# Discovering New Classes of *Brugia malayi* Asparaginyl-tRNA Synthetase Inhibitors and Relating Specificity to Conformational Change

Sai Chetan K. Sukuru[1,2], Thibaut Crepin[4], Youli Milev[3], Liesl C. Marsh[5], Jonathan B. Hill[5], Regan J. Anderson[5], Jonathan C. Morris[6], Anjali Rohatgi[1], Gavin O'Mahony[7], Morten Grøtli[7], Franck Danel[8], Malcolm G.P. Page[8], Michael Härtlein[9], Stephen Cusack[4], Michael A. Kron[3,†] and Leslie A. Kuhn[1,2]

[1]Department of Biochemistry and Molecular Biology, [2]Quantitative Biology and Modeling Initiative, and [3]Department of Medicine and Institute of International Health, Michigan State University, East Lansing, MI 48824, USA.

[4]European Molecular Biology Laboratory, Grenoble Outstation, Grenoble, France.

[5]Department of Chemistry, University of Canterbury, Christchurch, New Zealand.

[6]School of Chemistry and Physics, The University of Adelaide, Adelaide, Australia.

[7]Department of Chemistry, Göteborg University, Göteborg, Sweden.

[8]Basilea Pharmaceutica Limited, Basel, Switzerland.

[9]Large Scale Structures Group, Institute Laue-Langevin, Grenoble, France.

[†]Current address: Department of Medicine and International Health Program, Biotechnology and Bioengineering Center, Medical College of Wisconsin, Milwaukee, WI 53226.

*Corresponding Author*: Leslie A. Kuhn, 502C Biochemistry Building, Michigan State University, East Lansing, MI 48824, USA; TEL: +1-517-353-8745; FAX: +1-517-353-9334; URL: http://www.bch.msu.edu/labs/kuhn

## Abstract

SLIDE, which models the flexibility of protein and ligand side chains while docking, was used to screen several large databases to identify inhibitors of *Brugia malayi* asparaginyl-tRNA synthetase (AsnRS), a target for anti-parasitic drug design. Seven classes of compounds identified by SLIDE were confirmed as having micromolar inhibition constants against the enzyme. Analogs of one of these classes of inhibitors, the long side-chain variolins, cannot bind to the adenosyl pocket of the closed conformation of AsnRS due to steric clashes, though the short side-chain variolins identified by SLIDE apparently bind isosterically with adenosine. We hypothesized that an open conformation of the motif 2 loop also permits the long side-chain variolins to bind in the adenosine pocket and that their selectivity for *Brugia* relative to human AsnRS can be explained by differences in the sequence and conformation of this loop. Loop flexibility sampling using ROCK confirms this possibility, while scoring of the relative affinities of the different ligands by SLIDE correlates well with the compounds' ranks in inhibition assays. Combining ROCK and SLIDE provides a promising approach for exploiting conformational flexibility in structure-based screening and design of species selective inhibitors.

## Keywords

Conformational flexibility, docking, inhibitor design, sampling, scoring, screening, anti-parasitic agents, MSU ProFlex, ROCK, SLIDE

## Abbreviations

AARS – aminoacyl-tRNA synthetases, ASNAMS – asparagine sulfamoyl adenylate, AsnRS – asparaginyl-tRNA synthetase, CSD – Cambridge Structural Database, LBHAMP – L-aspartate-$\beta$-hydroxamate adenylate, MD – molecular dynamics, PDB – Protein Data Bank, SMEVAR – S-methyl-deoxyvariolin B

# Introduction

Lymphatic filariasis, also known as elephantiasis, is caused by the nematode worms *Wuchereria bancrofti* and *Brugia malayi*. It is a debilitating human disease that afflicts more than 200 million people worldwide. More than 1.2 billion people in 80 countries reside in areas where the disease is actively transmitted and are at a great risk of contracting the disease [1-3]. The crippling physical effects of the disease have a huge economic and social impact, which is why lymphatic filariasis is one of the top 10 tropical diseases being targeted by the World Health Organization. Strategies to control the disease include administration of drugs like ivermectin, diethylcarbamazine and albendazole, which reduce the level of infection and prevent transmission. Most of these drugs were discovered decades ago as chemotherapeutic agents to combat human filariasis. However, the use of these drugs has been plagued with concerns about their inability to kill the adult worms even after long treatment durations, severe side effects, and the emergence of drug resistance in humans [4, 5].

Aminoacyl-tRNA synthetases (AARS) have been acknowledged as rational targets for anti-infective drug development [6] because these enzymes are essential for viability. AARS are one of several new drug targets in human filarial parasites that have been proposed in recent years [1]. They differ significantly in sequence and structure between the parasite and host organism, although sharing a common catalytic site topology. AARS are responsible for the specific aminoacylation of transfer RNAs (tRNAs). The two-step catalytic reaction involves the ATP-based activation of the amino acid and the transfer of activated amino acid to the 3'-end of the cognate tRNA. The 20 AARS (one for each amino acid) are divided into two classes on the basis of their active site architecture and conserved sequence identity. The class I AARS possess two signature amino-acid sequences (HIGH and KMSKS) located in the active site with its characteristic nucleotide-binding Rossmann fold,

which consists of alternating pattern of β-strands and α-helices [7] . The active site of class II AARS is comprised of a six-stranded antiparallel β-sheet flanked by an additional parallel strand and 3 α-helices [8]. Three signature sequence motifs – motifs 1, 2 and 3 – characterize this class (Figure 1B). Motif 1 consists of: +G(F/Y)XX(V/L/I)PΦΦ, where + is a positively charged residue, Φ is a hydrophobic residue and X is any residue. Motif 2 consists of: +ΦΦXΦXXXFRxE. Motif 3 consists of: ΦGΦGΦGΦΦERΦΦΦΦ. Several exceptions to these classes have been reported [9]. It is possible to further divide class II AARS into subclasses IIa, IIb, and IIc based on the presence of specific domains that play a role in anticodon recognition. Given the sequence and structural differences between the prokaryotic and eukaryotic AARS, and that they are essential for the viability of all organisms, AARS make attractive targets for developing selective inhibitors [6, 10]. For example, Pseudomonic acid (mupirocin), a natural product synthesized by *Pseudomonas fluorescens*, inhibits isoleucyl-tRNA synthetase (IleRS) from Gram-positive infectious bacteria, including antibiotic-resistant *S. aureus* [11]. This molecule has been shown to have almost 8000-fold selectivity for pathogen IleRS over mammalian IleRS [12].

We are targeting the *Brugia malayi* asparaginyl tRNA synthetase (AsnRS) for drug development against filariasis because it is an essential enzyme in protein synthesis that is expressed in both sexes of the nematode and in several stages of the life cycle - adults, bloodborne microfilariae, and infective larvae [13]. Recent studies have shown that the expression levels of AsnRS in *Brugia* females are significantly higher than those of other AARS [14]. AsnRS is also specifically associated with chemokine activity towards human cells that may play a role in the massive inflammatory response associated with lymphatic filariasis [15]. AsnRS has been well characterized biochemically and structurally and can be recombinantly expressed to facilitate *in vitro* studies[16, 17]. AsnRS is a class IIb aminoacyl tRNA synthetase, together with AspRS and LysRS. Class IIb

AARS have a distinct N-terminal-beta-barrel domain (OB fold) that binds the tRNA anticodon stem loop.

Here, we present results of applying structure-based computational ligand screening and design to develop inhibitors against *Brugia* AsnRS, using a 1.9 Å resolution structure of the enzyme in complex with a non-hydrolyzable analog of asparaginyl adenylate, which likely mimics the enzymatic product. The homodimeric structure (Figure 1A) is N-terminally truncated, with each monomer lacking the first 111 amino-acid residues to enable crystallization of the complex; this truncation, however, retains catalytic activity. Two more crystal structures of the enzyme, one a dimer with the monomers bound to two different ligands and another a ligand-free dimer, further aided our analysis. Structure-based drug design [18] has led to the development of potent and specific new drugs, such as the widely used HIV protease inhibitor, Viracept. Computational screening is useful for identifying new leads for drug design as well as narrowing down the search domain and focusing *in vitro* screening toward appropriate, often novel molecular scaffolds. Shoichet and co-workers have shown that computational screening using molecular docking for identifying new scaffolds typically has a significantly higher hit rate than *in vitro* high-throughput screening alone [19].

The efficiency of structure-based screening methods depends on the speed with which they eliminate infeasible ligand candidates and how accurately they predict the binding modes and affinities of the docked ligands. Both protein and ligand flexibility need to be incorporated in computational drug design to improve the accuracy of the results [20, 21]. Modeling protein flexibility using representative protein conformations to screen against each conformer of each ligand candidate is computationally very expensive, requiring some compromise between speed and

accuracy. A number of methods have been developed in recent years to include such flexibility in drug design. Kuntz and co-workers averaged the information from multiple crystallographic and NMR structures of the same protein to describe its conformational variability [22], and the developers of AutoDock have used a similar strategy [23]. FlexE software [24] incorporates protein flexibility by retaining the varied orientations of flexible side chains of multiple crystal structures of the same protein and using compatible combinations of these conformations to dock ligands. Carlson and co-workers used several snapshots from molecular dynamics (MD) simulations to map the conserved, relatively immobile interaction sites in the dynamic protein binding site [25]. However, these methods are either limited to experimentally solved structures to explore the available conformational space for proteins or focus on the rigid regions within the protein.

SLIDE (Screening for Ligands by Induced-fit Docking, Efficiently) accommodates protein and ligand side-chain flexibility when screening databases of hundreds of thousands of small organic molecules to identify potential ligand candidates for a target protein [20, 26]. Knowledge-based representation of the protein binding site in SLIDE gives good sampling and identifies the correct binding modes of ligands [27, 28]. Steric misfit between the docked ligand and the protein is resolved in SLIDE through the minimal directed rotation of single bonds in the ligand and in the protein side chains [29]. Ligand candidates are assumed to be in a conformation close to the bioactive, bound conformation, or are input as libraries of low-energy conformers. Here we employ flexibility modeling to identify and model the interactions of new ligands for *Brugia* AsnRS using SLIDE screening of databases of small organic molecules with drug-like molecular weights and atomic compositions. The observed affinities and specificities of known and predicted ligands for *Brugia* AsnRS are compared with the structure-based predictions. We also present a case in which large-scale conformational change of an active-site loop must (and can) be modeled in order to

predict the mode of ligand binding and explain its reasonable selectivity for *Brugia* relative to human AsnRS.

The graph theoretic algorithm ProFlex, successor to the FIRST software [30], was used to identify the coupled networks of covalent and non-covalent bonds within the target protein to predict the flexible regions in *Brugia* AsnRS. Diverse conformers of these flexible regions were generated using ROCK, a random-walk sampling algorithm [31, 32]. In order to assess the contributions of the residues of the flexible loop regions to specificity and relative binding affinities of ligands, detailed docking studies were performed. Flexible low-energy conformers generated by Omega (OpenEye Software) for the top-scoring ligand candidates were docked into the most open ROCK conformer of the protein using SLIDE, and SLIDE modeled additional side-chain flexibility in the ligand and protein upon binding. This approach has provided insights into the role of main-chain flexibility in ligand binding for other systems (cyclophilin A, estrogen receptor, dihydrofolate reductase, and HIV protease) [31, 32], and, in the case of *Brugia* AsnRS, helps identify the structural elements that determine ligand specificity and binding for long side-chain variolins.

## Materials and Methods

### Asparaginyl-tRNA Synthetase Structures

A 1.9 Å resolution closed structure of *Brugia* AsnRS in complex with a non-hydrolyzable analog of asparaginyl adenylate (ASNAMS) was used for structure-based ligand screening and design (Figure 1A and Table 1). Additionally, two crystal structures (Table 1) providing the ligand-free (apo) conformation, and a structure with one monomer bound to ATP and the other bound to L-aspartate-β-hydroxamate adenylate (LBHAMP) were used for analyzing the conformational flexibility of the protein. Data collection and refinement statistics are provided in Table 1. Details of crystallization and structure determination will be presented elsewhere, and X-ray coordinates will be deposited in the Protein Data Bank [33]. In the meantime, X-ray coordinates can be obtained by contacting Stephen Cusack (cusack@embl-grenoble.fr).

### Experimental Assay

We standardized the malachite green assay for phosphate release [42-45] for use in monitoring inhibition of aminoacylation using colorimetric measurement of pyrophosphate generation from the first step in the aminoacylation reaction:

$$E + AA + ATP \longleftrightarrow E(AA\text{-}AMP) + PP_i \qquad (1)$$

$$E(AA\text{-}AMP) + tRNA \longleftrightarrow E + AA\text{-}tRNA + AMP \qquad (2)$$

This assay was used to measure inhibition constants for inhibitors predicted by SLIDE and their analogs.

**Screening and Docking with SLIDE**

SLIDE (Screening for Ligands by Induced-Fit Docking Efficiently) was used to screen databases of small organic molecules to find potential inhibitors of *Brugia* AsnRS. SLIDE [20, 26, 34] is a screening and docking tool that uses distance geometry to screen and dock ligand candidates into the binding site of the target protein. SLIDE represents the binding site of the protein by a template consisting of points identified as the most favorable positions for ligand atoms to form hydrogen bonds or make hydrophobic interactions with the neighboring protein atoms [27]. The ligand candidates in the database are similarly represented by a set of interaction points, assigned to polar atoms or centers of hydrophobic atom clusters. For each ligand candidate, all possible triplets of its interaction points are mapped onto all geometrically and chemically compatible template triangles. The anchor fragment of the ligand is defined by the triplet of interaction points that match with a template triangle. Any portion of the ligand outside this anchor fragment is considered flexible by SLIDE. After finding a feasible match between the ligand candidate and protein template, SLIDE models induced-fit by resolving steric overlaps between the flexible portion of the ligand and the protein side chains using minimal rotations determined by mean-field optimization [26, 29] . Collision-free docked ligand orientations are scored based on the number of hydrogen bonds and degree of hydrophobic complementarity with the protein.   The SLIDE software is available to academic and commercial researchers; see Software at http://www.bch.msu.edu/labs/kuhn

**Scoring Protein-Ligand Interactions**

Several comparative studies of docking and scoring methods [28,35-39] have shown that no one scoring function for predicting ligand binding and affinity performs consistently well across diverse protein families. Hence, to develop a scoring protocol that can distinguish ligands from non-ligands, reliably detect the correct conformation and binding mode for known ligands and score them in the

order of their relative affinity for *Brugia* AsnRS, a panel of three different scoring functions was tested: SLIDE score [27], DrugScore [40], and X-Score [41]. SLIDE score is a weighted sum of hydrophobic and hydrogen-bond interaction terms, trained to match affinity values in known complexes. DrugScore is a knowledge-based scoring function that uses structural information from the Protein Data Bank (PDB) to score protein-ligand complexes based on the preferred distances observed between different ligand and protein atom pairs. X-Score is an empirical scoring function that calculates the binding affinity of a protein-ligand complex by using terms that account for van der Waals interactions, hydrogen bonding, deformation, and the hydrophobic effect. Scoring accuracy was determined by how well the scoring functions assessed the binding modes and relative affinities of known *Brugia* AsnRS ligands, whereas the enrichment accuracy was determined by their ability to select true ligands from a large number of decoys (1000 diverse, random drug-like molecules obtained from the website of Dr. Didier Rognan, CNRS: http://bioinfo-pharma.u-strasbg.fr/bioinformatics-cheminformatics-group.html). Low-energy conformers of these decoy molecules were generated using Omega (OpenEye Software) as input to SLIDE screening. All conformers within 7.5 Kcal/mol of the minimum energy conformer sampled (using the MMFF force field) were included. For each docked compound (known ligands as well as non-ligand decoys), only the top scoring binding orientation was ranked. Any two docked compounds scoring identically were given identical ranks.

**Modeling Main-chain Flexibility**

The active sites of human and *Brugia* AsnRS are very similar, with only 3 amino acid differences in the first shell of amino acids surrounding the active site, including all residues within 9Å of any atom of ASNAMS (Figure 2A). Because these three side chains point away from the binding site, it is most likely that they influence the conformations of residues that interact with asparagine or

adenosine or the conformations of the motif 2 adenine-binding loop, which has residue 224 (Ala in *Brugia* and Thr in human) near its hinge. To examine alternate conformations accessible to the known flexible active-site loops in *Brugia* AsnRS, we used ProFlex software [30] to identify the flexible regions in the protein, and ROCK [31, 32] to sample them. ProFlex predicts the flexible and rigid regions in a given structure (which bonds are constrained and which bonds remain free to rotate) based on analysis of constraints posed by the protein's network of covalent bonds, hydrogen bonds, salt bridges, and hydrophobic interactions. ProFlex calculations are fast and have been shown to predict the conformational flexibility of a protein reliably from a single 3D structure [46, 47]. ROCK (Rigidity Optimized Conformational Kinetics) uses a restricted random-walk sampling to search the conformational space available to proteins given the flexible regions defined by ProFlex as input. A conformer generated by ROCK is either accepted or rejected, depending upon whether it maintains the non-covalent bond network and results in no van der Waals overlaps between atoms. The most distinct main-chain conformers generated by ROCK are selected based on the RMSD values relative to the initial structure. *Brugia* AsnRS active-site loop conformers representing favorable open conformations of the protein were used to interpret the observed affinities and specificities for *Brugia* AsnRS relative to human AsnRS for inhibitors that could not bind (based on steric clashes) with the closed loop conformation of *Brugia* AsnRS. ProFlex and ROCK software are available to academic and commercial researchers; see Software under http://www.bch.msu.edu/labs/kuhn.

## Results

**Scoring *Brugia* AsnRS-Ligand Complexes**

The results of the scoring analysis by SLIDE score, DrugScore and X-Score (Figure 3) show that both SLIDE score and DrugScore do a reliable job of assessing the right conformation, binding mode, and relative affinity of known AsnRS ligands, and also distinguish three known ligands from 1000 diverse drug-like molecules. The first of these known ligands discovered for *Brugia malayi* independently of SLIDE, with an experimentally characterized binding mode, is ASNAMS, which was designed as a product analog for step 1 in the aminoacylation reaction and had previously been shown to bind in the crystal structure of *Thermus thermophilus* AsnRS [8]. The structure of Complex I (Figure 1) shows that this compound binds in the same orientation in *Brugia malayi* AsnRS, with an IC50 value of 4.5 μM, as measured by the malachite green assay (Table 1). L-aspartate-β-hydroxamate was used as a reagent during development of the malachite green assay [45]. When the third known ligand, the substrate ATP, was added during the assay, L-aspartate-β-hydroxamate adenylate (LBHAMP) was formed and remained bound to AsnRS, as confirmed by the structure of Complex II (Table 1). This complex contains LBHAMP and pyrophosphate bound to one of the monomers in the dimer, whereas ATP alone was observed in the other monomer. The IC50 value of LBHAMP is 4 μM (Table 2), very similar to that of the product analog ASNAMS. The three known ligands, ASNAMS, LBHAMP, and ATP, were all docked by SLIDE to within 1 Å RMSD of their crystallographically observed positions, and the scoring functions ranked them correctly according to their experimentally determined IC50 values against *Brugia* AsnRS; adenosine is not observed to inhibit AsnRS at a concentration of 500 μM. The enrichment plot in Figure 3 shows that SLIDE score performs particularly well in ranking the three *Brugia* AsnRS known ligands within the top 5 scoring compounds and clearly distinguishes them from the vast majority of the molecules used as non-ligand decoys.

12

**Screening the Databases**

A template was generated by SLIDE to represent the active site of *Brugia* AsnRS and its interactions with ASNAMS and used to screen databases of small organic molecules. The Cambridge Structural Database (CSD) [48] and the National Cancer Institute (NCI) Plated Compounds Database [49] were filtered to retain compounds with appropriate atom types (no bound metals or inorganic atoms other than halogens), molecular weights ($\leq$ 500Da), lipophilicity values (logP $\leq$ 5) , flexibility ($\leq$ 5 rotatable bonds) and polar character ($\leq$ 5 hydrogen bond donor atoms and $\leq$ 10 hydrogen bond acceptor atoms), to focus screening on the most drug-like compounds [50]. The final set of compounds for screening included about 110,000 compounds from the CSD and about 78,000 compounds from the NCI Plated Compounds Database. The CSD contains at least one low-energy crystallographic conformation for each of its compounds, and these conformations were used for screening. For the NCI compounds, low-energy 3-dimensional conformers were generated using Omega (OpenEye Software). All conformers within 7.5 Kcal/mol of the minimum energy conformer sampled (using the MMFF force field) were included. Both SLIDE score and DrugScore were used to score the docked orientations of the potential ligand candidates screened by SLIDE. Compounds were selected for experimental assays based on these scores and molecular graphics inspection of their interactions with *Brugia* AsnRS.  Out of the high-scoring candidates, we selected for assays those having several of the following desirable features: matching the hydrogen-bond and pi-cation interactions of known AsnRS ligands, filling the same volume as the product analog ASNAMS in the adenine and ribose pockets, having amine or halogen groups that would lend to ready substitution, and having no known tendency to self-assemble at typical inhibitor concentrations.  Upon assaying, the following compounds were found to inhibit *Brugia* AsnRS significantly at concentrations in the low to mid-micromolar range.

## Results from Screening the CSD

**Variolin B**

Variolin B (CSD code LEPWIM), a pyrrolopyrimidine, was originally isolated from an Antarctic sea sponge and has been shown to have antitumor and antiviral activity [51]. In earlier work, SLIDE identified variolin B as a potential inhibitor [3], and it was confirmed to inhibit ~50% (47%±16%) of *Brugia* AsnRS activity at a concentration of 50μM. This marine natural product contains five- and six-membered rings that are isosteric and share chemistry with adenine, suggesting that it could bind similarly to adenosine to the AsnRS structure. Indeed, SLIDE indicates (Figure 4A) that it binds in the same pocket as the adenosyl portion of ASNAMS in the *Brugia* AsnRS crystallographic complex [3]. Pyrrolopyrimidines and their analogs have been shown to compete with ATP to inhibit cyclin dependent kinases [52, 53]. Here, as a follow-up study, three available, synthesized derivatives of variolin B [54-56] were tested for *Brugia* and human AsnRS inhibition; their IC50 values appear in Table 4. One of these derivatives, SMEVAR, shows significant inhibition of AsnRS at 50 μM concentration, whereas LCM01 and LCM02 [57-59] showed somewhat weaker (125-175 μM) IC50 values, with the advantage of 3- to 8-fold selectivity for *Brugia* over human AsnRS.


Although variolin B and its derivatives show promising inhibition of *Brugia* AsnRS enzymatic activity, they are highly cytotoxic in human Namalwa cell lines (F. Danel, data not shown). The binding mode predicted by SLIDE indicates that the ligand binds in the adenosyl pocket of the binding site. Because variolins contains five- and six-membered rings that are isosteric and share chemistry with adenine, these compounds may be toxic because they bind to ATP sites in general. This interpretation is supported by research showing that pyrrolopyrimidines and their variants inhibit human protein kinases [60-63]. The key to address cytotoxicity would be to design in

selectivity for *Brugia* AsnRS by conjugating an asparagine side-chain to the variolin scaffold via an appropriate linker, and this work is in progress. Since no crystallographic information is available on the variolin B derivatives, but the Cambridge Structural Database provides a crystal structure for the variolin B scaffold (CSD code: LEPWIM), the 3D structures of the variolin B derivatives were built on this scaffold using CORINA [64],  followed by generating all low-energy conformers using Omega (OpenEye Software). All conformers within 7.5 Kcal/mol of the minimum energy conformer sampled (using the MMFF force field) were included. Results of docking these conformers will be presented in the section entitled "Impact of Main-chain Conformational Flexibility on Ligand Binding".


**Rishirilide B**

Rishirilide B (CSD code CUQZUJ), isolated from *Streptomyces rishiriensis*, has been shown to have antithrombotic activity through selective $\alpha_2$-macroglobulin inhibition, leading to the activation of plasmin [65]. It has a tricyclic scaffold that is relatively rigid and is isosteric with the adenosine moiety of ASNAMS, according to the SLIDE docking. This orientation (Figure 4B) was scored highly by both SLIDE and DrugScore. An alkyl side-chain protrudes from the binding site and is in a polar environment surrounded by residues Arg411, Glu310 and His219. The scaffold of this compound, generated by pruning the alkyl chain and a carboxymethyl group (-COOMe) from the ligand, was also docked using SLIDE. The scaffold docked with improved isostericity with the adenine portion of ASNAMS but with a poor interaction score, apparently due to loss of the side chains. The malachite green assay was carried out on a sample of rishirilide B, provided generously by Dr. Samuel Danishefsky (Sloan-Kettering Institute, New York), to test for inhibition of *Brugia* AsnRS. The compound has + and − enantiomers, and only the enantiomeric mixture showed weak inhibitory activity (Table 3), while the − enantiomer showed no inhibition. This suggests that the +

enantiomer is responsible for inhibition. Given the unavailability of purified + enantiomer for assays and the apparently weak binding of the enantiomeric mix, we chose to focus on tighter-binding compounds, as described in the following sections.

**Cycloadenosine**

Also identified by SLIDE as a top-scoring potential inhibitor was 8, 2'-cycloadenosine (CSD code for the crystal structure of its trihydrate: CYADOT). Cycloadenosine is a modified nucleoside cyclized at the C(8) and O(2') atoms and is known to be active against leukemic and other tumor cells [66]. Derivatives of cycloadenosine are also known to inhibit other tRNA synthetases: PheRS, SerRS, LysRS, ValRS, IleRS and ArgRS [67]. This compound was docked by SLIDE in a position that is isosteric with the adenosine of ASNAMS and was considered to have a potential advantage over adenosine because the bridged ring system in cycloadenosine could potentially reduce the entropic cost of binding to AsnRS. The binding mode predicted by SLIDE is shown in Figure 4C. This compound can be made more stable by replacing the oxygen in the bridge between the adenine and ribose moieties by a methylene group. While this cycloadenosine did not show inhibitory activity even at a concentration of 500 μM, this is also true of the native substrate, adenosine. The corresponding sulfamoyl asparagine derivative (CYADOT-S-Asn) was then synthesized and assayed for inhibition of *Brugia* AsnRS. The corresponding sulfamoyl asparagine derivative of cycloadenosine showed moderate inhibition, with IC50 values of 70 μM and 90 μM against *Brugia* and human AsnRS, respectively (Table 4). This represents weaker binding than indicated by IC50 values of the un-cyclized analog, ASNAMS (4.5 μM and 1.7 μM, respectively). Strain caused by cyclization of the ribose moiety to the adenine, resulting in the 5' sulfamoyl asparagine group being directed somewhat out of the asparagine pocket of the binding site, may have weakened the binding relative to non-cyclized ASNAMS. This can be addressed by redesigning the linker.

16

## Results from Screening the NCI Plated Compounds Database

### Phenanthridinol

8-Chloro-3-(hydroxy(oxido)amino)-6-phenanthridinol (NCI code: NSC114691) has a rigid tricyclic scaffold (Figure 4D) that fills the adenine pocket of AsnRS. Although this compound showed a favorable 65 μM inhibition of *Brugia* AsnRS in the experimental assay (Table 3), the planarity and aromaticity of the scaffold are a cause for concern. Planar, tricyclic scaffolds are potentially toxic because of their ability to intercalate DNA. Compounds sharing similar scaffolds have been shown to possess inhibitory activity against *Brugia* AsnRS [68] and phosphodiesterase 4 (PDE4) [69].

### Triazinylamine

4-(3-(4-amino-6-isopropenyl-1,3,5-triazin-2-yl)phenyl)-6-isopropenyl-1,3,5-triazin-2-ylamine (NCI code: NSC363624) has a symmetric structure with two substituted triazine rings connected by a phenyl group. The SLIDE-predicted orientation of the compound (Figure 4E) places one of the triazine rings isosteric with the 6-membered ring of ASNAMS and mimics its interactions with the surrounding binding site residues as well. The bridging phenyl ring in the center is docked in the ribose pocket of the binding site, but is unable to mimic the polar interactions of the ribose ring owing to its hydrophobic character. Results of the malachite green assay (Table 3) show that this compound inhibits 50% of *Brugia* AsnRS activity and 80% of human AsnRS activity at 25 μM concentration. Further studies will assess if structure-based substitutions can make it an even more potent and selective inhibitor of *Brugia* AsnRS. 1,3,5-triazine-substituted-polyamines have been shown to be active against the malarial parasite, *Plasmodium falciparum* [70].

**Phenanthrylethanone**

2-(3-methyl-1 $\lambda^5$-pyridin-1-yl)-1-(2-phenanthryl)ethanone (NCI code: NSC35467) is a charged pyridine derivative. A keto group bridges between the tricyclic moiety and the pyridine ring. The SLIDE-predicted orientation (Figure 4F) shows the tricyclic scaffold in the adenine pocket which, though isosteric with adenine, does not form any specific hydrogen bonds with the surrounding binding site residues. Planarity of the tricyclic scaffold in this compound increases concerns about potential toxicity associated with DNA intercalation. Results from the malachite green assay (Table 3) show that at 200 μM, this compound inhibits 53% *Brugia* AsnRS activity but does not inhibit human AsnRS. Although it is weak inhibitor of *Brugia* AsnRS, its selectivity for *Brugia* relative to human AsnRS is attractive. We are focusing on identifying the source of specificity within this compound to guide the optimization of more potent inhibitor scaffolds.

**Dimethylmalonamide**

$N^1,N^3$-bis(4-amino-2-methyl-6-quinolinyl)-2,2-dimethylmalonamide (NCI code: NS12156) is a symmetric compound with two bicyclic ring systems. Disubstituted malonamides are known to have weak trypanocidal activity against *Trypanosoma brucei* [71]. SLIDE docked one of the bicyclic groups, which shares chemistry and shape with the 6-membered ring of adenine, into the adenine pocket of the binding site (Figure 4G), mimicking the hydrogen bonds formed by N3 and N6 of adenine. However, in this predicted binding mode, the other bicyclic ring system could not be docked favorably. Results from the inhibition assay (Table 3) show that this compound weakly inhibits both *Brugia* and human AsnRS (~ 50% inhibition at 200 μM ligand concentration). However, a single malonamide group could be substituted to allow binding in the ribose and asparagine pockets as well as the adenine pocket.

**Success Rate of Screening**

From SLIDE screens on the CSD and NCI drug-like compounds, 45 compounds altogether were tested for *Brugia* and human AsnRS inhibition. Out of the compounds tested, seven classes of compounds predicted by SLIDE were confirmed as low- to mid-micromolar inhibitors: rishirilide and cycloadenosine from the CSD, four NCI plated compounds, and variolin B and its analogs (Tables 3 and 4 and Figure 4). Some of these compounds and their analogs (particularly the long-chain variolins, Table 4) selectively inhibit *Brugia* relative to human AsnRS. Compounds that docked with good complementarity scores and were selected for assaying but proved not to inhibit *Brugia* AsnRS are listed in Table S1 of the Supplementary Materials. The success rate in screening by SLIDE for AsnRS inhibitors is thus 7 out of 45 compounds (~15%). The best published hit rate for structure-based screening is 34% [19], involving visual screening by medicinal chemists as well as using docking scores as a guide.

**Modeling the Conformational Flexibility of *Brugia* AsnRS**

Modeling protein flexibility in ligand binding is important to improve the accuracy of results in computational ligand screening and design, as even a small change in the protein binding site conformation can introduce large changes in ligand interactions and computed binding affinities. Understanding conformational differences can also enable the design of substituents that improve binding and specificity. The results from the enzyme inhibition assays performed on ligand candidates identified by SLIDE indicate that they bind to *Brugia* AsnRS with a range of binding affinities, from low ($\geq 200$ μM) to moderate ($\leq 25$ μM). The variolin B derivatives are of particular interest because they show some selectivity towards *Brugia* AsnRS. Given the absence of active-site sequence differences between *Brugia* and human AsnRS, we sought to understand how the

flexibility of active-site loops coupled with neighboring sequence differences (Fig. 2A) could influence ligand binding through conformational differences.

To model flexibility that might contribute to inhibitor specificity, the known flexible regions in *Brugia* AsnRS were mapped by comparing the ASNAMS-bound and ligand-free (apo) crystal structures (Table 5). Analysis of the apo crystal structure indicates high mobility of the adenine-binding loop, and comparison of *Brugia* AsnRS structures bound to LBHAMP and ATP indicates the amino acid recognition loop adopts significantly different conformations depending on the type of ligand bound. To assess alternative conformations for the adenine-binding loop, ProFlex flexibility analysis [30] was performed on the ASNAMS-bound conformation of *Brugia* AsnRS to identify the coupled networks of covalent and non-covalent bonds within the protein. The ligand was removed from the protein before running ProFlex, and only those hydrogen bonds and salt bridges with energies of $\leq -1.0$ Kcal/mol were included to avoid including hydrogen bonds that are too weak to influence protein flexibility. The results of ProFlex analysis for *Brugia* AsnRS included the relative flexibility for each bond (from rigid/non rotatable through entirely flexible) and lists of which bond rotations were coupled through rings of covalent and noncovalent interactions. This information and the structure of Complex I with ASNAMS removed were used as the input to ROCK. ROCK then generated alternative low-energy conformations that preserved the non-covalent bond network, by sampling favored main-chain dihedral angles in the flexible regions of *Brugia* AsnRS [31, 91]. 500 main-chain conformers were generated, spanning from closed to very open conformations. Out of the 500 conformers generated by ROCK, there were 14 conformers with a significant main-chain deviation (more than 4.5 Å) in the adenine-binding loop. To select from these conformations the most open conformer of the protein, overall, we computed the minimum and maximum distances of the three known flexible loops (Table 5 and Figure 2A) from

20

the centroid of the co-crystallized ligand, ASNAMS, and chose the AsnRS conformation with the greatest sum of these loop distances. Thus, the open conformation analyzed not only had a significant main-chain deviation in the adenine-binding loop, but also reflected a feasible, highly open conformation of the protein overall, when compared to the closed conformation.

The most open conformation generated by ROCK for *Brugia* AsnRS shows a significant opening of the adenine-binding loop (residues K213 – L220; Figure 5B) connecting the two anti-parallel β-strands near the binding site (Figure 1B). This loop is involved in the binding of ATP and the acceptor end of the cognate tRNA and has been reported to play a significant role in conformational changes associated with other class II AARS, such as AspRS [72], LysRS [73], SerRS [74], ProRS [75] and HisRS [76]. Mutations in this loop have also been shown to affect the tRNA dependent amino acid recognition by SerRS [77]. Motion of this loop in *Brugia* AsnRS, as simulated by ROCK, exposes a new cavity near the adenosine pocket of the binding site, leading to an open conformation that emulates the apo crystal structure of the protein (Figure 5B). A shift in the position of His 219 facilitates the significant change in backbone conformation of this loop. In the ASNAMS-bound crystal structure, representing the closed conformation of the protein, His 219 is docked between Glu 310 and Arg 411 and blocks access to the cavity that is exposed in the open conformation (Figure 5C).

**Impact of Main-chain Conformational Flexibility on Ligand Binding: Interpreting the Observed Affinities and Specificities**

The binding modes predicted by SLIDE, using the closed, crystal structure *Brugia* AsnRS conformation, could explain the observed binding affinities and specificities of all the compounds except the two long side-chain derivatives of variolin B (LCM01 and LCM02 in Table 4). All low-

energy conformations of the variolin B derivatives, generated using Omega, were tested for docking into the closed conformation with SLIDE. LCM01 and LCM02 could not be docked into the binding site of the closed conformation, due to unresolvable steric collisions between the long side chains of the variolins and the backbone atoms of the adenine-binding loop (Figure 5E). However, with the same docking protocol, SLIDE was able to dock these compounds into the binding site of the open conformation generated by ROCK (Figure 5F). (The apo crystal structure was not used for docking these compounds because the adenine-binding loop is so flexible in the apo structure that its atomic coordinates could not be determined, and the interactions between the ligand and the loop therefore could not be assessed.) The binding mode of LCM01 and LCM02 docked in the ROCK-generated open conformation of *Brugia* AsnRS was in good agreement with the SLIDE-predicted binding mode of the unsubstituted variolin B in the closed conformation of the protein.

Long side-chain variolins LCM01 and LCM02 have IC50 values of 173±90 µM and 123±54 µM, respectively, indicating they are moderate to weak binders of *Brugia* AsnRS, whereas variolin B and its short side-chain derivative SMEVAR have IC50 values of ~50 µM against the enzyme. However, the assay data on the long side-chain variants of variolin B indicates they bind to *Brugia* AsnRS 3- to 8-fold more tightly than to human AsnRS. While the structure of human AsnRS has not yet been determined, there are only three sequence differences (A224T, A335S, and L353V; Figure 2A) near the active site. One of these residues, Ala 224 in *Brugia* AsnRS, is near the hinge of the adenine-binding loop, and may favor a different, more open conformation in *Brugia* than in human AsnRS, since the less bulky and non-hydrogen bonding alanine side chain is likely to restrict the motion of this loop less than threonine. The *Brugia* selectivity of long side-chain variolins may therefore be explained by their fitting only into the open conformation of the binding site, with this conformation being more readily accessible in *Brugia* than in human AsnRS due to the sequence

substitution at the base of the loop.  Similarly, the greater potency of the short side-chain variolins relative to the long side-chain analogs in *Brugia* could also be accounted for by this conformational model, since the closed conformation of the adenine-binding loop allows more favorable contacts with inhibitors.

## Discussion

A realistic expectation of structure-based drug screening is to find low affinity binders with novel scaffolds that can be further optimized by adding substituents to develop tight and selective inhibitors. Low micromolar affinity is typical of such lead compounds, especially in the case of *Brugia* AsnRS, where even the product mimic has a low micromolar (4.5 μM) IC50 (Table 2). While aminoacyl-tRNA synthetases are acknowledged as rational drug targets [6], this is the first published account of discovering new classes of AARS inhibitors by structure-based screening. Screening and docking algorithms previously had been used to model the binding and relative affinity values of known inhibitors of synthetases and their analogs. Goddard and co-workers used the HierDock virtual screening protocol to dock and predict the relative binding energies of phenylalanine analogs to the *T. thermophilus* PheRS crystal structure [78]. Lee and Kim used comparative molecular field analysis [79] to dock four known inhibitors of *S. aureus* MetRS and develop a predictive quantitative structure-activity relationship [80]. Most of the highly potent and selective AARS inhibitors discovered in recent years have come from *in vitro* screening and optimization studies [81-84]. Here we show that structure-based screening against an AARS target can identify several new classes of inhibitors.

SLIDE has identified seven classes of inhibitors showing 50% inhibition of *Brugia* AsnRS at 25-240 μM concentrations. Analogs of variolin B showed 3- to 8-fold selectivity for *Brugia* relative to human AsnRS. This success rate for identifying new ligands based on SLIDE virtual screening (~15%) supports the benefits of including 3-dimensional structural information in high-throughput screening, since structure-blind *in vitro* screening typically has a success rate of <0.1% [19]. In the process of screening, SLIDE predicts the binding mode of the docked ligand in the binding site of

the protein, which aids in optimizing the new ligands for higher affinity and selectivity for the target protein.

Incorporating complete protein flexibility during the screening of large molecular databases is sufficiently computationally intensive that it is not yet feasible. Various methods developed in recent years have shown that selecting a small ensemble of protein structures can satisfactorily represent the conformational space available to the flexible regions of a protein binding site. In particular, crystallographic snapshots, representing structures of the same protein in different conformational states, have been used to represent protein flexibility in the screening and design of ligand candidates [22, 24, 25, 85]. However, this approach is limited to experimentally observed states rather than fully representing the low-energy conformations of a protein. Molecular dynamics simulations can provide a sample of low-energy states, but remain limited to sampling motions on the sub-millisecond timescale, typically reflecting small-scale motions. However, Gorfe and Caflisch have used explicit-water MD simulations in a similar application to ours, to assess the flexibility of the substrate binding site between apo and inhibitor-bound structures of β-secretase [86]. Their results indicate that the open- and closed-flap conformations of the protein are accessible at room temperature; hence, the open conformation could also be used for β-secretase inhibitor design. ROCK is designed to sample flexible regions in a protein using a non-forcefield approach, in which the motions maintain the non-covalent bond network and avoid steric overlaps. Unlike MD, ROCK does not ascribe timescales to modeled motions nor assess the relative likelihood/energy of the generated conformers; this can be done by coupling ROCK and MD, however. By preserving non-covalent interactions, ROCK tends to sample low-energy states and follow low-barrier paths between conformations. Furthermore, the ProFlex software used with ROCK can automatically define interactions that are coupled within the protein, without the need

for expensive normal modes or essential dynamics calculations [32]. This approach can also assess how flexibility in a protein changes or redistributes upon complex formation, as has been analyzed for HIV protease [30] and the Ras-Raf complex [87].

The active sites of *Brugia* and human AsnRS have high sequence identity, with only three amino acid differences adjacent to the substrate binding sites. One of these substitutions occurs at the base of the adenine-binding loop (residue 224 is Ala in *Brugia* and Thr in human AsnRS), which likely alters its conformational flexibility in *Brugia* relative to human AsnRS. Designing inhibitor substituents that optimally fill the pocket created when this loop opens could improve inhibitor binding affinity and selectivity for *Brugia* AsnRS. This approach is supported by the work of others. For instance, Bursavich and Rich have proposed that stabilizing the conformational ensemble of an enzyme, including less-populated open conformations, can explain a range of ligand binding events that cannot be explained by lock-and-key or induced fit to a single target structure [88]. Stroud and co-workers also suggest, based on their crystallographic analysis of *C. neoformans* and *E. coli* thymidylate synthase [89], that differences in flexibility or dynamics can be employed for species-specific inhibition. Thus, we envision that considering conformational differences of active-site loops between species, rather than only considering residue differences in the static parts of binding pockets, will open a range of new possibilities for gaining specificity between closely homologous enzymes.

## Conclusions

Using a template designed to represent the active site of *Brugia* AsnRS and its interactions with known ligands, SLIDE has successfully identified seven diverse compounds that mimic the interactions between adenosine and the protein and bind with micromolar affinity. All the CSD and NCI compounds docked into the adenosyl pocket of the binding site. This protein is highly specific for binding asparagine in its aminoacyl pocket, as is generally true for AARS and their cognate amino acids. As a consequence, a productive strategy for AARS inhibitor design is to find promising scaffolds that bind strongly in the adenosyl pocket and can be linked appropriately to the cognate aminoacyl group. SLIDE identification of variolin B as an inhibitor led to the testing of variolin derivatives, which prove to be similarly potent and show selectivity for the *Brugia* enzyme. The impact of main-chain conformational flexibility on ligand binding in *Brugia* AsnRS has been modeled, providing insights into the binding of long side-chain variolins and their selectivity for the parasite AsnRS. The motions of active-site loops sampled by ROCK enable us to assess the contributions of protein conformational flexibility to ligand binding and specificity and provide a potent tool to develop even more selective inhibitors of the protein.

## Acknowledgements

## References

1.  Lazdins, J., and Kron, M., Parasitol. Today, 15 (1999) 305.

2.  Melrose, W. D., Int. J. Parasitol., 32 (2002) 947.

3.  Kron, M. A., Kuhn, L. A., Sanschagrin, P. C., Härtlein, M., Grøtli, M., and Cusack, S., J. Parasitol., 89(Suppl.) (2003) S226.

4.  Brown, K. R., Ricci, F. M., and Ottesen, E. A., Parasitology, 121 (2000) S133.

5.  Horton, J., Witt, C., Ottesen, E. A., Lazdins, J. K., Addiss, D. G., Awadzi, K., Beach, M. J., Belizario, V. Y., Dunyo, S. K., Espinel, M., Gyapong, J. O., Hossain, M., Ismail, M. M., Jayakody, R. L., Lammie, P. J., Makunde, W., Richard-Lenoble, D., Selve, B., Shenoy, R. K., Simonsen, P. E., Wamae, C. N., and Weerasooriya, M. V., Parasitology, 121 (2000) S147.

6.  Schimmel, P., Tao, J. S., and Hill, J., FASEB J., 12 (1998) 1599.

7.  Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D., Nature, 347 (1990) 203.

8.  Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., and Leberman, R., Nature, 347 (1990) 249.

9.  Woese, C. R., Olsen, G. J., Ibba, M., and Soll, D., Microbiology and Molecular Biology Reviews, 64 (2000) 202.

10. Brown, M. J. B., Mensah, L. M., Doyle, M. L., Broom, N. J. P., Osbourne, N., Forrest, A. K., Richardson, C. M., O'Hanlon, P. J., and Pope, A. J., Biochemistry, 39 (2000) 6003.

11. Casewell, M. W., and Hill, R. L. R., J. Antimicrob. Chemoth., 15 (1985) 523.

12. Hughes, J., and Mellows, G., Biochem. J., 191 (1980) 209.

13. Nilsen, T. W., Maroney, P. A., Goodwin, R. G., Perrine, K. G., Denker, J. A., Nanduri, J., and Kazura, J. W., Proc. Natl. Acad. Sci. U. S. A., 85 (1988) 3604.

14. Kron, M., Petridis, M., Milev, Y., Leykam, J., and Härtlein, M., Mol. Biochem. Parasitol., 129 (2003) 33.

15. Ramirez, B. L., Howard, O. M. Z., Dock, H. F., Edamatsu, T., Gao, P., Härtlein, M., and Kron, M., in press, J. Infect. Dis. (2006).

16. Kron, M., Marquard, K., Härtlein, M., Price, S., and Leberman, R., FEBS Lett., 374 (1995) 122.

17. Beaulande, M., Tarbouriech, N., and Härtlein, M., Nucleic Acids Res., 26 (1998) 521.

18.	Doucet, J. P., and Weber, J., Molecular Similarity, in Computer-Aided Molecule Design: Theory and Applications, Springer, pp. 328, 1996.

19.	Doman, T. N., McGovern, S. L., Witherbee, B. J., Kasten, T. P., Kurumbail, R., Stallings, W. C., Connolly, D. T., and Shoichet, B. K., J. Med. Chem., 45 (2002) 2213.

20.	Schnecke, V., Swanson, C. A., Getzoff, E. D., Tainer, J. A., and Kuhn, L. A., Proteins, 33 (1998) 74.

21.	Carlson, H. A., and McCammon, J. A., Mol. Pharmacol., 57 (2000) 213.

22.	Knegtel, R. M. A., Kuntz, I. D., and Oshiro, C. M., J. Mol. Biol., 266 (1997) 424.

23.	Osterberg, F., Morris, G. M., Sanner, M. F., Olson, A. J., and Goodsell, D. S., Proteins, 46 (2002) 34.

24.	Claussen, H., Buning, C., Rarey, M., and Lengauer, T., J. Mol. Biol., 308 (2001) 377.

25.	Carlson, H. A., Masukawa, K. M., Rubins, K., Bushman, F. D., Jorgensen, W. L., Lins, R. D., Briggs, J. M., and McCammon, J. A., J. Med. Chem., 43 (2000) 2100.

26.	Schnecke, V., and Kuhn, L. A., Proc. Int. Conf. Intell. Syst. Mol. Biol. (1999) 242.

27.	Zavodszky, M. I., Sanschagrin, P. C., Korde, R. S., and Kuhn, L. A., J. Comput. Aid. Mol. Des., 16 (2002) 883.

28.	Zavodszky, M. I., and Kuhn, L. A., Proteins, in review (2005).

29.	Zavodszky, M. I., and Kuhn, L. A., Protein Sci., 14 (2005) 1104.

30.	Jacobs, D. J., Rader, A. J., Kuhn, L. A., and Thorpe, M. F., Proteins, 44 (2001) 150.

31.	Lei, M., Zavodszky, M. I., Kuhn, L. A., and Thorpe, M. F., J. Comput. Chem., 25 (2004) 1133.

32.	Zavodszky, M. I., Lei, M., Thorpe, M. F., Day, A. R., and Kuhn, L. A., Proteins, 57 (2004) 243.

33.	Berthet-Colominas, C., Crepin, T., Haertlein, M., Kron, M., and Cusack, S., *in preparation*.

34.	Schnecke, V., and Kuhn, L. A., Perspect. Drug Discov., 20 (2000) 171.

35.	Ferrara, P., Gohlke, H., Price, D. J., Klebe, G., and Brooks, C. L., J. Med. Chem., 47 (2004) 3032.

36.	Perola, E., Walters, W. P., and Charifson, P. S., Proteins, 56 (2004) 235.

37.	Halperin, I., Ma, B. Y., Wolfson, H., and Nussinov, R., Proteins, 47 (2002) 409.

38.	Stahl, M., and Rarey, M., J. Med. Chem., 44 (2001) 1035.

39.	Bissantz, C., Folkers, G., and Rognan, D., J. Med. Chem., 43 (2000) 4759.

40.	Gohlke, H., Hendlich, M., and Klebe, G., J. Mol. Biol., 295 (2000) 337.

41.  Wang, R. X., Lai, L. H., and Wang, S. M., *J. Comput. Aid. Mol. Des.,* 16 (2002) 11.

42.  Hess, H. H., and Derr, J. E., *Anal. Biochem.,* 63 (1975) 607.

43.  Baykov, A. A., Evtushenko, O. A., and Avaeva, S. M., *Anal. Biochem.,* 171 (1988) 266.

44.  Cogan, E. B., Birrell, G. B., and Griffith, O. H., *Anal. Biochem.,* 271 (1999) 29.

45.  Danel, F., Walle, C., Kron, M., Haertlein, M., Cusack, S., and Page, M. G. P., *International Conference on Aminoacyl tRNA Synthetases,* Seoul, Korea (2004) 115.

46.  Hespenheide, B. M., Rader, A. J., Thorpe, M. F., and Kuhn, L. A., *J. Mol. Graph. Model.,* 21 (2002) 195.

47.  Rader, A. J., Hespenheide, B. M., Kuhn, L. A., and Thorpe, M. F., *Proc. Natl. Acad. Sci. U. S. A.,* 99 (2002) 3540.

48.  Taylor, R., *Acta Crystallogr. D,* 58 (2002) 879.

49.  Ihlenfeldt, W. D., Voigt, J. H., Bienfait, B., Oellien, F., and Nicklaus, M. C., *J. Chem. Inf. Comp. Sci.,* 42 (2002) 46.

50.  Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J., *Adv. Drug Deliver. Rev.,* 46 (2001) 3.

51.  Perry, N. B., Ettouati, L., Litaudon, M., Blunt, J. W., Munro, M. H. G., Parkin, S., and Hope, H., *Tetrahedron,* 50 (1994) 3987.

52.  Evers, D. L., Breitenbach, J. M., Borysko, K. Z., Townsend, L. B., and Drach, J. C., *Antimicrob. Agents Ch.,* 46 (2002) 2470.

53.  Gompel, M., Leost, M., Joffe, E. B. D., Puricelli, L., Franco, L. H., Palermo, J., and Meijer, L., *Bioorg. Med. Chem. Lett.,* 14 (2004) 1703.

54.  Anderson, R. J., and Morris, J. C., *Tetrahedron Lett.,* 42 (2001) 8697.

55.  Anderson, R. J., and Morris, J. C., *Tetrahedron Lett.,* 42 (2001) 311.

56.  Anderson, R. J., Hill, J. B., and Morris, J. C., *J. Org. Chem.,* 70 (2005) 6204.

57.  Anderson, R. J., *Total Synthesis of Variolin B*, Ph.D. thesis, Department of Chemistry, University of Canterbury, Christchurch, NZ, 2002.

58.  Marsh, C. L., *Synthetic Studies on Marine Natural Products*, M.Sc. thesis, Department of Chemistry, University of Canterbury, Christchurch, NZ, 2005.

59.  Hill, J. B., *Deoxyvariolins and Polymer Therapeutics*, Ph.D. thesis, Department of Chemistry, University of Canterbury, Christchurch, NZ, 2005.

60.  Saffer, J. D., and Glazer, R. I., *Mol. Pharmacol.,* 20 (1981) 211.

61. Davies, L. P., Jamieson, D. D., Baird-Lambert, J. A., and Kazlauskas, R., Biochem. Pharmacol., 33 (1984) 347.

62. Recchia, I., Rucci, N., Festuccia, C., Bologna, M., MacKay, A. R., Migliaccio, S., Longo, M., Susa, M., Fabbro, D., and Teti, A., Eur. J. Cancer, 39 (2003) 1927.

63. Recchia, I., Rucci, N., Funari, A., Migliaccio, S., Taranta, A., Longo, M., Kneissel, M., Susa, M., Fabbro, D., and Teti, A., Bone, 34 (2004) 65.

64. Sadowski, J., and Gasteiger, J., Chem. Rev., 93 (1993) 2567.

65. Allen, J. G., and Danishefsky, S. J., J. Am. Chem. Soc., 123 (2001) 351.

66. Neidle, S., Taylor, G. L., and Cowling, P. C., Acta Crystallogr. B, 35 (1979) 708.

67. Freist, W., Wiedner, H., and Cramer, F., Bioorg. Chem., 9 (1980) 491.

68. Dhananjeyan, M. R., Milev, Y. P., Kron, M. A., and Nair, M. G., J. Med. Chem., 48 (2005) 2822.

69. Burnouf, C., and Pruniaux, M. P., Curr. Pharm. Design, 8 (2002) 1255.

70. Klenke, B., Barrett, M. P., Brun, R., and Gilbert, I. H., J. Antimicrob. Chemother., 52 (2003) 290.

71. Goble, F. C., J. Pharmacol. Exp. Ther., 98 (1950) 49.

72. Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J. C., and Moras, D., Science, 252 (1991) 1682.

73. Shiba, K., Stello, T., Motegi, H., Noda, T., Musier-Forsyth, K., and Schimmel, P., J. Biol. Chem., 272 (1997) 22809.

74. Cusack, S., Yaremchuk, A., and Tukalo, M., EMBO J., 15 (1996) 2834.

75. Burke, B., Yang, F., Chen, F., Stehlin, C., Chan, B., and Musier-Forsyth, K., Biochemistry, 39 (2000) 15540.

76. Yaremchuk, A., Tukalo, M., Grøtli, M., and Cusack, S., J Mol Biol, 309 (2001) 989.

77. Lenhard, B., Filipic, S., Landeka, I., Skrtic, I., Soll, D., and Weygand-Durasevic, I., J. Biol. Chem., 272 (1997) 1136.

78. Wang, P., Vaidehi, N., Tirrell, D. A., and Goddard, W. A., J. Am. Chem. Soc., 124 (2002) 14442.

79. Cramer, R. D., Patterson, D. E., and Bunce, J. D., J. Am. Chem. Soc., 110 (1988) 5959.

80. Kim, S. Y., and Lee, J., Bioorgan. Med. Chem., 11 (2003) 5325.

81. Finn, J., Mattia, K., Morytko, M., Ram, S., Yang, Y. F., Wu, X. M., Mak, E., Gallant, P., and Keith, D., Bioorg. Med. Chem. Lett., 13 (2003) 2231.

82. Lee, J., Kim, S. E., Lee, J. Y., Kim, S. Y., Kang, S. U., Seo, S. H., Chun, M. W., Kang, T., Choi, S. Y., and Kim, H. O., Bioorg. Med. Chem. Lett., 13 (2003) 1087.

83. Jarvest, R. L., Erskine, S. G., Forrest, A. K., Fosberry, A. P., Hibbs, M. J., Jones, J. J., O'Hanlon, P. J., Sheppard, R. J., and Worby, A., Bioorg. Med. Chem. Lett., 15 (2005) 2305.

84. Bernier, S., Akochy, P. M., Lapointe, J., and Chenevert, R., Bioorgan. Med. Chem., 13 (2005) 69.

85. Cavasotto, C. N., and Abagyan, R. A., J. Mol. Biol., 337 (2004) 209.

86. Gorfe, A. A., and Caflisch, A., Structure, 13 (2005) 1487.

87. Gohlke, H., Kuhn, L. A., and Case, D. A., Proteins, 56 (2004) 322.

88. Bursavich, M. G., and Rich, D. H., J. Med. Chem., 45 (2002) 541.

89. Finer-Moore, J. S., Anderson, A. C., O'Neil, R. H., Costi, M. P., Ferrari, S., Krucinski, J., and Stroud, R. M., Acta Crystallogr. D, 61 (2005) 1320.

90. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L., Acta Crystallogr. D, 54 (1998) 905.

91. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M., J. Appl. Crystallogr., 26 (1993) 283.

**Table 1**: Data collection and refinement statistics of *Brugia* AsnRS crystal structures[a]

| | Complex I<br>AsnRS: ASNAMS | Complex II<br>AsnRS:LBHAMP:ATP | Apo AsnRS |
|---|---|---|---|
| **Data collection** | | | |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell dimensions (Å) | 55.5 125.7 144.3 | 57.9 106.4 161.4 | 58.8 108.8 162.4 |
| Resolution range (Å) | 30-1.9 (2.0-1.9) | 50-2.4 (2.49-2.4) | 49-2.3 (2.42-2.3) |
| Rsym I [b, c] (%) | 8.7 (43.7) | 9.5 (42.3) | 7.1 (33.3) |
| Completeness [b] (%) | 89.5 (69.6) | 97.6 (95.4) | 97.4 (91.9) |
| **Refinement statistics**[d] | | | |
| *R*-factor (%) | 22.6 | 21.1 | 26.9 |
| $R_{\text{free}}$ (%) | 26.3 | 28.6 | 32.2 |
| **Ramachandran plot**[e] | | | |
| Favoured (%) | 86.9 | 79.1 | 86.5 |
| Additional (%) | 12 | 19.2 | 12.4 |

[a]Determined by Carmen Berthet-Colominas, Michael Härtlein, Thibaut Crepin and Stephen Cusack, EMBL Outstation, Grenoble, France [33]. X-ray coordinates will be provided upon request to cusack@embl-grenoble.fr.

[b]Values in parentheses are for the highest-resolution shell.

[c]$R\text{sym} (I) = [\sum_{hkl}\sum_i |<I_{hkl}> - I_{hkl,i}|]/[\sum_{hkl} \sum_i |I_{hkl}|]$, where *i* is the number of reflection *hkl*.

[d]Refinement with CNS [90]

[e]Ramachandran diagram has been calculated with PROCHECK [91]

**Table 2**: Predicted AsnRS-inhibitor complementarity scores (SLIDE and DrugScore) and experimentally determined affinity values of known ligands of *Brugia* AsnRS. The SLIDE and DrugScore values were computed for dockings into chain A of the ASNAMS-bound *Brugia* AsnRS structure.

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x $10^5$) | IC$_{50}$ ($\mu$M) *Brugia* AsnRS | Human AsnRS | Structure[c] |
|---|---|---|---|---|---|---|
| ASNAMS |  | 59 | -7.25 | 4.5 | 1.7 | Complex I |
| LBHAMP |  | 47 | -7.39 | 4 | ND[d] | Complex II |
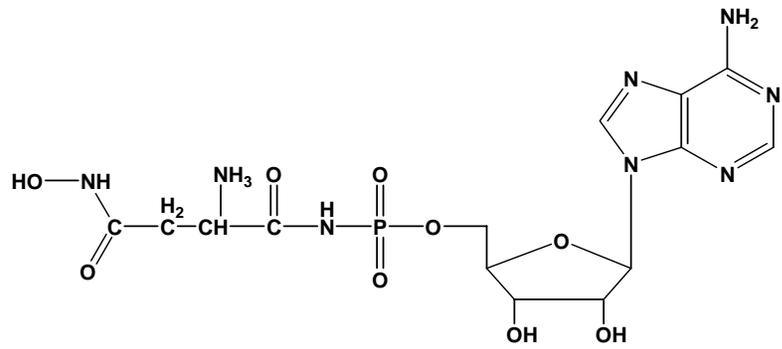
34

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x $10^5$) | IC$_{50}$ (µM) | | Structure[c] |
| | | | | Brugia AsnRS | Human AsnRS | |
| --- | --- | --- | --- | --- | --- | --- |
| ATP | | 42 | -5.91 | ND | ND | Complex II |
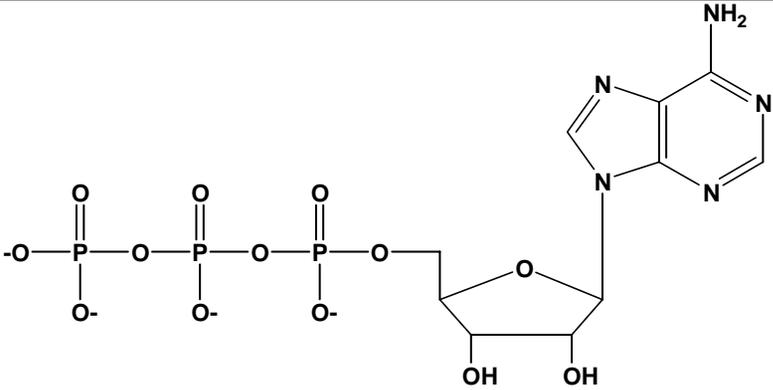
[a]A higher value of SLIDE score is more favorable.

[b]A more negative value of DrugScore is more favorable.

[c]Refer to Table 1 for more information on the structural complexes. In complex I, ASNAMS is bound to both chains of AsnRS, while in complex II LBHAMP is bound to one chain of AsnRS and ATP to the other.

[d]Not determined

**Table 3**: Predicted AsnRS-inhibitor complementarity scores (SLIDE and DrugScore) and experimentally determined affinity values of SLIDE-predicted inhibitors of *Brugia* AsnRS. Experimental affinities are based on using the malachite green assay for *Brugia* and human AsnRS, while the SLIDE and DrugScore values were computed for dockings into chain A of Complex I (see Table 1).

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x $10^5$) | IC$_{50}$ (µM) | |
| | | | | *Brugia* AsnRS | Human AsnRS |
| --- | --- | --- | --- | --- | --- |
| Rishirilide |  | 46 | -5.49 | 240 | 400 |
| NSC114691 |  | 46 | -3.70 | 65 | 50 |

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x $10^5$) | IC$_{50}$ (µM) | |
|---|---|---|---|---|---|
| | | | | *Brugia* AsnRS | Human AsnRS |
| NSC363624 |  | 28 | -5.10 | 25 | ND[c] 80% inhibition at 25µM |
| NSC35467 |  | 42 | -4.30 | ND 53% inhibition at 200µM | ND No inhibition at 200µM |

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x $10^5$) | IC$_{50}$ (µM) | |
| --- | --- | --- | --- | --- | --- |
| | | | | *Brugia* AsnRS | Human AsnRS |
| NSC12156 | | 41 | -5.26 | ND 47% inhibition at 200µM | ND 54% inhibition at 200µM |
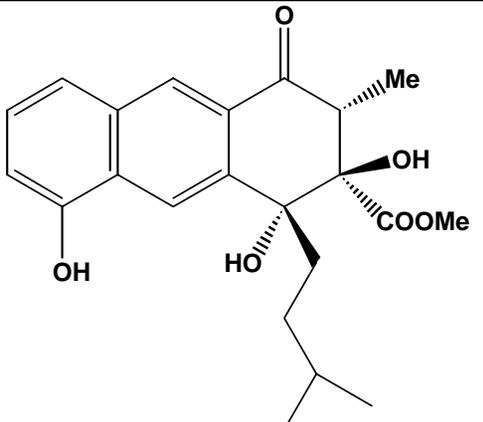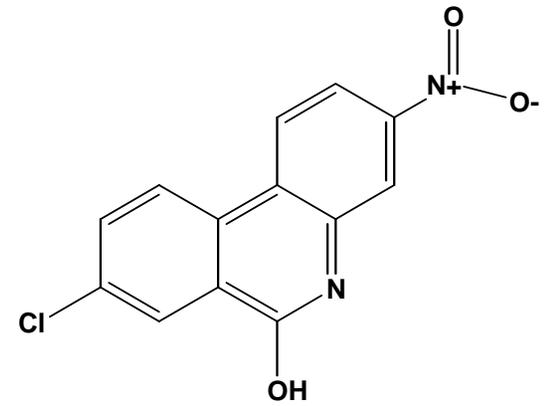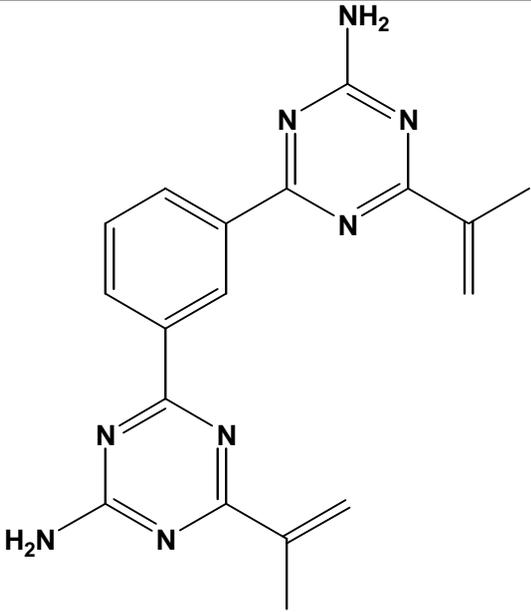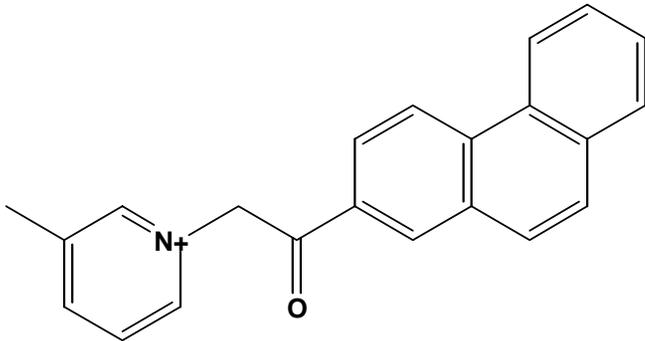
[a]A higher value of SLIDE score is more favorable.

[b]A more negative value of DrugScore is more favorable.

[c]Not determined

**Table 4**: Predicted AsnRS-inhibitor complementarity scores (SLIDE and DrugScore) and experimentally determined affinity values of synthesized analogs of the SLIDE-discovered *Brugia* AsnRS inhibitors variolin B and cycloadenosine. Experimental affinities are based on using the malachite green assay for *Brugia* and human AsnRS, while the predicted scores were computed for dockings into chain A of the ASNAMS-bound *Brugia* AsnRS structure.

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x 10$^5$) | IC$_{50}$ (µM) | |
| --- | --- | --- | --- | --- | --- |
| | | | | *Brugia* AsnRS | Human AsnRS |
| SMEVAR |  | 31 | -4.67 | 51±8 | 101±10 |
| LCM01 |  | 23 | ND[c] | 173±90 | >500 |

| Ligand | 2D Structure | SLIDE Score[a] | DrugScore[b] (x $10^5$) | IC$_{50}$ (μM) Brugia AsnRS | IC$_{50}$ (μM) Human AsnRS |
|---|---|---|---|---|---|
| LCM02 | | 26 | ND | 123±54 | >1000 |
| CYADOT-S-ASN[d] | | ND | ND | 76 | 90 |
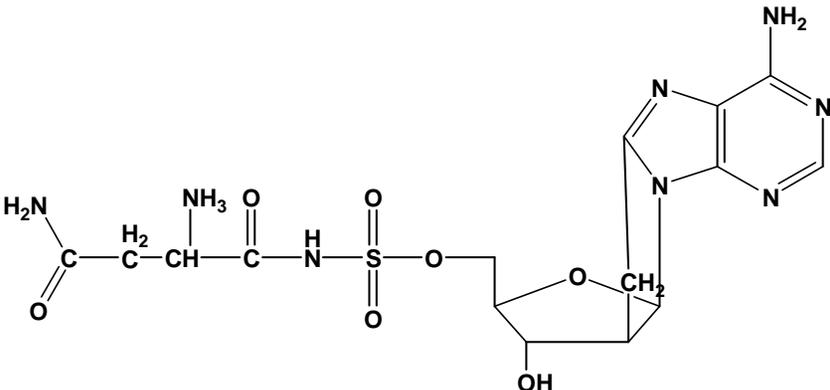
[a]A higher value of SLIDE score is more favorable.

[b]A more negative value of DrugScore is more favorable.

[c]Not determined

[d]This analog was synthesized based on the SLIDE docking of cycloadenosine (CYADOT), with addition of the asparaginyl group guided by the sulfamoyl group in ASNAMS (Table 2).

**Table 5**: Known flexible regions in *Brugia* AsnRS. See structures in Figure 2.

| Residue Range of Flexible Loop | Remarks |
| --- | --- |
| K213 – L220<br>Adenine binding loop | Coordinates could not be determined due to mobility of these residues in both chains A and B of the apo crystal structure. |
| E297 – F302<br>Distal loop | Residues were mobile (no coordinates determined) in chain A of the apo crystal structure. |
| Q163 – L172<br>Amino acid substrate recognition loop | Significant conformational differences were observed for these residues between the LBHAMP-bound monomer and the ATP-bound monomer of the same crystal structure. |

**Figure Legends**

**Figure 1**: (A) *Brugia* AsnRS dimer, with one chain of the dimer colored yellow while the other is colored cyan. ASNAMS is rendered as a space filling model, colored by atom type (carbon, green; oxygen, red; and nitrogen, blue; sulfur, yellow; magnesium, white). The associated magnesium ions are shown as white spheres. (B) The three class II AARS sequence motifs are mapped onto the structure. Motif 1 is shown in purple, motif 2 in yellow (including the adenine binding loop at bottom center), and motif 3 in magenta. ASNAMS is shown in atom colored tubes and interacts primarily with the motif 2 residues.

**Figure 2**: (A) Known flexible regions of *Brugia* AsnRS, based on comparison of the three crystal structures, are shown in green ribbons, while the rest of the monomer is shown in grey ribbons (see Table 5 for details; residue ranges are given for each flexible loop). The three residues that differ near the active site in *Brugia* and human AsnRS are rendered in space filling models colored by atom type, with carbon atoms colored grey. ASNAMS is rendered in atom-colored tubes with carbon atoms colored orange. The template generated by SLIDE to represent the binding site is rendered in small stars. The template points, representing ligand chemistry that would be favored at that site, are colored red for hydrogen-bond acceptor; blue for hydrogen-bond donor; white for hydrogen-bond donor and/or acceptor; and, green for hydrophobic interactions. (B) ROCK flexible loop conformations generated from the ASNAMS-bound *Brugia* AsnRS crystal structure. The diverse conformations shown were chosen from 500 ROCK conformers by selecting those conformations with the largest pairwise RMSDs in main-chain dihedral angles, as described in [31]. The monomer is rendered in ribbons colored by the flexibility index, with blue being the most rigid and red being the most flexible regions of the structure, according to ProFlex. The ligand is shown in atom colored tubes with carbon atoms in orange. The three active-site-neighboring residues that differ between *Brugia* and human AsnRS are rendered as space filling models colored by flexibility index. Different conformers for the flexible regions of *Brugia* AsnRS, generated by ROCK, are shown in green ribbons, with the adenine binding loop shown at bottom.

**Figure 3**: Enrichment plot for the three different scoring functions - SLIDE score, DrugScore and X-Score. The scoring functions were assessed for their ability to distinguish as top-scoring compounds the three known *Brugia* AsnRS ligands (ASNAMS, LBHAMP, ATP) when mixed into

a set of 1000 drug-like small molecules (from http://bioinfo-pharma.u-strasbg.fr/bioinformatics-cheminformatics-group.html). The top-scoring binding orientation and conformer for each ligand candidate was ranked relative to the other candidates.

**Figure 4**: The SLIDE-predicted binding mode of (A) variolin B, shown in thick atom colored tubes (carbon, green; oxygen, red; and nitrogen, blue), is compared with ASNAMS bound in the crystal structure, shown in atom colored tubes but with carbon atoms colored orange. Side chains of binding-site residues rotated by SLIDE, to model induced-fit while docking, are shown in purple tubes, while their original positions in the crystal structure are shown in grey. Predicted binding modes are also shown for the SLIDE-discovered inhibitors, in atom-colored tubes, compared with the crystallographic binding modes of ASNAMS. (B) Rishirilide B, (C) cycloadenosine, (D) NSC114691 (phenanthridinol), (E) NSC363624 (triazinylamine), (F) NSC35467 (phenanthrylethanone), and (G) NSC12156 (dimethylmalonamide).

**Figure 5**: (A) The ASNAMS-bound closed crystal structure conformation of *Brugia* AsnRS is compared with the ligand-free (apo) crystal structure. ASNAMS is shown as atom colored tubes, with carbon atoms colored orange. The Connolly solvent-accessible molecular surface of the apo conformation, lacking atomic coordinates of the mobile adenine binding loop, is rendered as a solid surface (off-white), while that of the closed conformation is rendered as mesh in cyan, showing the closed adenine binding loop as a cyan-colored ribbon. (B) The ROCK-generated most open conformation of *Brugia* AsnRS is compared with the apo crystal structure. The Connolly surface of the apo crystal structure conformation is again rendered as solid, while that of the open conformation is rendered as green mesh. The most open conformation of the adenine binding loop is rendered as a green ribbon. (C) The adenine binding loop residue His 219 (Connolly surface colored by atom type, with carbon atoms colored cyan) is docked between Glu 310 and Arg 411 (atom-colored Connolly surface, with carbons in green) in the closed crystal structure conformation of *Brugia* AsnRS. ASNAMS is shown in atom colored tubes with carbons colored orange. (D) His 219 (Connolly surface colored by atom type) undergoes a significant motion away from adenosine, due to reorientation of the main chain in the ROCK-generated open conformation. (E) LCM02, a long side-chain variolin B derivative shown in atom colored tubes, manually docked in the closed crystal structure conformation of *Brugia* AsnRS (with Connolly surface colored purple) according to the predicted binding mode of variolin B (Figure 4A). ASNAMS is shown in atom colored tubes,

with carbons colored orange. The steric clashes between the side-chain of LCM02 at the lower left and the closed conformation of the adenine binding loop (purple ribbon) could not be resolved by any single bond rotations in the ligand or protein. (F) LCM02, long-chain variolin B derivative shown in atom colored tubes, docked in the ROCK-generated open conformation of *Brugia* AsnRS (Connolly surface colored cyan) to match the variolin B binding mode. For reference, ASNAMS is shown in atom colored tubes with carbons colored orange. There were no steric clashes between LCM02 and the open conformation of the adenine binding loop (cyan ribbon), and the long side chain with dimethyl amino group (lower-left) fits well into the channel uncovered by opening the adenine binding loop and the proposed His 219 gate.

**Figure 1**

**A**

A335S

L353V

Q163 – L172

E297 – F302

K213 – L220

A224T

**B**
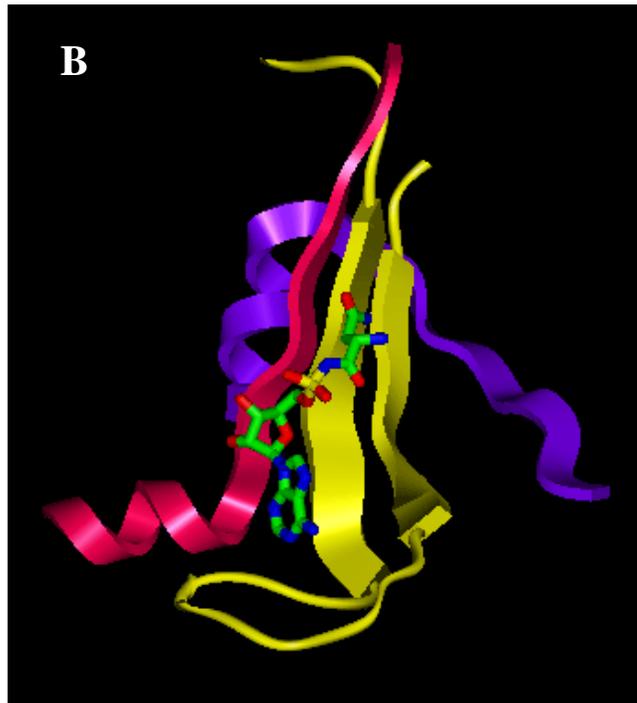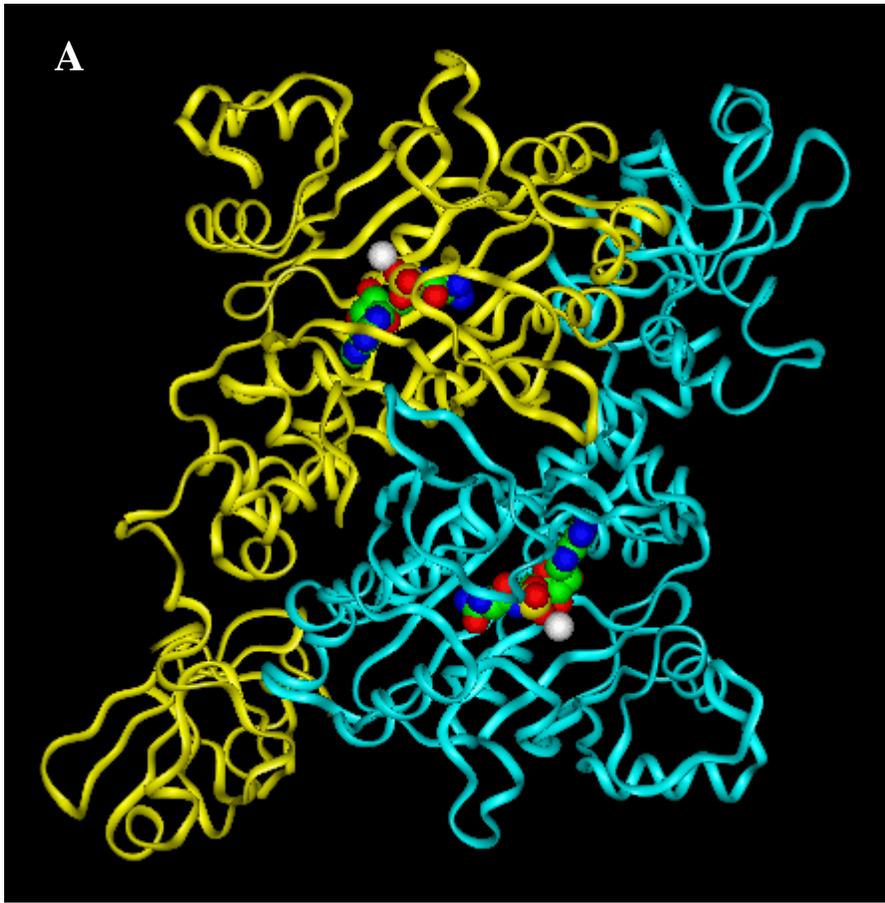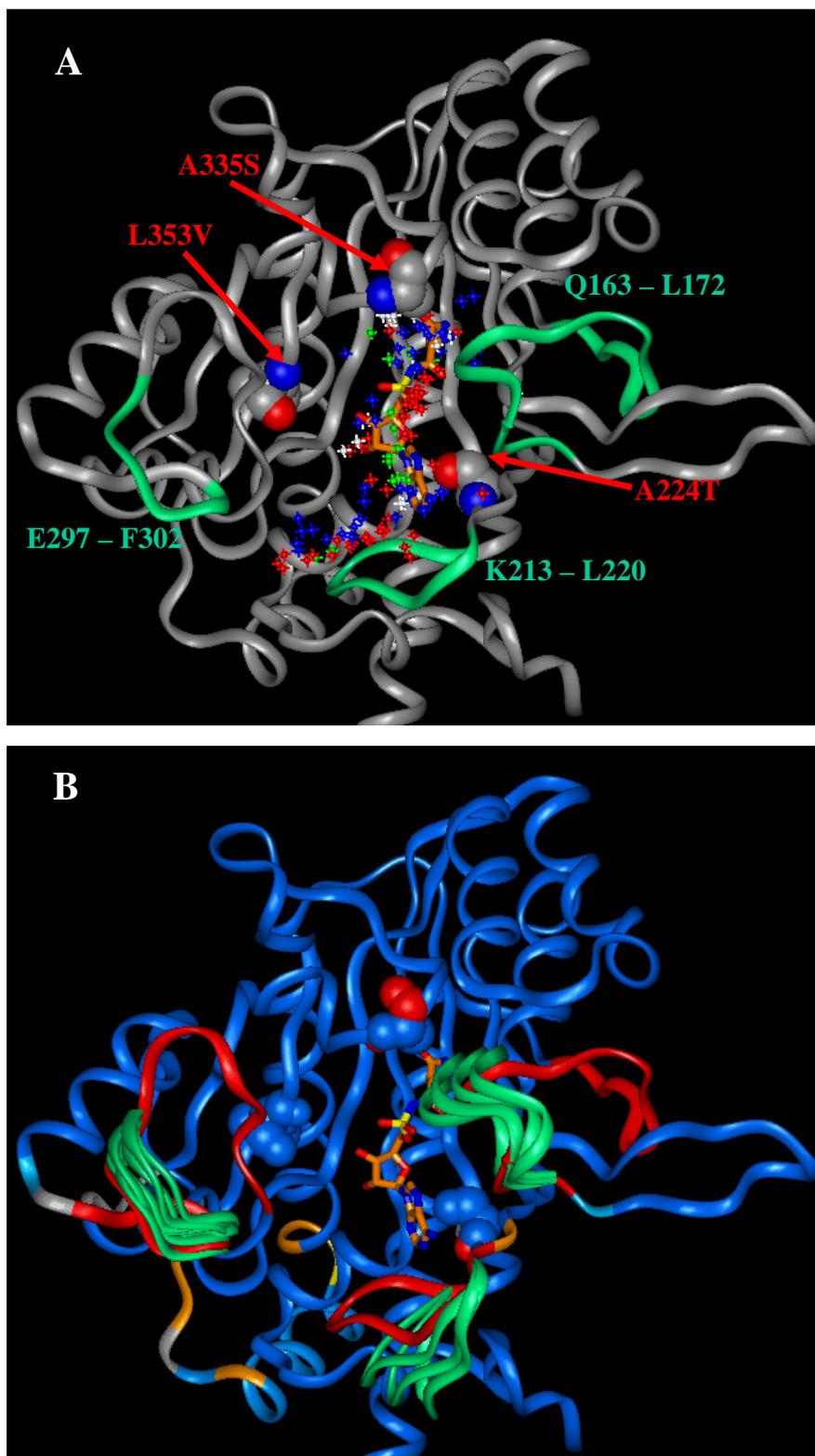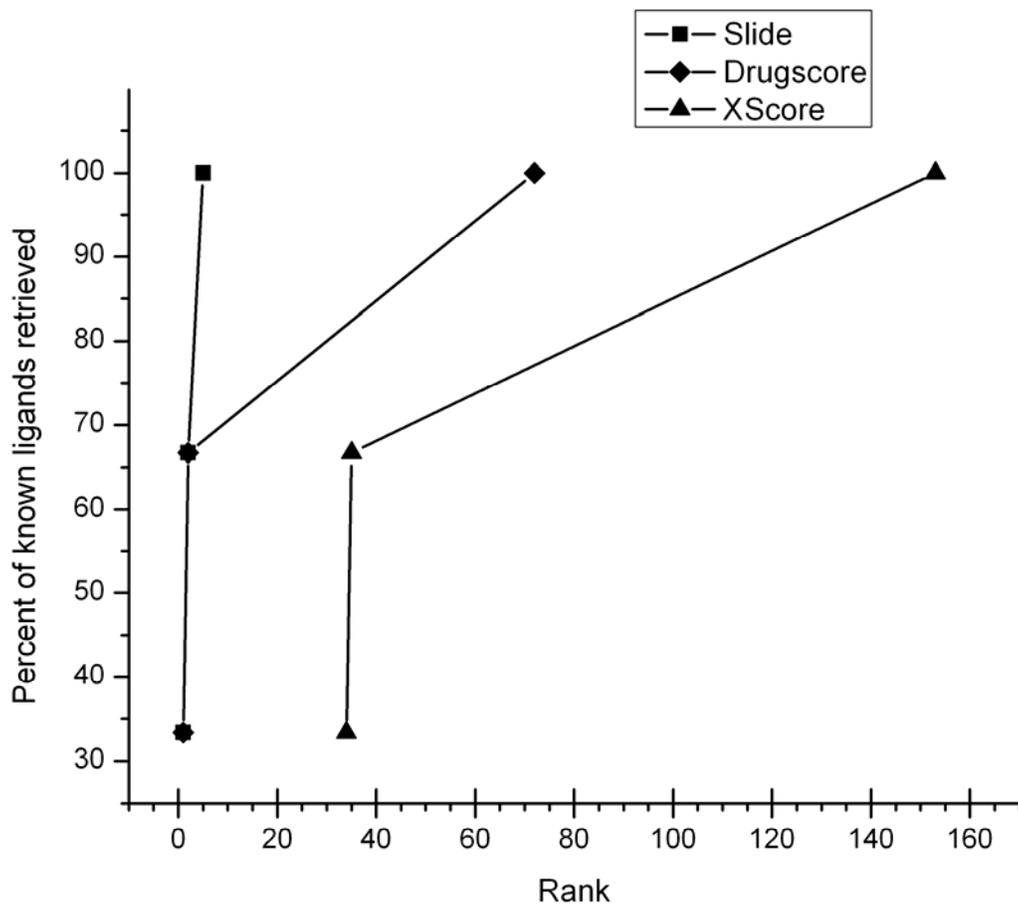
**Figure 2**

**Figure 3**

**Figure 4**

**Figure 5**