map of 68 clusters of 43,643 cells. For RNAi of hnf4, all experiments utilized freshly perfused male parasites (separated from females) unless otherwise noted. dsRNA treatments were all carried out at 30 μg/ml in Basch Medium 169. dsRNA was generated by *in vitro* transcription.

2.6 qPCR and RNAseq

RNA collection was performed as previously described [39] with the following modifications. Parasites were treated with dsRNA, and whole parasites were collected in Trizol. RNA was purified from samples utilizing Direct-zol RNA miniprep kits (Zymo Research R2051). Quantitative PCR analyses were performed as previously described [39, 40]. cDNA was synthesized using iScript™ cDNA synthesis kit (Bio-Rad 1708891), and qPCR was performed as previously

described [41] utilizing iTaq™ Universal SYBR® Green Supermix (Bio-Rad 1725122) and a QuantStudio 3 Real-Time PCR System (Applied Biosystems). RNAseq on hnf4(RNAi) parasites was performed as previously described [41] using TruSeq Stranded mRNA Library Prep (illumina 20020594) to prepare libraries, which were sequenced on a NextSeq 550 (illumina). The total number of reads per gene was determined by mapping the reads to the *S. mansoni* genome (v7) using STAR (version 020201) [42]. *S. mansoni* genome sequence and GTF files used for mapping were acquired from Wormbase Parasite [43]. Pairwise comparisons of differential gene expression were performed with DESeq2 (version 1.12.2) [44]. Volcano plots were made using the "volc" function from ggplot2. To remove genes expressed at lower levels, genes with a base-mean expression value less than 50 were excluded from the analysis. Furthermore, genes that were differentially expressed (padj < 0.05) that were not assigned to the automatically assigned to the "gut" cluster during initial clustering were manually examined in the single-cell RNAseq data, and those that were expressed in the gut were reclassified to the "gut" cluster. Raw data from hnf4 RNAi RNAseq experiments are available at NCBI GEO with accession number GSE146737.

2.7 Cysteine protease assay on RNAi-treated *S. mansoni* lysates

10 worms of positive control and RNAi treated *S. mansoni* were homogenized using a pestle connected to a handheld motor in 2 mL tubes with 300 μL of 100 mM Sodium Citrate, pH 5.5, followed by a 30s sonication at 30% power by a sonic dismembrator (Fisher Scientific, FB120110). Lysates were centrifuged for 15 mins at 15,000 ×g at 4 °C (Eppendorf 5418R, 5401000137). The pellet and upper lipid layer were discarded, and the protein concentration of the remaining supernatant was determined using the Pierce BCA kit (Thermo Fisher Scientific, 23225). Each experimental condition was assayed in 30 μL of total volume reactions containing 1

μg of protein and 10 μM of each substrate (H-Arg-AMC, z-Phe-Arg-AMC, z-Val-Arg-AMC, z-Arg-Arg-AMC, H-Gly-Arg-AMC) in 100 mM Sodium Citrate, pH 5.5 with 2 mM Dithiothreitol, as triplicate wells on a black Nunc® Polystyrene 384-Well Plate (Thermo Scientific, 262260). The E-64 negative control group was incubated with 10 μM E-64 (Sigma-Aldrich, E3132) for 30 mins at room temperature. For Z-Phe-Arg-AMC and Z-Arg-Arg-AMC, an additional CA-047 control group was set up by 10 μM CA-047 (Caymen Chemical, 24679) for 30 mins at room temperature. The release of the AMC fluorophore was recorded in a Synergy HTX multi-mode reader (BioTek Instruments, Winooski, VT) with excitation and emission wavelengths at 340 nm and 460 nm, respectively. The maximum velocity was calculated in RFU/sec over 10 sequential points on the linear part of the progress curve. The activity of proteases was quantified as RFU/sec/1 μg.

2.8 Surgical transplantation of schistosomes

Seven days prior to surgery, 5-week-old parasites were recovered from mice and treated with 30 μg/ml dsRNA for 7 days in Basch Medium 169 (see Table S6 "strategy 8"). Before mice were anesthetized, 10 pairs (male and female) were sucked into a 1ml syringe, the syringe was fitted with a custom 25G extra thin wall hypodermic needle (Cadence, Cranston, RI), the air and all but ~200 μL of media were purged from the needle, and the syringe was placed needle down in a test tube to settle the parasites to the bottom of the syringe. Mice were kept on infrared heating pads for the duration of the surgery. Following wound closure, mice received a single subcutaneous 20 μL dose of a 1 mg/mL solution of Buprenorphine SR-LAB CIII for analgesia and were allowed to recover on a warm heating pad. Mice were group-housed, and individual mice were tracked by ear punches. On either day 22 or day 30, post-transplantation mice were sacrificed and perfused to recover parasites. Male and female parasites were counted and fixed for 4 hours

in 4% formaldehyde in phosphate buffered saline with 0.01% TritonX-100 (PBSTx). The recipient's livers were removed and fixed for 72 hours in 4% formaldehyde in PBS. The percentage parasite recovery was determined by dividing the total (male and female) number of worms transplanted by the total number of parasites recovered following perfusion. Livers from individual mice were sectioned and processed for Haematoxylin and Eosin staining by the UT Southwestern Molecular Pathology Core.

2.9 Aspartic protease assay on RNAi-treated *S. mansoni* lysates

10 worms of positive control and RNAi treated *S. mansoni* were homogenized using a pestle connected to a handheld motor in 2 ml tubes with 300 μL of 100 mM Sodium Citrate, pH 5.5, followed by a 30s sonication at 30% power by a sonic dismembrator (Fisher Scientific, FB120110). Lysates were centrifuged for 15 mins at 15,000 ×g at 4 °C (Eppendorf 5418R, 5401000137). The pellet and upper lipid layer were discarded, and the protein concentration of the remaining supernatant was determined using the Pierce BCA kit (Thermo Fisher Scientific, 23225). Each experimental condition was assayed in 30 μL of total volume reactions containing 1 μg of protein and 10 μM of the FRET substrate (Mca-Gly-Lys-Pro-Ile-Leu-Phe-Phe-Arg-Leu-Lys(DNP)-DArg-NH2) in 100 mM Sodium Citrate, pH 3.5, as triplicate wells on a black Nunc® Polystyrene 384-Well Plate (Thermo Scientific, 262260). The E-64 control group and Pepstatin A control group were incubated with 10 μM E-64 (Sigma-Aldrich, E3132) and 2 μM Pepstatin A (Caymen Chemical, 9000469) respectively for 30 mins at room temperature. The release of fluorescence was recorded in a Synergy HTX multi-mode reader (BioTek Instruments, Winooski, VT) with excitation and emission wavelengths at 320 nm and 400 nm, respectively. The maximum velocity was calculated in RFU/sec over 10 sequential points on the linear part of the progress

curve. The activity of proteases was quantified as RFU/sec/1 μg.

2.10 Tutuilamide dose-response Assay

Stocks of porcine pancreatic elastase, Type 1 (Sigma-Aldrich, P2308), human chymotrypsin (Sigma-Aldrich, 230900), bovine trypsin (Sigma-Aldrich, T-7409), and proteinase K (Sigma-Aldrich P-2308) were stored at −20°C and diluted to the appropriate concentration in assay buffer consisting of Dulbecco's phosphate buffered saline, pH 7.4, 0.01% Tween-20 and 2 mM dithiothreitol. All inhibitors were stored in DMSO and diluted to 20 μM in assay buffer. Compounds were then serially diluted 3-fold in assay buffer and mixed with an equal volume of the enzyme. After 30 minutes of incubation at 22°C, 15 μL of the enzyme-inhibitor mixture was added to wells of black 384-well plates (Thermo Scientific, 262260) containing 15 μL of substrate diluted in assay buffer. Fluorescence was measured in 3-minute intervals at 22°C using a Synergy HTX Multi-Mode Microplate Reader (BioTek, Winooski, VT) with excitation and emission wavelengths of 360 and 460 nm, respectively. Protease activity was reported as the change in relative fluorescent units per second over a 30 minute time interval. The final concentration of inhibitor ranged from 20 μM to 339 pM for trypsin and chymotrypsin assays and from 5 μM to 85 pM for elastase and proteinase K assays. The final enzyme and substrate concentration in each assay were as follows: bovine trypsin (10 nM) and z-Leu-Arg-Arg-AMC (100 μM); human chymotrypsin (10 nM) and Suc-Ala-Ala-Pro-Phe-AMC (100 μM); porcine pancreatic elastase (10 nM) and MeOSuc-Ala-Ala-Pro-Val-AMC (223 μM), proteinase K (10 nM) and Suc-Ala-Ala-Pro-Phe-AMC (100 μM).

The research described referring to the RNAi of hnf4 for *Schistosoma mansoni* is a modified reprint of the material as it appears in the article published in Science. 2020 Sep

25;369(6511):1644-1649 by George Wendt, Lu Zhao, Rui Chen, Chenxi Liu, Anthony J. O'Donoghue, Conor R. Caffrey, Michael L Reese, James J Collins 3rd. The thesis author is a co-author of the article.

The research described referring to Tutuilamides is a modified reprint of the material as it appears in the article published in ACS Chem Biol. 2020 Mar 20;15(3):751-757 by Lena Keller, Kirley Marques Canuto, Chenxi Liu, Brian M Suzuki, Jehad Almaliti, Asfandyar Sikandar, C Benjamin Naman, Evgenia Glukhov, Danmeng Luo, Brendan M Duggan, Hendrik Luesch, Jesko Koehnke, Anthony J. O'Donoghue, William H. Gerwick. The thesis author is a co-author of the article.

The research described referring to the development of substrates for *S. mansoni* proteasome is a full reprint of the material as it appears in the manuscript prepared to be submitted by Zhenze Jiang, Chenxi Liu, Pavla Fajtova, Nelly El-Sakkary, Danielle Skinner, Lawrence Liu, Ali Syed, Steven C. Wang, Conor R. Caffrey, Anthony J. O'Donoghue. The thesis author is a co-author of the article.

RESULTS

3.1 Deconvoluting and characterizing *S. mansoni* proteasome subunit activity

S. mansoni proteasome (Sm20S) consists of three major catalytic subunits: β1, β2, and β5. Each subunit possesses distinct substrate specificities. The fluorogenic substrates LLVY-AMC, LLE-AMC, and LRR-AMC can be used to report the activity of β5, β1, and β2, respectively. CFZ selectively inhibits β5 activity of the proteasome, whereas BTZ selectively inhibits β1 and β5 activity [12]. To deconvolute the activity of the three subunits in MSP-MS, we employed carfilzomib (CFZ) and bortezomib (BTZ) in an IC_{50} assay with pure *S. mansoni* proteasome (Figure 6B).

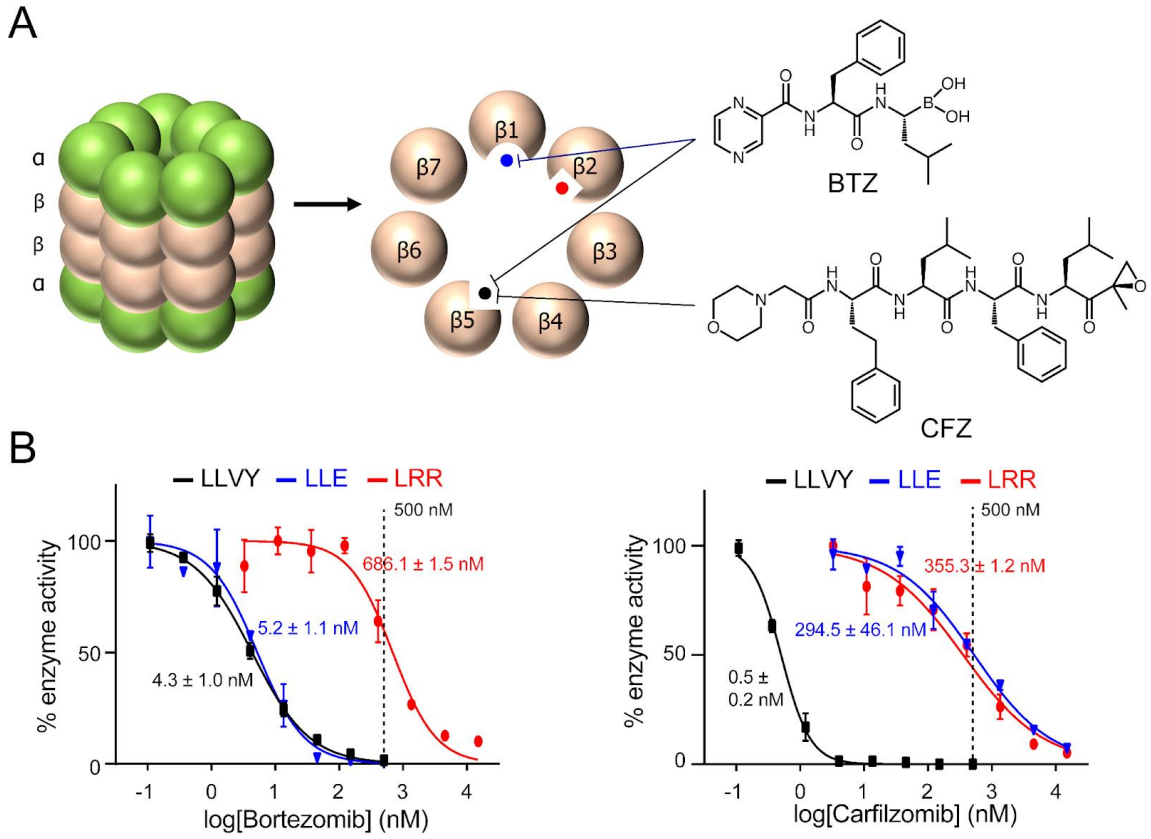


Figure 6. Characterization of substrate specificities of Sm20S proteasome subunits. A. Bortezomib (BTZ) and carfilzomib (CFZ) were used in MSP-MS assay to selectively inhibit proteasome subunits to reveal their specificities. **B.** The potency of bortezomib and carfilzomib was determined. 500 nM of BTZ and CFZ were used to achieve selective inhibition of β1&β5 and β5 subunits, respectively.

IC_{50} values of BTZ and CFZ are shown in Fig. 1B. BTZ has an IC_{50} of 4.3 nM, 5.2 nM,

and 686.1 nM using LLVY-AMC (β5), LLE-AMC (β1), and LRR-AMC (β2), respectively. CFZ has an IC_{50} of 0.5 nM, 294.5 nM, and 355.3 nM for LLVY-AMC, LLE-AMC, and LRR-AMC, respectively (Figure 6B). The results indicate that we can isolate β2 activity by the addition of 500 nM BTZ to inhibit both β5 and β1 and isolate for β1 and β2 activity by using 500 nM CFZ to inhibit β5.

3.2 Mapping the substrate profile of *S. mansoni* proteasome subunit

Based on the IC_{50} data described above, we first performed MSP-MS in the presence of 500 nM BTZ to identify the cleavage profile of β2 (Figure 7). To obtain the cleavage profile of β5, we performed MSP-MS in the presence and absence of 500 nM CFZ and performed a subtractive analysis to identify the profile due specifically to β5 (Figure 7). Finally, the β1 subunit specificity was determined by subtracting CFZ-sensitive (β5) cleavage profile from the BTZ-sensitive (β5 and β1) cleavage profile. Following the MSP-MS assay, the results were analyzed using software Peak and R Studio. The β1 subunit showed a strong preference for Glu and Asp at P1. It also preferred Ser and Glu at P4; Tyr, Pro and Ala at P3; and Leu, Gly, Trp, and Arg at P2. Glu and Ala were also found to be favored at P3′ position. The β2 subunit showed a broad preference for Arg (Lys) between P3 and P2′ favoring Lys and Arg at P1. Finally, the β5 subunit exhibited a strong preference for hydrophobic residues at P4 and P3, charged residues at P2, and hydrophobic residues at P1. Between P1′ and P3′, hydrophobic residues were preferred.

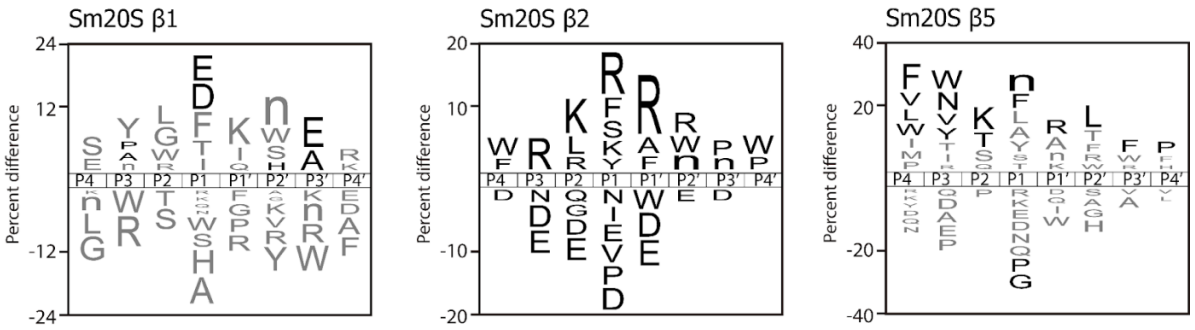


Figure 7. Substrate profiles of β1, β2, and β5 catalytic subunit of *S. mansoni* proteasome (Sm20S). The cleavage profile of β2 was obtained in the presence of 500 nM BTZ. We performed MSP-MS in the presence and absence of 500 nM CFZ and performed a subtractive analysis to identify the profile due specifically to β5. The β1 subunit specificity was determined by subtracting CFZ-sensitive (β5) cleavage profile from the BTZ-sensitive (β5 and β1) cleavage profile.

3.3 Designing novel fluorogenic substrates based on MSP-MS and evaluating their efficacy

Through GenScript**,** we designed four substrates for each of the three Sm20S catalytic subunits: β1-AMC (SPLE-AMC) and β1-IQ1 (Mca-EYADSK(DNP)-NH2) for β1; β2-IQ (Mca-FnKRRK(DNP)-OH) for β2 and β5-AMC (FNKL-AMC) for β5. The IQ substrates were designed by utilizing the prime-side specificities we observed in MSP-MS assays. In addition, we evaluated a commercially available substrate, β1-IQ2 (Mca-VDQMDGWK(DNP)-NH2), which has been used to detect human caspase 3 activity (PMID: 17610031, 15078480). We first evaluated the efficacy of newly designed and previously reported substrates (PMID: 29182146, 24301521) with purified Sm20S. Proteasome activity was measured by fluorescence increasing rate, as shown in Figure 8A. We detected activities with β1-IQ2, β2-IQ, and β5-AMC substrates, and these provided reaction velocities that were 2x, 10x, and 3x greater than the respective 'classical' LLE, LRR, LLVY substrates. The inhibition profiles confirmed that these substrates are subunit selective. Thus, with the β1-IQ2 substrate, BTZ had an IC_{50} of 49.84 ± 0.24 nM, whereas CFZ did not inhibit the activity, suggesting the substrate is β1 selective. With the β2-IQ substrate, CFZ had an IC50 of 288.0 ± 90 nM, whereas BTZ was inactive, suggesting that the substrate is β2 selective. The

cleavage site of β2-IQ substrate was determined to be between two Arg residues by mass spectrometry (Fig. S3), which perfectly matched our specificity profile. With β5-AMC, BTZ and CFZ had IC_{50} values of 17.97 ± 2.05 nM and IC50 = 2.495 ± 0.51 nM, respectively, suggesting the substrate is β5-selective. However, we failed to detect any activity with β1-AMC and β1-IQ1 showed a very weak activity when assayed with Sm20S. One explanation for this failure is that the β1 activity of Sm20S is weakest among all three subunits. There were only a few substrates cleaved in our MSP-MS assay by β1 subunit, which was not enough to allow us to generate an accurate specificity profile.

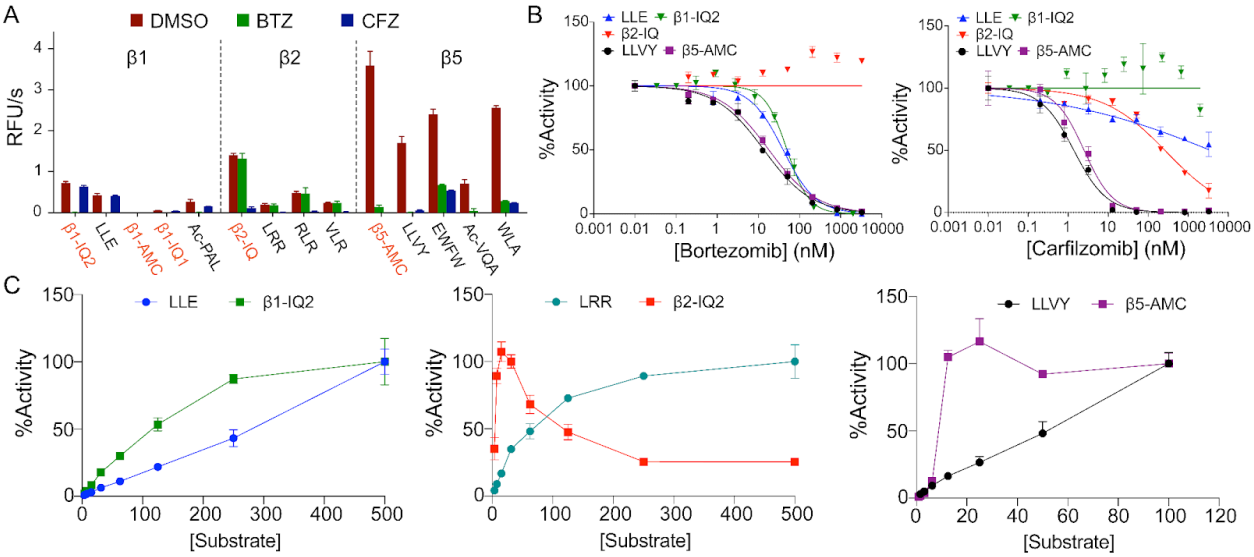


Figure 8. Sm20S subunit activity assays with novel fluorescent substrates. A. Comparison of previously reported as well as newly designed proteasome substrates when screened with purified Sm20S. BTZ and CFZ (2 μM) were pre-incubated with Sm20S to test the selectivity of the substrates. **B**. Dose-response of proteasome inhibitors with selective Sm20S proteasome substrates. Semi-pure S. mansoni lysate was used in the assays. All assays were performed in triplicate. **C.** KM studies of previously reported proteasome substrates and newly designed Sm20S substrates. Activities were normalized to the highest activity for each substrate. All data are shown as triplicate values ± SD.

Considering that the AMC and IQ substrates use different fluorophores which excite and emit at different wavelengths and are measured at different gain settings, we performed assays to determine the Michaelis-Menten kinetics for each substrate. As is shown in Fig. 8C, all our newly designed substrates exhibited stronger affinities for Sm20S comparing to traditional LLE, LRR, and LLVY substrates. β2-IQ activity decreased with increasing substrate concentrations starting

from 50 μM. This is likely due to the intermolecular quenching effect of internally quenched substrates or product inhibition.

3.4 Generating inhibitory profile of three *Schistosoma* species using novel fluorogenic substrates

Our novel substrates were designed based on the specificity profile of *S. mansoni*. However, there are other two major *Schistosoma* species that infect humans: *S. haematobium, S. japonicum*. To determine whether our substrates were suitable for detecting proteasome activities of three different *Schistosoma* species and whether our inhibition profile on *S. mansoni* can be well translated to the other two species, we evaluated their proteasome activities with our best subunit-selective substrates and several traditional *S. mansoni* proteasome inhibitors (Figure 9). LLE-AMC is a traditionally used substrate for β1 of the proteasome. We found that LLE activity was inhibited by BTZ in all three species at 125 nM, 250 nM, and 500 nM, as expected. As for the β2 substrate (FnKRR), CFZ inhibited the activity across all three species at 250 nM, 500 nM, 1000 nM. Lastly, to evaluate the activity of β5 across three species, we selected several commercial inhibitors and analogs of natural products that have been previously shown to inhibit β5 in *S. mansoni*. ONX 0914, Carmaphycin, 78A1, 1-45, 1-47 are previously described proteasome inhibitors [45, 46]. Each inhibitor showed a similar inhibitory pattern in all three species at 5 nM, 25 nM, and 125 nM with the novel β5 substrate FnKL-AMC. These results indicate these three substrates share very similar substrate specificities and active site structures *in S. mansoni, S. haematobium,* and *S. japonicum*. As we increase BTZ's concentration, LLE and β5-AMC's activity decreases across all three species, whereas an increase of CFZ concentration results in a decrease of β5 activity. They were observed to inhibit the β5-AMC's activity in *S. mansoni, S. haematobium,* and *S. japonicum,*

following a similar trend.

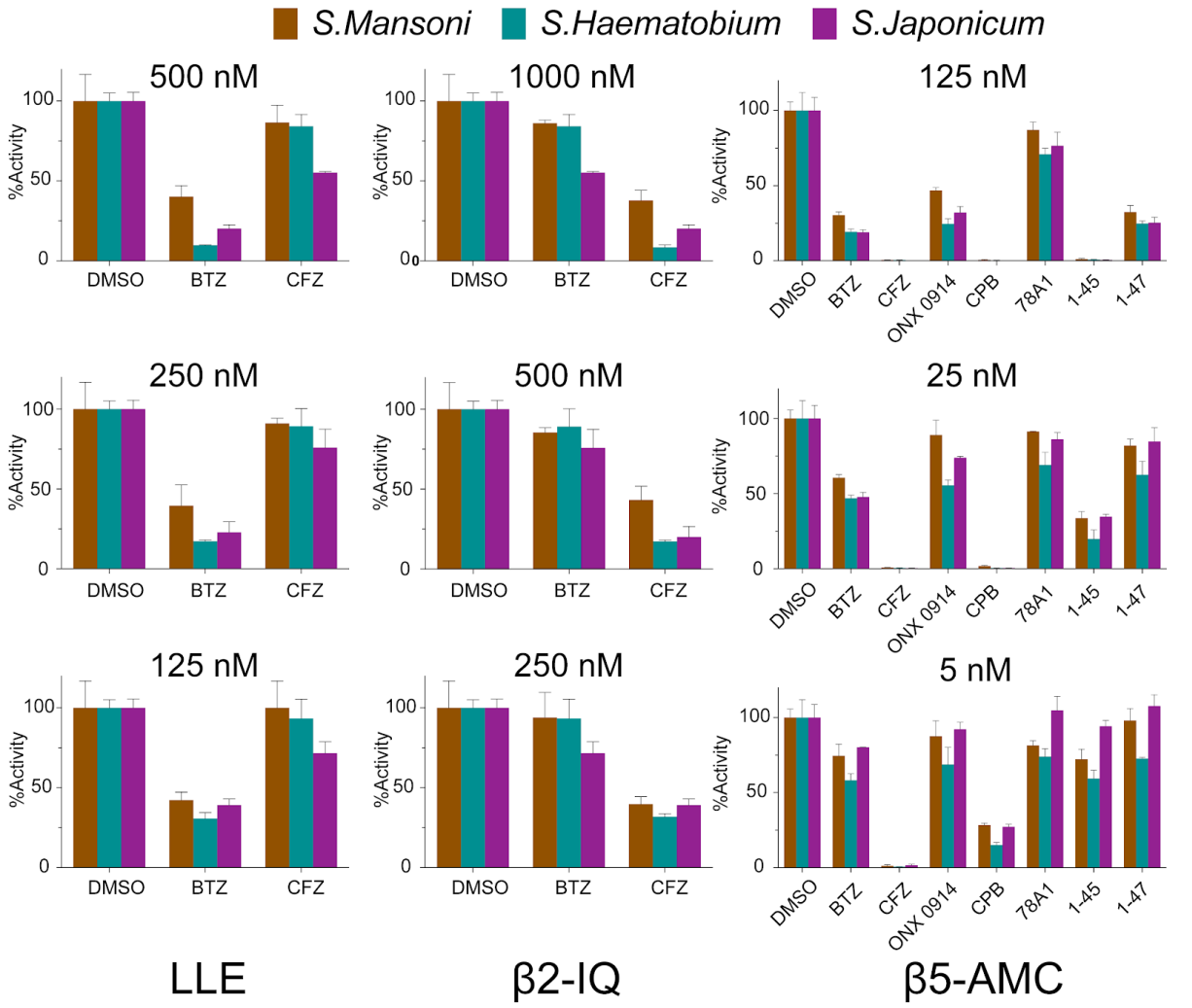


Figure 9. Comparison of proteasome activities of three *Schistosoma* species. Proteasome inhibitors were pre-incubated with semi-pure proteasomes of different *Schistosoma* species to obtain their cross-species inhibitory profile. The concentrations of the inhibitors (BTZ, CFZ) were shown on top of each chart. Data presented as value ± SD in triplicate.

3.5 hnf4 RNAi results in transcriptional gut abnormalities

Based on the localization of a subpopulation of hnf4-expressing neoblasts (eled+ neoblasts) on the UMAP projection plot (Supp. Fig. 3), and that hnf4 is a marker of gut stem cells in planarians [47], we hypothesized that hnf4 is required for gut maintenance, blood feeding, and

pathology *in vivo*. Following RNAi of hnf4, we found expression of several definitive gut markers such as ctsl (Smp_343260) and ctsb (Smp_103610) was decreased (Figure 10B). We next performed in situ hybridization (ISH) to examine the localization of transcripts. We found that several gut transcripts were no longer expressed, their expression was reduced, or their expression pattern was significantly altered (Figure 10B). To examine the extent of gut dysfunction, we performed RNAseq on hnf4(RNAi) animals. We found that over 70% of all transcripts expressed in the "gut" cluster in our single-cell data set were significantly downregulated following hnf4 RNAi (Figure 10 C). Indeed, a look at the top 25 most downregulated genes in the RNAseq experiment revealed that all were expressed in the gut and 21 were expressed almost exclusively in the gut (Figure 10D).

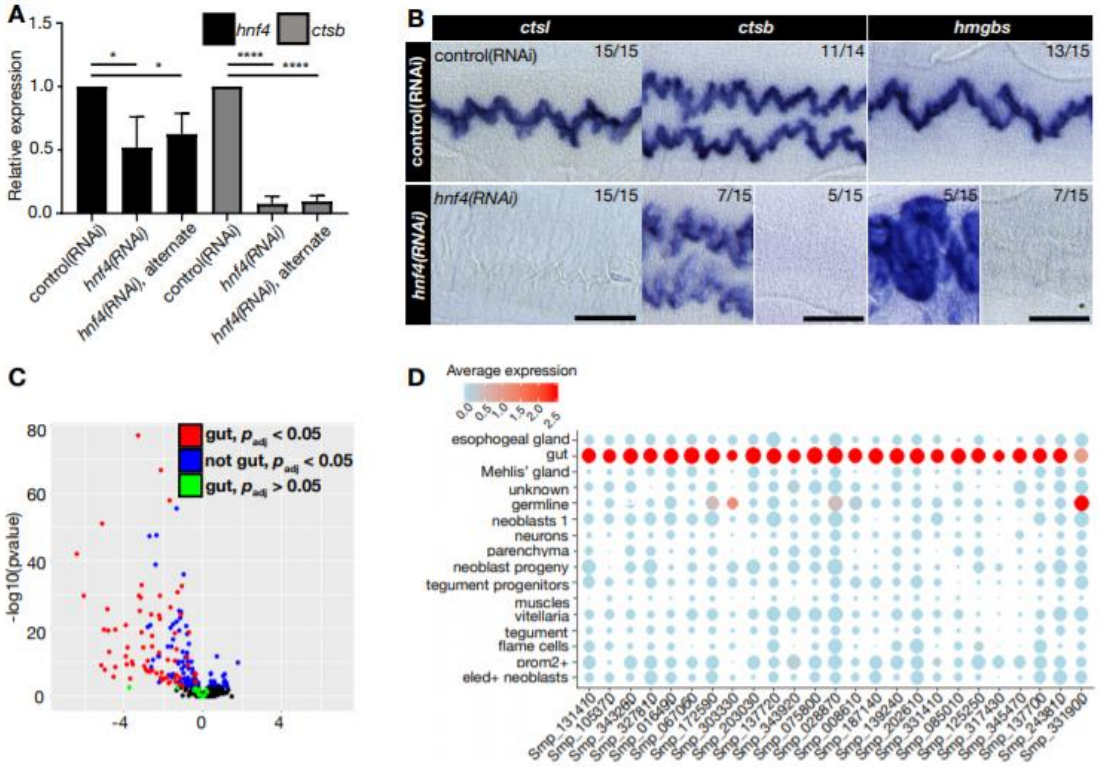


Figure 10. hnf4 RNAi results in transcriptional gut abnormalities. **A.** Graph of relative quantification of hnf4 mRNA (black) or ctsb mRNA (grey) as determined by qPCR in either "control(RNAi)," "hnf4(RNAi)", or "hnf4(RNAi) alternate" animals. Four biological replicates. **B.** For the "gut"-specific genes ctsl, ctsb, and hmgbs: WISH of the indicated gene in either control RNAi conditions or hnf4 RNAi conditions. The number of parasites similar to representative images is indicated in the upper right of each panel. n ≥ 14 parasites, three biological replicates. **C.** Volcano plot of data from an RNAseq experiment comparing gene expression of control(RNAi) animals to that of hnf4(RNAi) animals. "gut," genes expressed in the "gut" cluster, "not gut," genes not expressed in the "gut" cluster. Significance determined as padj < 0.05 by Benjamini and Hochberg-corrected Wald test. **D.** A dot-plot summarizing the cluster-specific expression of each of the top 25 down-regulated genes in hnf4(RNAi) animals. Cluster IDs are on the vertical axis, and gene IDs are on the horizontal axis. Expression levels are colored by gene expression (blue = low, red = high). The percentage of cells in the cluster expressing the gene is indicated by the size of the circle (small = few, large = many). Scale bars, 100 μm. *p < 0.05, ****p < 0.0001 (Welch's t-test).

3.6 hnf4 is essential for blood feeding and pathology of *S.mansoni*

We asked whether there were any functional consequences of hnf4 RNAi. Although glucose can be absorbed across the parasite's tegument, parasites rely on the gut to digest host blood cells [48]. To test the digestive capability of hnf4(RNAi) parasites, we added red blood cells to the media and observed the parasites' ability to uptake and digest the cells (Figure 10A). While

the vast majority of control (RNAi) parasites (67/69) were able to ingest and digest red blood cells as evidenced by black pigmentation in the gut [49], hnf4(RNAi) parasites either failed to ingest red blood cells (15/69) or ingested red blood cells but couldn't digest them as evidenced by red pigmentation in the gut (54/69). These data suggest a decrease in the blood ingestion and digestion capacity of the hnf4(RNAi) animals but do not address the mechanism of any digestive defects. Mice receiving hnf4(RNAi) parasites had morphologically normal livers compared to abundant egg-induced granulomata in livers of control parasite recipients (Figure 10C). Additionally, recovered male hnf4(RNAi) parasites were significantly shorter than controls (2.87 versus 5.21 mm, respectively; $p < 0.0001$; Welch's t-test) (Figure 10B). These results show that hnf4 is at least indirectly required for parasite growth and egg-induced pathology in vivo.

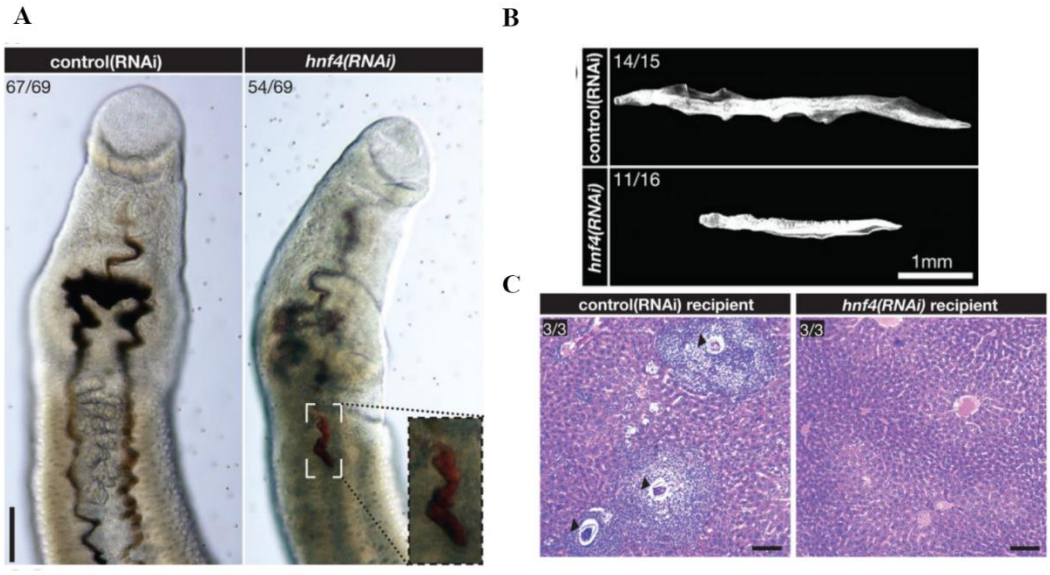


Figure 11. hnf4 is required for blood feeding and pathology. A. Brightfield microscopy images of control(RNAi) or hnf4(RNAi) animals cultured with RBCs. The inset shows a magnification of the boxed area. **B**. Parasites recovered from transplant recipients. n > 15 parasites from three recipients. Nuclei are white. The number of parasites or sections similar to the representative micrograph is in the upper-left corner of each panel. **C**. Hematoxylin and eosin (H&E)–stained mouse liver sections 22 days after transplant with RNAi-treated parasites. Arrowheads show granulomata. Sections are from n = 3 recipients.

To determine the remaining cysteine protease activity in *S. mansoni* worms that have

received the RNAi treatment of hnf4, a plate reader assay with the fluorogenic substrate z-Phe-Arg-AMC was conducted. The data in Figure 12 demonstrates three principal findings. Protease activity is represented by RFU/s/µg of protein. Firstly, compared to control, cysteine protease activity has decreased 8.2-fold in the RNAi group. Secondly, both groups' activity with the substrate z-Phe-Arg-AMC is completed eliminated upon the addition of 10 µM of E-64, a broad cysteine protease inhibitor. This suggests that Z-Phe-Arg-AMC indeed measures cysteine protease activity. Additionally, CA-074, a cathepsin-B specific inhibitor, was added, which decreases the activity of both groups to approximately the same level (0.76 ± 0.38 RFU/sec and 0.72 ± 0.38 RFU/sec respectively). This result shows us that most of the cysteine protease activity in lysates is due to cathepsin B (94%). Furthermore, we also sought to test the level of aspartic protease activity in lysates, which was determined by the ability to cleave the fluorogenic substrate mca-GKPILFFRL-K(dnp). Results in Figure 12 suggest that RNAi of hnf4 did not result in a decrease of aspartic protease activity in the lysate (33.1 ± 2.14 RFU/sec for the control group and 40.2 ± 6.12 RFU/sec for RNAi group). The addition of pepstatin A, a broad aspartic acid protease inhibitor, was added to confirm that aspartic protease activity is represented by cleaving mca-GKPILFFRL-K(dnp).

A **B**

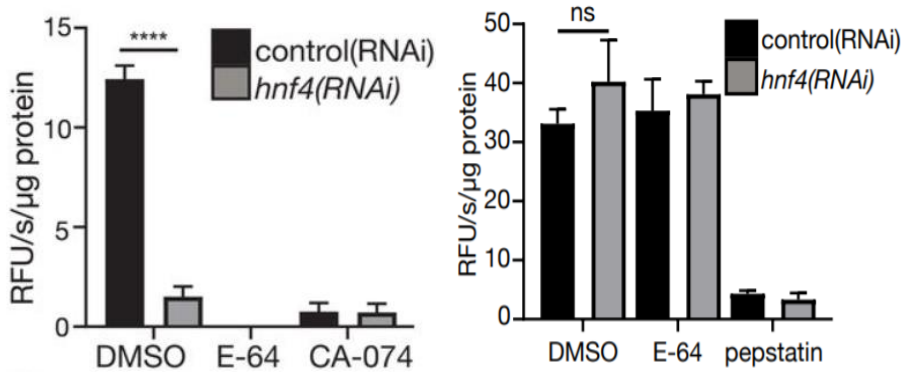


Figure 12. **Cathepsin and aspartyl protease activity of lysates from control(RNAi) or hnf4(RNAi) animals. A**. Determination of cysteine protease activity by measuring Z-Phe-Arg turnover in lysates. **B.** Evaluating aspartyl protease activity in worms by measuring MeOSuc-Ala-Ala-Pro-Val-AMC turnover in lysates**.** Data presented as triplicate with standard deviation.

3.7 Obtaining inhibition profile of Tutuilamide compounds

We predicted that pancreatic elastase would be potently inhibited by Tutuilamide A and B. 5 µM to 85 pM of Tutuilamide A-C were incubated with porcine pancreatic elastase, and their inhibition profile was compared to Lyngbyastatin 7 and Symplostatin 2. For Tutuilamide A-C, we calculated IC_{50} values of 1.18 nM, 2.05 nM, and 4.93 nM, respectively. In comparison, Symplostatin 2 and Lyngbyastatin 7 had IC_{50} values of 5.41 nM and 11.50 nM, respectively. Taken together, we show here that Tutuilamide A and Tutuilamide B are the most potent inhibitors of pancreatic elastase in this group, while Lyngbyastatin was the least potent inhibitor.

Pancreatic elastase, trypsin, and chymotrypsin are members of the S1 family of serine proteases that form a double β-barrel structure. Therefore, we wanted to test if Tutuilamide A-C can inhibit these enzymes. After incubating bovine trypsin with 10 µM of Tutuilamide A-C, no inhibition was observed. For chymotrypsin, we calculated IC_{50} values of 1014 nM, 576.6 nM, and 542 nM, respectively. Therefore, within the S1 family of serine proteases, Tutuilamide A preferentially targets elastase-type enzymes over chymotrypsin and has no activity against trypsin.,

The 2-fold reduction in potency can be explained by chymotrypsin appearing to prefer valine as found in Tutuilamide B and Tutuilamide C over isoleucine found in Tutuilamide A.

The S8 family of serine protease is also commonly researched. These enzymes are not structurally related to S1 family serine proteases. When Tutuilamide A and Tutuilamide B were incubated with proteinase K, a representative S8 protease isolated from the soil fungi, *Engyodontium album*, IC_{50} values of these compounds were calculated to be 103.7 nM and 87.6 nM, respectively. Tutuilamide C had an IC_{50} value that was greater than 50 µM. This could mean that longer peptides possess more potency for proteinase K. Taken together, these studies show that Tutuilamide A and B can inhibit enzymes in two unrelated serine protease families.

The cytotoxicity of Tutuilamide A-C toward H-460 lung cancer cell was also calculated to have IC_{50} values of 0.53 ± 0.04 µM, 1.27 ± 0.21 µM, and 4.78 ± 0.45 µM.

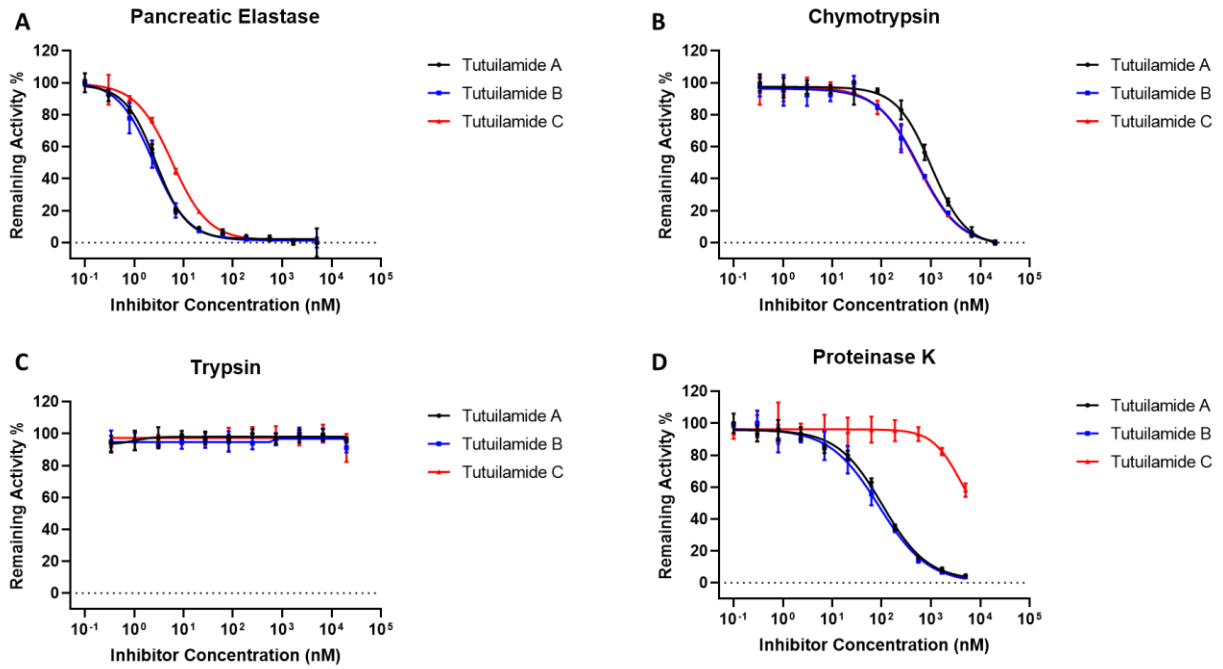


Figure 13. IC50 curves of Tuilamide A-C toward serine proteases. A-C. The potency of inhibition of Tutuilamide A-C toward pancreatic elastase, chymotrypsin, and trypsin (S1 family). **D.** Potency of inhibition of Tutuilamide A-C toward proteinase K (S8 family)

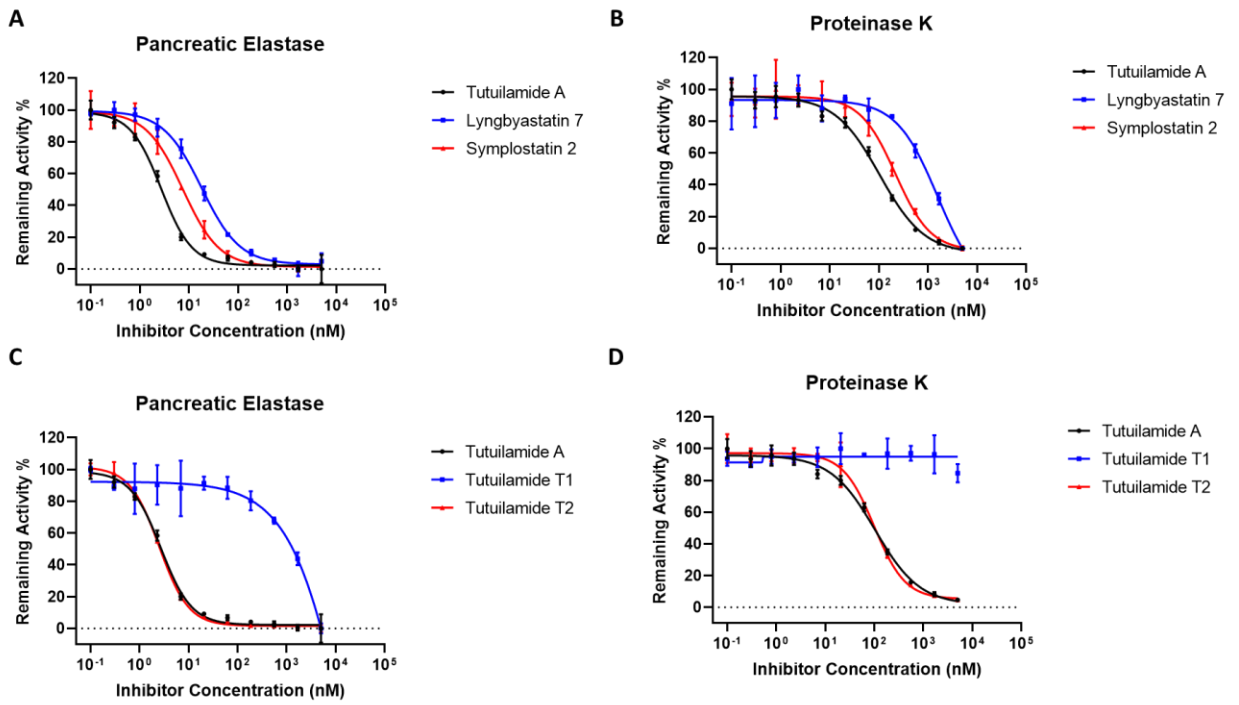


Figure 14. Comparison between Turuilamide A and Lyngbyastasin 7, Symplostatin 2, and two Tutuilamide analogues. A-B. The potency of inhibition of Lyngbyastatin 7 and Symplostatin 2 toward pancreatic elastase and proteinase K, compared to Tutuilamide A. **C-D.** The potency of inhibition of Tutuilamide T1 (linear) and Tutuilamide T2 (methyl) toward pancreatic elastase and proteinase K.

3.8 Synthesis of semi-synthetic analogues of Tutuilamide A

Two semi-synthetic analogues of Tutuilamide A were also produced. In the first semi-synthetic analogue, the depsipeptide ester in Tutuilamide A was hydrolyzed using $LiOH/H_2O$ to synthesize the acyclic analogue Tutuilamide T1 (linear). The IC_{50} of this compound for pancreatic elastase was 3994 nM, which is 3,384-fold less potent when compared to the cyclic Tutuilamide A. This analogue also did not proteinase K activity with an IC_{50} value greater than 5 µM. In addition, Tutuilamide T2 (methyl) was obtained by methylating the hemi-aminal moiety in Tutuilamide A with pyridinium p-toluenesulfonate (PPTS). This analog demonstrated satisfactory inhibition potency toward pancreatic elastase and proteinase K, with its IC_{50} values being 1.83 nM

and 98.6 nM, respectively.

Table 2: Inhibition potency of compounds for selected serine proteases with 30-minute incubations

IC50 (nM)/(95 %Confidence Interval Values)	Elastase	Chymotrypsin	Trypsin	Proteinase K
Tutuilamide A	1.179/(0.3271 - 1.965 nM)	1014/(853.0 -1227 nM)	>10 µM/(>10 µM)	103.7/(84.55 - 129.2 nM)
Tutuilamide B	2.051/(1.676 - 2.459 nM)	576.6/(447.0 - 767.3 nM)	>10 µM/(>10 µM)	87.62/(65.16 - 122.2 nM)
Tutuilamide C	4.930/(2.721 - 7.477 nM)	542.0/(445.0 - 671.2 nM)	>10 µM/(>10 µM)	>10 µM/(1458 nM to >10 µM)
Lyngbyastatin 7	11.50/(9.252 - 14.28 nM)	Not Done	Not Done	1574/(806.0 nM to >10 µM)
Symplostatin 2	5.406/(4.301 - 6.686 nM)	Not Done	Not Done	216.5(155.2 - 326.0 nM)
Tutuilamide T1	31197/(>10 µM)	Not Done	Not Done	>10 µM/(>10 µM)
Tutuilamide T2	1.832/(1.558 - 2.124 nM)	Not Done	Not Done	98.62(94.36 - 100.1 nM)

The research described referring to the RNAi of hnf4 for *Schistosoma mansoni* is a modified reprint of the material as it appears in the article published in Science. 2020 Sep 25;369(6511):1644-1649 by George Wendt, Lu Zhao, Rui Chen, Chenxi Liu, Anthony J. O'Donoghue, Conor R. Caffrey, Michael L Reese, James J Collins 3rd. The thesis author is a co-author of the article.

The research described referring to Tutuilamides is a modified reprint of the material as it appears in the article published in ACS Chem Biol. 2020 Mar 20;15(3):751-757 by Lena Keller, Kirley Marques Canuto, Chenxi Liu, Brian M Suzuki, Jehad Almaliti, Asfandyar Sikandar, C Benjamin Naman, Evgenia Glukhov, Danmeng Luo, Brendan M Duggan, Hendrik Luesch, Jesko Koehnke, Anthony J. O'Donoghue, William H. Gerwick. The thesis author is a co-author of the article.

The research described referring to the development of substrates for *S. mansoni* proteasome is a full reprint of the material as it appears in the manuscript prepared to be submitted by Zhenze Jiang, Chenxi Liu, Pavla Fajtova, Nelly El-Sakkary, Danielle Skinner, Lawrence Liu,

Ali Syed, Steven C. Wang, Conor R. Caffrey, Anthony J. O'Donoghue. The thesis author is a co-author of the article.

Discussion

The proteasome is essential for Schistosome's entry of host and degradation of host proteins, making it a potent druggable target [12]. Traditionally, Suc-Leu-Leu-Val-Tyr-AMC (LLVY-AMC) has been the commonly used fluorogenic substrate for detecting proteasome activity. It is widely used in the human constitutive proteasome (c20s) and immunoproteasome (i20s) assays. This thesis describes the development of *S. mansoni*-specific substrates using the multiplex substrate profiling assay by mass spectrometry (MSP-MS). Specifically, FNKL-AMC and FnKRR proved to be much more potent than their commercial counterparts (LLVY-AMC and LRR-AMC). They also demonstrated a similar inhibition profile in *S. mansoni*, *S. japonicum*, and *S. haemabotium* when incubated with different concentrations of bortezomib and carfilzomib. We anticipate these new substrates to be potential replacements for the commercial substrates in our future studies. This project outlines a comprehensive method to design species-specific fluorogenic substrates for the proteasome.

Dr. James Collin's group discovered hnf4, a gene expressed by a gut cell lineage in *S. mansoni* as a result of single-cell RNAseq of 43,642 cells. We anticipated that RNAi treatment of this gene would result in a decrease in cathepsin-mediated hemoglobin digestion in *S. mansoni*. I conducted enzyme assays on RNAi-treated and control worm lysates to determine the change in cathepsin level. To distinguish broad cysteine protease activity and cathepsin B activity, I incorporated a cathepsin B-selective inhibitor CA-074 and a broad cysteine protease inhibitor E-64. As a result, it was revealed that cathepsin B activity was decreased 8.2-fold relative to the control group. Cathepsin is essential in facilitating the blood-feeding of the parasite. This result is consistent with the gene expression studies conducted by the Collins group. Additionally, aspartyl protease activity of lysates was determined using the fluorogenic substrate mca-GKPILFFRLK-

K(dnp) with the presence and absence of pepstatin A, an aspartyl protease-specific inhibitor. We were not able to observe a statistically significant change in aspartic protease activity. This means that RNAi of hnf4 does not alter the level of expression of aspartic protease in *S. mansoni*. Taken together, we concluded that hnf4 is essential for cathepsin B-facilitated hemoglobin digestion in *S. mansoni*. Additionally, microscopic studies revealed that RNAi of hnf4 results in deformed gut tissues, smaller worm size. It was also observed that RNAi worms have resulted in considerably less granuloma in mice liver, which indicates that the RNAi-treated worms cannot sustain successful infection in mice. These results suggest that hnf4 is essential in *Schistosoma mansoni*'s gut maintenance, blood digestion, and pathology. We have identified a potent drug target in *S. mansoni* that we can perturb to disrupt its pathology in mammalians. This paper is one of the most comprehensive single-cell RNAseq studies of any parasite to date. More importantly, by combining RNAseq, RNAi, and protein assay techniques, this publication provides a blueprint for future investigations of novel drug targets in parasites.

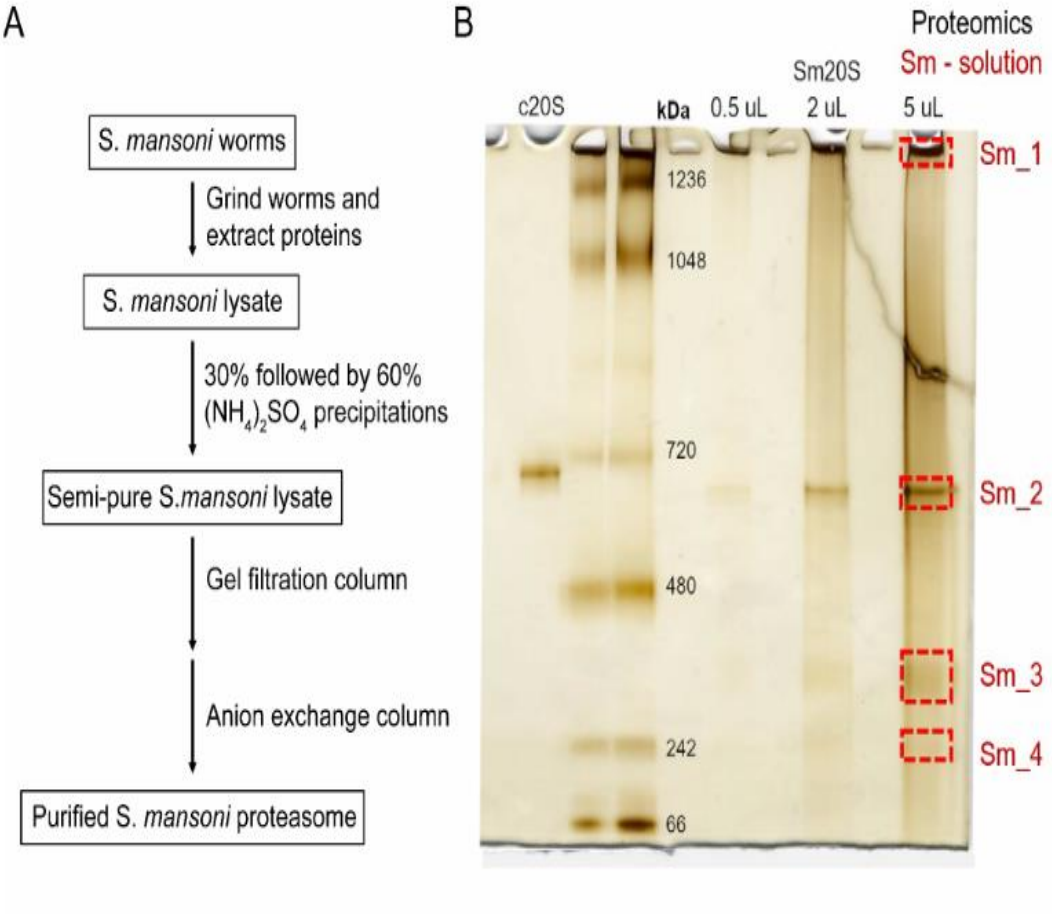
Natural marine products have been shown to produce promising compounds with diverse biological activities, such as antiviral, antifungal, antibacterial, and anti-parasitic. In the Tutuilamide study in collaboration with Dr. William H. Gerwick, cyanobacteria samples from American Samoa and Palmyra Atoll produced novel cyclic peptides: Tutuilamide A-C. After their structures were characterized by NMR and mass spectrometry techniques, I was able to test these novel compounds and their analogues against different families of serine protease. Tutuilamide A-C are novel cyanopeptolin-like peptides isolated from Cyanobacteria. Tutuilamide A and B showcased outstanding inhibition potency against pancreatic elastase and proteinase K. The linear form (Tutuilamide T1) was a poor inhibitor of both pancreatic elastase and proteinase K, suggesting that the cyclic structure of the depsipeptide is essential in maintaining inhibition against

serine protease. Taken together, the newly characterized compounds are potent inhibitors of pancreatic elastase and fungal proteinase K. Compared to previously studied elastase inhibitor lyngbyastatin 7, we have shown that new compounds have approximately 2 to 4-fold elevated inhibition.

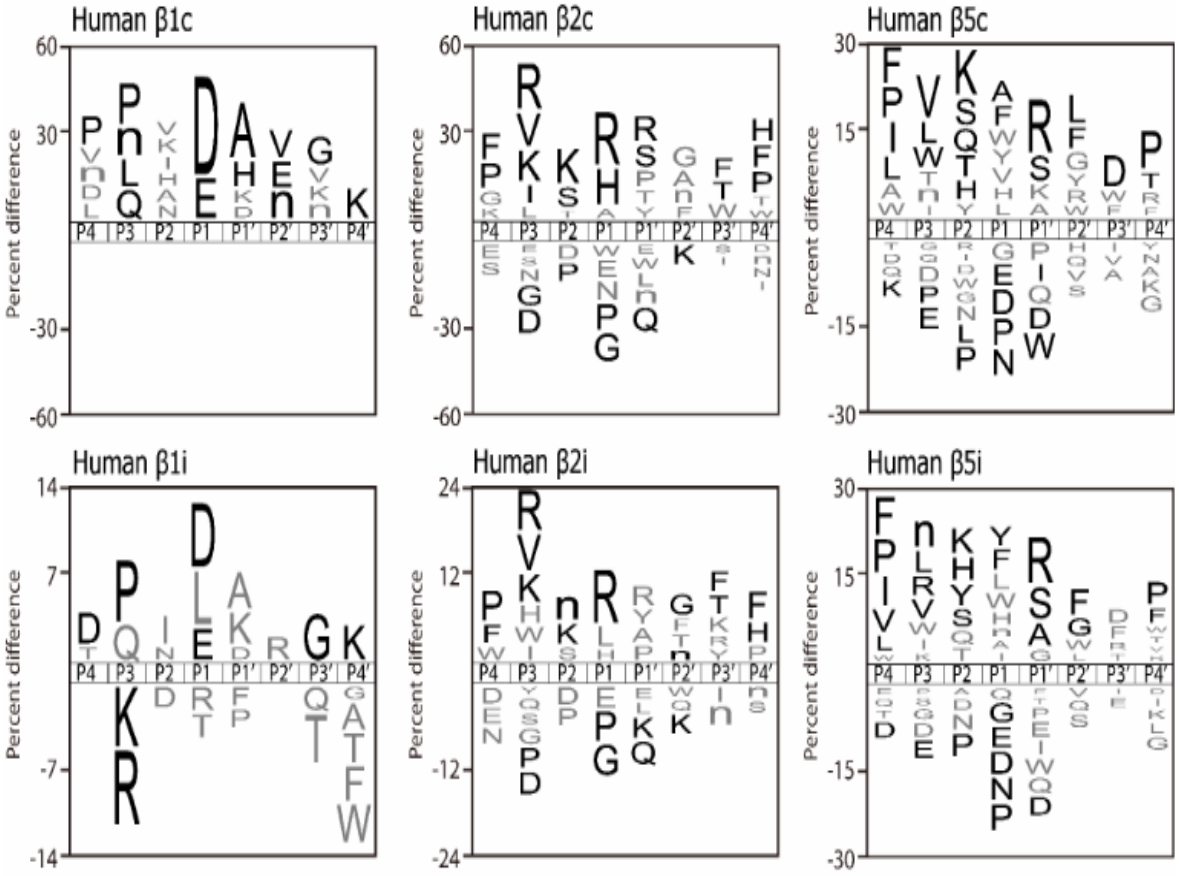
Conclusion

This thesis describes three projects in which protease plays a significant role. Utilizing a mass spectrometry-based assay, we developed novel fluorescent substrates for *S. manson*i proteasome based on its substrate specificities. I then evaluated their efficacies in 3 major *Schistosoma* species. Within *Schistosoma mansoni*, we also identified hnf4, a gene that is responsible for gut maintenance and successful pathology. Knocking out this gene with RNAi results in an 8.2-fold decrease in cysteine protease level in worm lysates, suggesting that this gene is essential in the parasite's cathepsin-mediated hemoglobin digestion. This thesis also describes Tutuilamides A-C, a series of potent pancreatic elastase and fungal proteinase K. These compounds showed increased potency compared to previously studied inhibitors of similar structures.

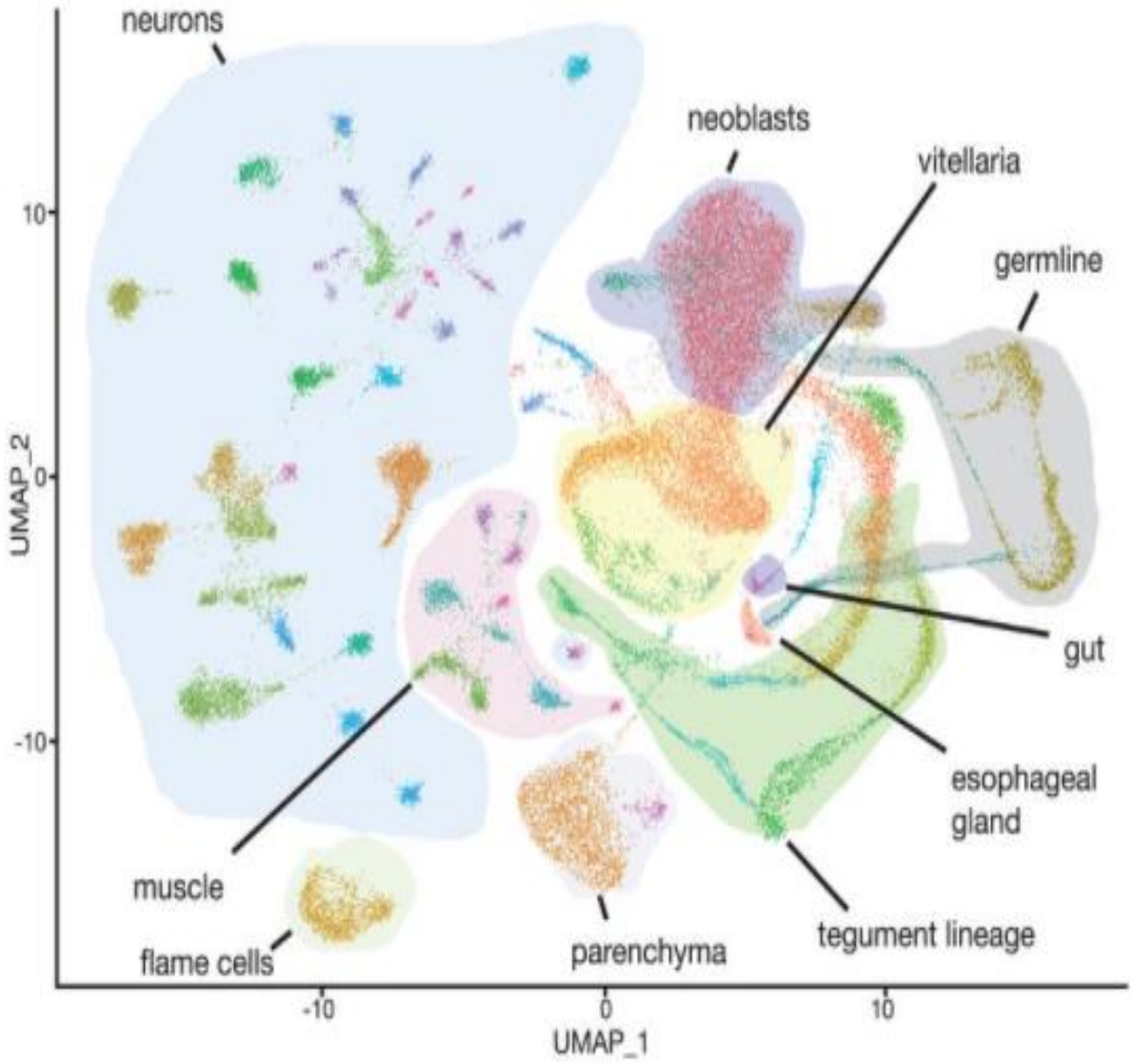
Appendix



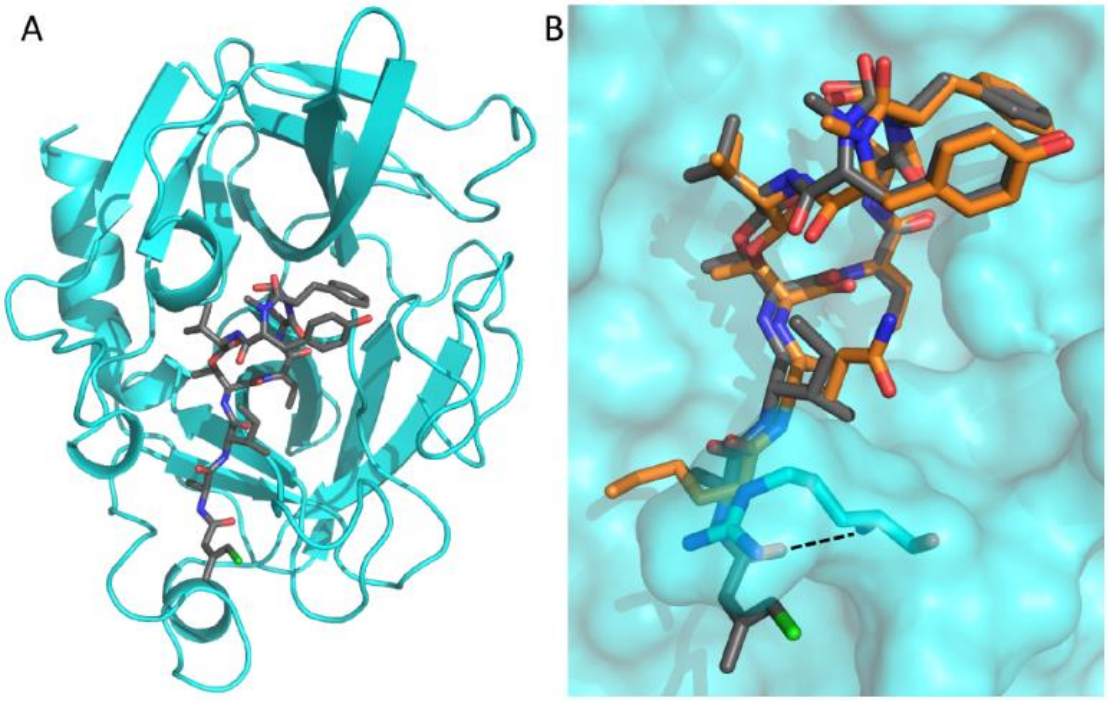
Supplemental Figure 1. Purification of Sm20S. A. Purification procedures. B. A native gel image of purified Sm20S sample.



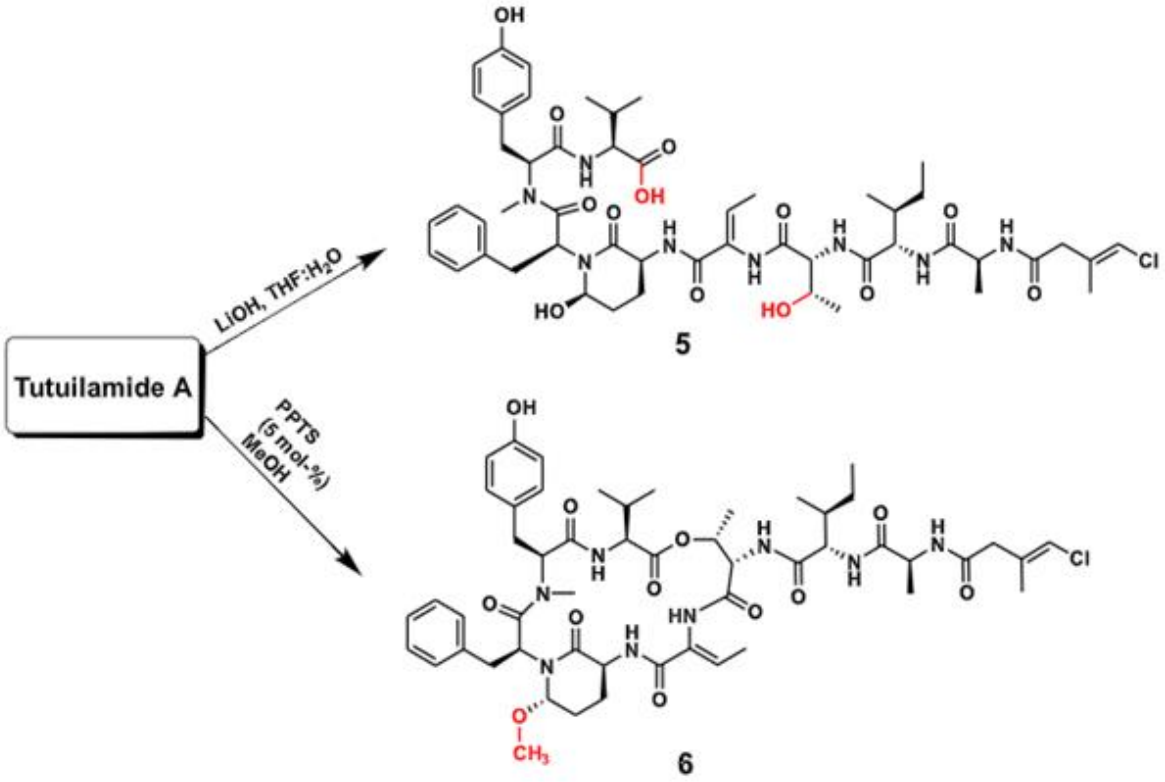
Supplemental Figure 2. Substrate specificity profiles of the human constitutive proteasome (C20S) and immunoproteasome (I20S) subunits.



Supplemental Figure 3. Uniform Manifold Approximation and Projection (UMAP) plot of the 68 scRNAseq clusters.



Supplementary Figure 4. Comparison of Tutuilamide A and Lyngbyastatin 7 binding to porcine pancreatic elastase. A Cartoon representation of PPE (cyan) in complex with Tutuilamide A (shown as grey sticks). **B.** Comparison of 1 (grey) and 4 (orange) binding to PPE (cyan surface). The side chains of PPE residues Ser216 and Arg217 are shown as sticks. A dashed line indicates the additional hydrogen bond found in the PPE-1 complex structure.



Supplemental Figure 5. Synthesis of analogues of Tutuilamide A: Tutuilamide T1 (linear) and Tutuilamide T2 (methyl)

Supplementary Table 1. Abbreviations of proteasome substrates.

	Sequence
β1-IQ1	Mca-EYADSK(DNP)-NH_2
β1-IQ2	Mca-VDQMDGWK(DNP)-NH_2
β1-AMC	SPLE-AMC
β2-IQ	Mca-FnKRRK(DNP)-OH
β5-AMC	FNKL-AMC

Supplementary Table 2. IC50 values of BTZ, CFZ with classical and novel substrates.

	BTZ	CFZ
LLE	40.04 ± 7.76 nM	> 9347 nM
β1-IQ	49.84 ± 0.24 nM	NA
β2-IQ	NA	288.0 ± 90 nM
LLVY	13.61 ± 2.54 nM	1.383 ± 0.23 nM
β5-AMC	17.97 ± 2.05 nM	2.495 ± 0.51 nM

References

1. LoVerde, P. T., Schistosomiasis. *Adv Exp Med Biol* **2019,** *1154*, 45-70.

2. Hotez, P. J.; Fenwick, A., Schistosomiasis in Africa: an emerging tragedy in our new global health decade. *PLoS Negl Trop Dis* **2009,** *3* (9), e485.

3. Colley, D. G.; Bustinduy, A. L.; Secor, W. E.; King, C. H., Human schistosomiasis. *Lancet* **2014,** *383* (9936), 2253-64.

4. Gray, D. J.; Ross, A. G.; Li, Y. S.; McManus, D. P., Diagnosis and management of schistosomiasis. *BMJ* **2011,** *342*, d2651.

5. Siqueira, L. D. P.; Fontes, D. A. F.; Aguilera, C. S. B.; Timoteo, T. R. R.; Angelos, M. A.; Silva, L.; de Melo, C. G.; Rolim, L. A.; da Silva, R. M. F.; Neto, P. J. R., Schistosomiasis: Drugs used and treatment strategies. *Acta Trop* **2017,** *176*, 179-187.

6. Cioli, D.; Pica-Mattoccia, L.; Basso, A.; Guidi, A., Schistosomiasis control: praziquantel forever? *Mol Biochem Parasitol* **2014,** *195* (1), 23-9.

7. Budenholzer, L.; Cheng, C. L.; Li, Y.; Hochstrasser, M., Proteasome Structure and Assembly. *J Mol Biol* **2017,** *429* (22), 3500-3524.

8. Gu, Z. C.; Enenkel, C., Proteasome assembly. *Cell Mol Life Sci* **2014,** *71* (24), 4729-45.

9. Bibo-Verdugo, B.; Jiang, Z.; Caffrey, C. R.; O'Donoghue, A. J., Targeting proteasomes in infectious organisms to combat disease. *FEBS J* **2017,** *284* (10), 1503-1517.

10. Groen, K.; van de Donk, N.; Stege, C.; Zweegman, S.; Nijhof, I. S., Carfilzomib for relapsed and refractory multiple myeloma. *Cancer Manag Res* **2019,** *11*, 2663-2675.

11. Zheng, Y.; Jiang, X.; Gao, F.; Song, J.; Sun, J.; Wang, L.; Sun, X.; Lu, Z.; Zhang, H., Identification of plant-derived natural products as potential inhibitors of the Mycobacterium tuberculosis proteasome. *BMC Complement Altern Med* **2014,** *14*, 400.

12. Bibo-Verdugo, B.; Wang, S. C.; Almaliti, J.; Ta, A. P.; Jiang, Z.; Wong, D. A.; Lietz, C. B.; Suzuki, B. M.; El-Sakkary, N.; Hook, V.; Salvesen, G. S.; Gerwick, W. H.; Caffrey, C. R.; O'Donoghue, A. J., The Proteasome as a Drug Target in the Metazoan Pathogen, Schistosoma mansoni. *ACS Infect Dis* **2019,** *5* (10), 1802-1812.

13. O'Donoghue, A. J.; Eroy-Reveles, A. A.; Knudsen, G. M.; Ingram, J.; Zhou, M.; Statnekov, J. B.; Greninger, A. L.; Hostetter, D. R.; Qu, G.; Maltby, D. A.; Anderson, M. O.; Derisi, J. L.; McKerrow, J. H.; Burlingame, A. L.; Craik, C. S., Global identification of peptidase specificity by multiplex substrate profiling. *Nat Methods* **2012,** *9* (11), 1095-100.

14. Lapek, J. D., Jr.; Jiang, Z.; Wozniak, J. M.; Arutyunova, E.; Wang, S. C.; Lemieux, M.

J.; Gonzalez, D. J.; O'Donoghue, A. J., Quantitative Multiplex Substrate Profiling of Peptidases by Mass Spectrometry. *Mol Cell Proteomics* **2019,** *18* (5), 968-981.

15. Beekman, C.; Jiang, Z.; Suzuki, B. M.; Palmer, J. M.; Lindner, D. L.; O'Donoghue, A. J.; Knudsen, G. M.; Bennett, R. J., Characterization of PdCP1, a serine carboxypeptidase from Pseudogymnoascus destructans, the causal agent of White-nose Syndrome. *Biol Chem* **2018,** *399* (12), 1375-1388.

16. Dvorak, J.; Fajtova, P.; Ulrychova, L.; Leontovyc, A.; Rojo-Arreola, L.; Suzuki, B. M.; Horn, M.; Mares, M.; Craik, C. S.; Caffrey, C. R.; O'Donoghue, A. J., Excretion/secretion products from Schistosoma mansoni adults, eggs and schistosomula have unique peptidase specificity profiles. *Biochimie* **2016,** *122*, 99-109.

17. Papalexi, E.; Satija, R., Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* **2018,** *18* (1), 35-45.

18. Luecken, M. D.; Theis, F. J., Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* **2019,** *15* (6), e8746.

19. Wendt, G.; Zhao, L.; Chen, R.; Liu, C.; O'Donoghue, A. J.; Caffrey, C. R.; Reese, M. L.; Collins, J. J., 3rd, A single-cell RNA-seq atlas of Schistosoma mansoni identifies a key regulator of blood feeding. *Science* **2020,** *369* (6511), 1644-1649.

20. Delcroix, M.; Sajid, M.; Caffrey, C. R.; Lim, K. C.; Dvorak, J.; Hsieh, I.; Bahgat, M.; Dissous, C.; McKerrow, J. H., A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *J Biol Chem* **2006,** *281* (51), 39316-29.

21. Caffrey, C. R.; Goupil, L.; Rebello, K. M.; Dalton, J. P.; Smith, D., Cysteine proteases as digestive enzymes in parasitic helminths. *PLoS Negl Trop Dis* **2018,** *12* (8), e0005840.

22. Jedlickova, L.; Dvorakova, H.; Dvorak, J.; Kasny, M.; Ulrychova, L.; Vorel, J.; Zarsky, V.; Mikes, L., Cysteine peptidases of Eudiplozoon nipponicum: a broad repertoire of structurally assorted cathepsins L in contrast to the scarcity of cathepsins B in an invasive species of haematophagous monogenean of common carp. *Parasit Vectors* **2018,** *11* (1), 142.

23. Withana, N. P.; Blum, G.; Sameni, M.; Slaney, C.; Anbalagan, A.; Olive, M. B.; Bidwell, B. N.; Edgington, L.; Wang, L.; Moin, K.; Sloane, B. F.; Anderson, R. L.; Bogyo, M. S.; Parker, B. S., Cathepsin B inhibition limits bone metastasis in breast cancer. *Cancer Res* **2012,** *72* (5), 1199-209.

24. Yasuda, Y.; Kageyama, T.; Akamine, A.; Shibata, M.; Kominami, E.; Uchiyama, Y.; Yamamoto, K., Characterization of new fluorogenic substrates for the rapid and sensitive assay of cathepsin E and cathepsin D. *J Biochem* **1999,** *125* (6), 1137-43.

25. Chen, D.; Chai, J.; Hart, P. J.; Zhong, G., Identifying catalytic residues in CPAF, a Chlamydia-secreted protease. *Arch Biochem Biophys* **2009,** *485* (1), 16-23.

26. Welker, M.; von Dohren, H., Cyanobacterial peptides - nature's own combinatorial biosynthesis. *FEMS Microbiol Rev* **2006,** *30* (4), 530-63.

27. Keller, L.; Canuto, K. M.; Liu, C.; Suzuki, B. M.; Almaliti, J.; Sikandar, A.; Naman, C. B.; Glukhov, E.; Luo, D.; Duggan, B. M.; Luesch, H.; Koehnke, J.; O'Donoghue, A. J.; Gerwick, W. H., Tutuilamides A-C: Vinyl-Chloride-Containing Cyclodepsipeptides from Marine Cyanobacteria with Potent Elastase Inhibitory Properties. *ACS Chem Biol* **2020,** *15* (3), 751-757.

28. Martin, C.; Oberer, L.; Ino, T.; Konig, W. A.; Busch, M.; Weckesser, J., Cyanopeptolins, new depsipeptides from the cyanobacterium Microcystis sp. PCC 7806. *J Antibiot (Tokyo)* **1993,** *46* (10), 1550-6.

29. Kang, H. S.; Krunic, A.; Orjala, J., Stigonemapeptin, an Ahp-containing depsipeptide with elastase inhibitory activity from the bloom-forming freshwater cyanobacterium Stigonema sp. *J Nat Prod* **2012,** *75* (4), 807-11.

30. Taori, K.; Matthew, S.; Rocca, J. R.; Paul, V. J.; Luesch, H., Lyngbyastatins 5-7, potent elastase inhibitors from Floridian marine cyanobacteria, Lyngbya spp. *J Nat Prod* **2007,** *70* (10), 1593-600.

31. Harrigan, G. G.; Luesch, H.; Yoshida, W. Y.; Moore, R. E.; Nagle, D. G.; Paul, V. J., Symplostatin 2: a dolastatin 13 analogue from the marine cyanobacterium Symploca hydnoides. *J Nat Prod* **1999,** *62* (4), 655-8.

32. Liu, L.; Rein, K. S., New peptides isolated from Lyngbya species: a review. *Mar Drugs* **2010,** *8* (6), 1817-37.

33. Salvador, L. A.; Taori, K.; Biggs, J. S.; Jakoncic, J.; Ostrov, D. A.; Paul, V. J.; Luesch, H., Potent elastase inhibitors from cyanobacteria: structural basis and mechanisms mediating cytoprotective and anti-inflammatory effects in bronchial epithelial cells. *J Med Chem* **2013,** *56* (3), 1276-90.

34. Matern, U.; Schleberger, C.; Jelakovic, S.; Weckesser, J.; Schulz, G. E., Binding structure of elastase inhibitor scyptolin A. *Chem Biol* **2003,** *10* (10), 997-1001.

35. Abdulla, M. H.; Lim, K. C.; Sajid, M.; McKerrow, J. H.; Caffrey, C. R., Schistosomiasis mansoni: novel chemotherapy using a cysteine protease inhibitor. *PLoS Med* **2007,** *4* (1), e14.

36. Abdulla, M. H.; Ruelas, D. S.; Wolff, B.; Snedecor, J.; Lim, K. C.; Xu, F.; Renslo, A. R.; Williams, J.; McKerrow, J. H.; Caffrey, C. R., Drug discovery for schistosomiasis: hit and lead compounds identified in a library of known drugs by medium-throughput phenotypic screening. *PLoS Negl Trop Dis* **2009,** *3* (7), e478.

37. Colley, D. G.; Wikel, S. K., Schistosoma mansoni: simplified method for the production of schistosomules. *Exp Parasitol* **1974,** *35* (1), 44-51.

38. Basch, P. F., Cultivation of Schistosoma mansoni in vitro. I. Establishment of cultures from cercariae and development until pairing. *J Parasitol* **1981,** *67* (2), 179-85.

39. Collins, J. J., 3rd; Wendt, G. R.; Iyer, H.; Newmark, P. A., Stem cell progeny contribute to the schistosome host-parasite interface. *Elife* **2016,** *5*, e12473.

40. Collins, J. J., 3rd; Wang, B.; Lambrus, B. G.; Tharp, M. E.; Iyer, H.; Newmark, P. A., Adult somatic stem cells in the human parasite Schistosoma mansoni. *Nature* **2013,** *494* (7438), 476-9.

41. Wendt, G. R.; Collins, J. N.; Pei, J.; Pearson, M. S.; Bennett, H. M.; Loukas, A.; Berriman, M.; Grishin, N. V.; Collins, J. J., 3rd, Flatworm-specific transcriptional regulators promote the specification of tegumental progenitors in Schistosoma mansoni. *Elife* **2018,** *7*.

42. Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013,** *29* (1), 15-21.

43. Howe, K. L.; Bolt, B. J.; Cain, S.; Chan, J.; Chen, W. J.; Davis, P.; Done, J.; Down, T.; Gao, S.; Grove, C.; Harris, T. W.; Kishore, R.; Lee, R.; Lomax, J.; Li, Y.; Muller, H. M.; Nakamura, C.; Nuin, P.; Paulini, M.; Raciti, D.; Schindelman, G.; Stanley, E.; Tuli, M. A.; Van Auken, K.; Wang, D.; Wang, X.; Williams, G.; Wright, A.; Yook, K.; Berriman, M.; Kersey, P.; Schedl, T.; Stein, L.; Sternberg, P. W., WormBase 2016: expanding to enable helminth genomic research. *Nucleic Acids Res* **2016,** *44* (D1), D774-80.

44. Love, M. I.; Huber, W.; Anders, S., Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **2014,** *15* (12), 550.

45. LaMonte, G. M.; Almaliti, J.; Bibo-Verdugo, B.; Keller, L.; Zou, B. Y.; Yang, J.; Antonova-Koch, Y.; Orjuela-Sanchez, P.; Boyle, C. A.; Vigil, E.; Wang, L.; Goldgof, G. M.; Gerwick, L.; O'Donoghue, A. J.; Winzeler, E. A.; Gerwick, W. H.; Ottilie, S., Development of a Potent Inhibitor of the Plasmodium Proteasome with Reduced Mammalian Toxicity. *J Med Chem* **2017,** *60* (15), 6721-6732.

46. Narayanan, S.; Cai, C. Y.; Assaraf, Y. G.; Guo, H. Q.; Cui, Q.; Wei, L.; Huang, J. J.; Ashby, C. R., Jr.; Chen, Z. S., Targeting the ubiquitin-proteasome pathway to overcome anti-cancer drug resistance. *Drug Resist Updat* **2020,** *48*, 100663.

47. van Wolfswinkel, J. C.; Wagner, D. E.; Reddien, P. W., Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell* **2014,** *15* (3), 326-339.

48. Skelly, P. J.; Da'dara, A. A.; Li, X. H.; Castro-Borges, W.; Wilson, R. A., Schistosome feeding and regurgitation. *PLoS Pathog* **2014,** *10* (8), e1004246.

49. Wasilewski, M. M.; Lim, K. C.; Phillips, J.; McKerrow, J. H., Cysteine protease inhibitors block schistosome hemoglobin degradation in vitro and decrease worm burden and egg production in vivo. *Mol Biochem Parasitol* **1996,** *81* (2), 179-89.