Docking-based Virtual Screening for the Discovery of RNA-targeting Molecules: Identification of Selective Riboswitch Binding Ligands.

Julia Stille[1,#], Nathania Takyi[2,3,#], Omma S. Ayon[1,#], Maira Rivera[1], Thershan Satkunarajah[1], Joshua Pottel[3], Anthony Mittermaier,[1] Maureen McKeague[*1,4] and Nicolas Moitessier[*1,3]

[1]Department of Chemistry, McGill University, 801 Sherbrooke St W, Montreal, QC, Canada H3A 0B8; [2]Department of Chemistry and Biochemistry, Faculty of Arts and Science, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada; [3]Molecular Forecaster Inc. 910-2075 Robert Bourassa, Montreal Quebec H3A 2L1, Canada and [4]Pharmacology and Therapeutics, Faculty of Medicine and Health Sciences, McGill University, Montreal, QC H3G 1Y6, Canada

ABSTRACT: We present herein one of the very first applications of molecular docking towards the discovery of novel RNA-targeting molecules. Our in-house docking program FITTED was used to perform independent virtual screening campaigns against the TPP, SAM, and FMN riboswitches. Based on the predicted docking scores and poses, between 14 and 20 compounds were selected from commercial libraries and purchased for experimental evaluation of binding to their respective riboswitches by surface plasmon resonance. Promisingly, two compounds displayed highly specific and dose-dependent binding to the TPP riboswitch and FMN riboswitch, with experimentally-determined K_{Ds} of 170 μM and 220 μM, respectively. This work highlights the promise of modifying and applying docking programs for the discovery of nucleic acid- targeting molecules.

INTRODUCTION

Docking small molecules to nucleic acids. The use of computational methods in drug discovery campaigns is well-established, and their application has contributed to the discovery of many active compounds and approved drugs.[1] In practice, as proteins have historically been favored as drug targets, docking programs have been primarily developed and optimized for use with protein targets.[2] However, nucleic acid structures such as riboswitches or splicing elements have emerged as promising therapeutic targets and existing computational programs developed for proteins often face difficulties when applied directly to nucleic acids, limiting their use.[2] Notably, nucleic acids are typically more flexible, polar, and solvated when compared to the more well-defined, lipophilic binding pockets of many protein drug targets. Additionally, relatively few nucleic acid-ligand complexes have been experimentally determined compared to protein-ligand complexes. This has implications for the parameterization of docking programs which often rely on fitting to/optimizing on experimental data. As a result, nucleic acid-specific docking programs, including rDock[3] and NLDock[4] have been devised but currently do not necessarily outperform protein-based docking programs.[5 6, 7]

Riboswitches as potential drug targets. The growing resistance of pathogenic bacteria to antibiotics and the emergence of multi-drug resistant bacteria represents a major threat to human welfare.[8] The issue of antibiotic resistance has been exacerbated by the overuse and improper use of antibiotics, driving the evolutionary development of resistance.[9] As most antibiotics target bacterial proteins, there has been a long-standing interest in alternative molecular targets, including nucleic acids. Among those are the bacterial RNA riboswitches[10-12] discovered in the early 2000's.[12-15] Riboswitches are regulatory segments of RNA that control gene expression in response to binding a small molecule or ion. The binding of these ligands by the riboswitch aptamer domain induces a conformational change in the riboswitch expression

platform that influences gene translation or transcription. Most known riboswitches act in a *cis* manner, in which ligand binding results in the inhibition of gene expression, either through the formation of a terminator loop or sequestration of the ribosome-binding site (Figure 1, Figure S1). Riboswitches bind their native substrate with high affinity and specificity, enabling selective gene expression control even in the presence of a large number of structurally similar metabolites.[16-18] This regulatory mechanism is found almost exclusively in bacteria and archaea – to date, roughly 56 classes of riboswitches have been identified in more than 6,000 bacterial species.[19, 20]

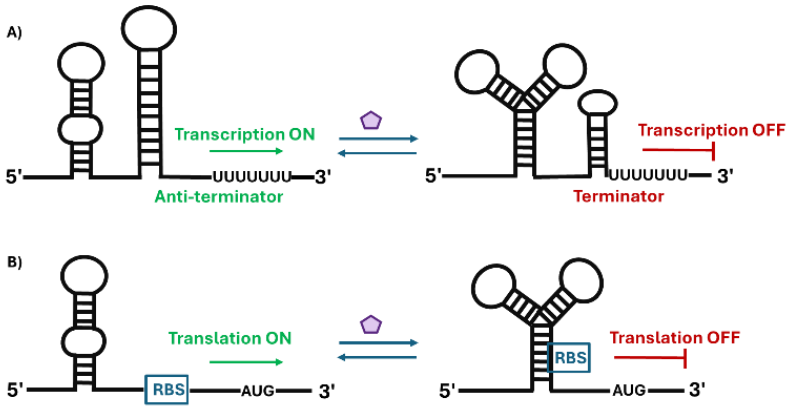

Figure 1 Two key riboswitch mechanisms for controlling gene expression. A) Transcription termination. B) Translation inhibition.

Targeting riboswitches with small molecules is still in its infancy, but preliminary reports support the potential for this novel antibiotic target. For example, the antimicrobial activity of roseoflavin, a chemical analogue of flavin mononucleotide (FMN), was a result of its ability to directly bind the FMN riboswitch, altering FMN riboswitch-regulated gene expression.[21] Furthermore, a phenotypic screen by researchers at Merck led to the discovery of Ribocil, a selective inhibitor of the FMN riboswitch capable of inhibiting bacterial cell growth, and it successfully protected mice from infection (Figure 2).[22] Importantly, Ribocil is structurally distinct from FMN and does not display the same off-target activity or cellular transport issues observed with roseoflavin, providing a promising proof-of-principle for riboswitch inhibition by synthetic small molecules.

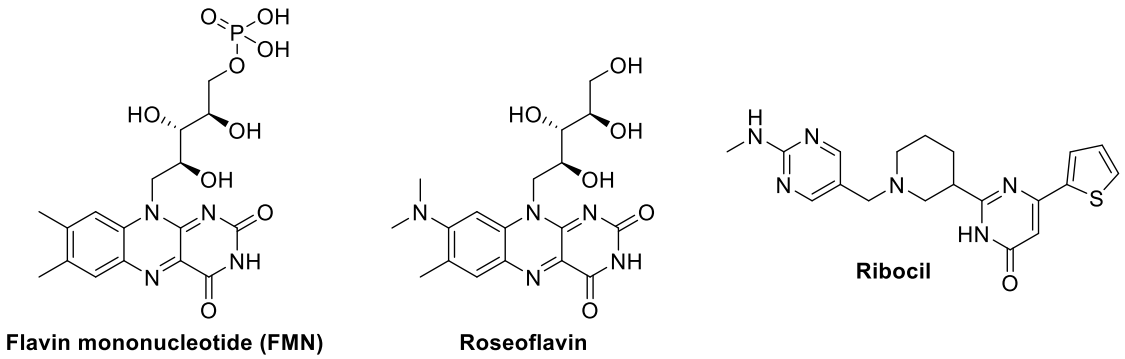

Figure 1. Chemical structures of the native ligand (FMN) and two inhibitors (Roseoflavin and Ribocil) for the FMN riboswitch.

Docking to riboswitches. Despite interest in developing riboswitch-binding molecules, little has been developed to improve their virtual screening. In one early example, more than ten years ago, the DOCK docking program with RNA-specific parameters was used to virtually screen small molecules against the adenine-binding riboswitch. Four ligands with mid-micromolar affinity were identified and their predicted binding poses were confirmed by x-ray crystallography.[23] Similarly, ligands binding to the Pre-Q1

riboswitch were discovered using FRED, HYBRID and FlexX docking programs that were developed for protein-ligand docking . However, to ensure applicability to nucleic acids, the authors used a pharmacophore model based on the binding mode of the Pre-Q1 riboswitch to its native ligand to aid in their selection of compounds for experimental testing.[7] The experimental results indicated that the number of matching interactions between the binding pose and the pharmacophore model was more predictive of binding affinity than the docking score alone.[7]

Here, we describe the application of our docking program FITTED (flexibility induced through targeted evolutionary description) to the discovery of novel riboswitch-targeting molecules. FITTED has been initially developed and optimized for protein targets and has been applied to the discovery of a number of enzyme inhibitors.[24-26] Following these successful prospective validations, we sought to evaluate FITTED on nucleic acids. Prior to its application, we modified our program to account for nucleic acids and challenged it for the discovery of small molecules binding RNA motifs. We report herein our discovery of riboswitch binding molecules through docking-based virtual screening on three prevalent riboswitches (TPP, FMN, and SAM riboswitches) present in a variety of bacteria.

RESULTS AND DISCUSSION

Modifying FITTED for Docking to Nucleic Acids. While FITTED and its accessory programs PREPARE and PROCESS were originally developed for docking to proteins,[27] they have since undergone several modifications and optimizations for docking to nucleic acids. In the most basic optimization, nucleotides were added to the programs as available "residues", allowing interaction sites to be generated around nucleotides by PROCESS that are employed by FITTED during the conformational search and scoring process. Additionally, the overall charge on phosphate anions was neutralized to account for the presence of counterions such that electrostatic interactions are not overestimated, a well-established approach.[28] Nucleic acid binding sites are typically more exposed and solvated when compared to those in proteins, and therefore water often plays a critical role in binding. To account for this, we previously implemented "displaceable waters", which enable ligand binding through bridging water molecules or displacement of water molecules in poses in which direct ligand-nucleic acid binding is more favourable.[28] In the cases where water molecules are not resolved in the crystal structure, particle waters (single bead model for water molecules) may be added using our program SPLASH'EM.[29]

Nucleic acid structures will often contain conserved metal ions in the binding site. For example, binding of TPP in the TPP riboswitch is mediated by the presence of two Mg^{2+} ions in the binding site (Figure 3A). To more accurately model the binding of ligands to metals, our docking program was further optimized by fitting the molecular mechanics scoring function to more accurate quantum mechanics-derived potential energy curves.[30] For Mg^{2+} ions specifically, the interaction is also modelled by a periodic function to more accurately capture its preference for an octahedral binding geometry.

Benchmarking FITTED for Docking to Riboswitches. To evaluate the ability of our program to identify riboswitch binding molecules, we selected three riboswitches with available structural information: the TPP, SAM-I and FMN riboswitches (see supporting information for detailed description of these riboswitches). It was envisioned that our modified in-house docking program FITTED could be applied toward the identification of novel TPP, FMN and SAM-I/IV riboswitch inhibitors by virtual screening. Tests of our docking program including self-docking, cross-docking and virtual screening with known binding molecules were carried out. The self-docking on all three riboswitches led to accurate predictions (TPP: RMSD = 1.64 Å, SAM-I: 1.22 Å, FMN: 1.30 Å). Cross-docking with FMN using 15 structures showed an accuracy of 79 % (metrics: RMSD < 2.0 Å) when 3F2Y was used, while the screening of a library of 9 known TPP riboswitch binders and 180 decoy compounds confirmed a reasonable accuracy of our scoring function to distinguish between active and inactive molecules (AUROC = 0.9, Figure S2). While this benchmark was carried out on a small library (due to limitations in available data), it was nonetheless promising and encouraged us to proceed with our virtual screening campaigns.

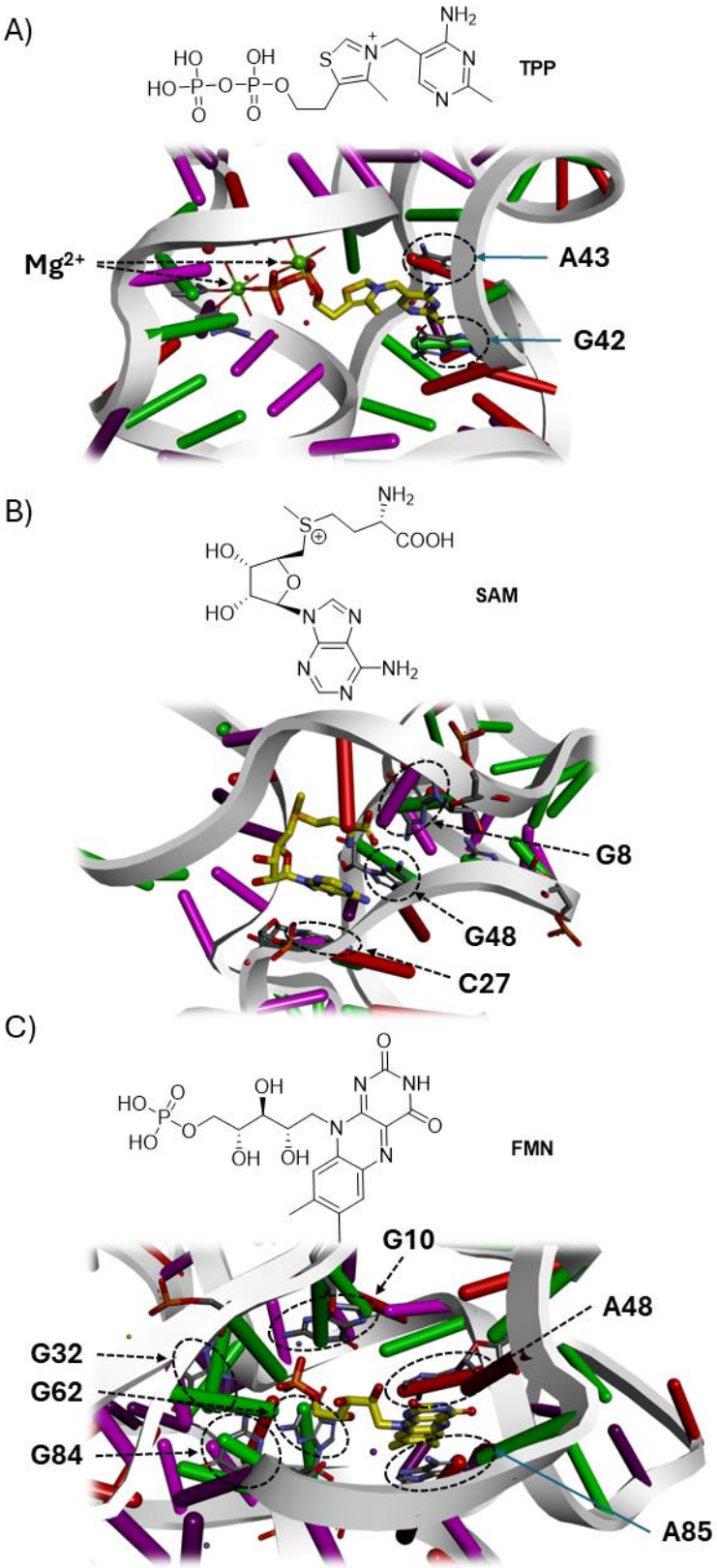

Figure 3. Binding to riboswitch for A) TPP (PDB ID: 2GDI), B) SAM I/IV (PDB ID: 4L81) and C) FMN (PDB ID: 3F2Y).

To select the molecules for docking and/or to orient the docking towards more promising solutions, we first analyzed the available structural information (Figure 3). In the case of the TPP riboswitch, it appears that the inhibitor should ideally bind both the pyrophosphate and pyrimidine sensor helices, mimicking the binding of the native TPP ligand to stabilize the riboswitch in the "OFF" state. As a means of

prioritizing potential inhibitors that can form both of these key interactions, a pharmacophore model was incorporated into the docking protocol. The pharmacophore model was based on the crystal structure of TPP bound to the aptamer domain of the TPP riboswitch (PDB: 2GDI). For a docked pose to be considered for scoring, it must contain at least one hydrophobic interaction within the pyrimidine sensor helix and an interaction with at least one of the two magnesium ions of the pyrophosphate sensor helix.

Similarly, for the SAM riboswitch, a significant hydrogen bond network between the SAM adenine and U47 as well as stacking with C27 were observed. In parallel, the amino acid of the methionine portion of SAM is making four hydrogen bonds with G8 and G48. We reasoned that prioritizing these key interactions will be critical when selecting compounds for experimental testing. Finally, for the FMN riboswitch, notable interactions include a large hydrogen bond network between the terminal phosphate and G10, G11, G32, G62 and G84 and the intercalation of the polycyclic aromatic portion of FMN between A48 and A85.

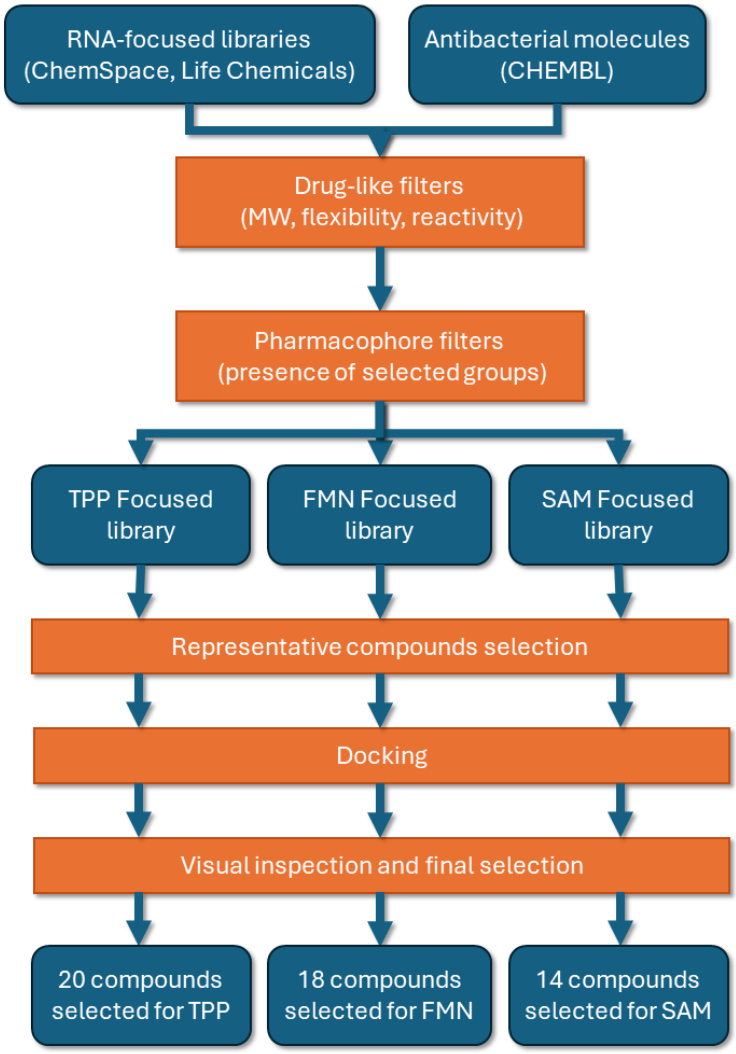

Figure 4. Computational protocol (see supporting information for details).

Using the combined structural information, we designed appropriate selection criteria for the library selections. For example, the presence of a metal binding motif and an aromatic ring was required for the TPP library. In addition to these structural filters, the computational protocol includes drug-like filters (Figure 4). These three focused libraries were then docked to the three riboswitches. The top scoring molecules were then visually inspected for the key interactions to select final compounds for experimental

testing. Through visual inspection, the top 200 scoring molecules were narrowed down to 20 (TPP, Table S1), 14 (SAM-I, Table S2) and 18 (FMN, Table S3) molecules (Figures 4 and 5).

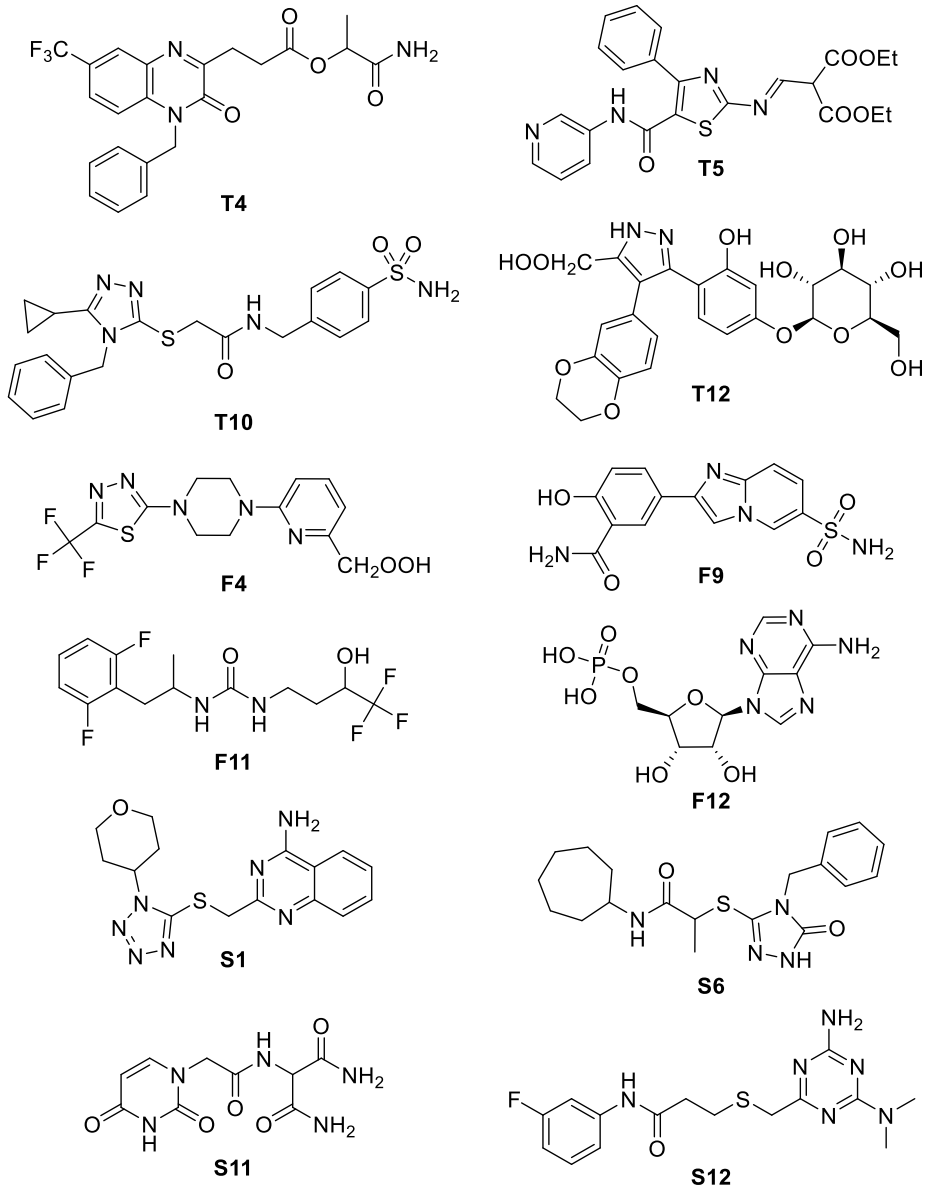

Figure 5. Example of compounds selected for experimental testing. T4, T5, T10 and T12 for TPP riboswitch, F4, F9, F11 and F12 for FMN riboswitch and S1, S6, S11 and S12 for SAM riboswitch.

Surface plasmon resonance for characterizing riboswitch-targeting molecules. To screen and validate binding of the selected small molecules to the riboswitches, we conducted an SPR assay.[31] This assay was chosen because it requires relatively little RNA sample, conducive for screening compared to other methods, provides highly sensitive detection of binding of low molecular weight molecules such as ligands that bind to riboswitches, and accounts for specificity by including a negative control "scrambled control RNA" on the reference flow cell (Figure 6).[31, 32]

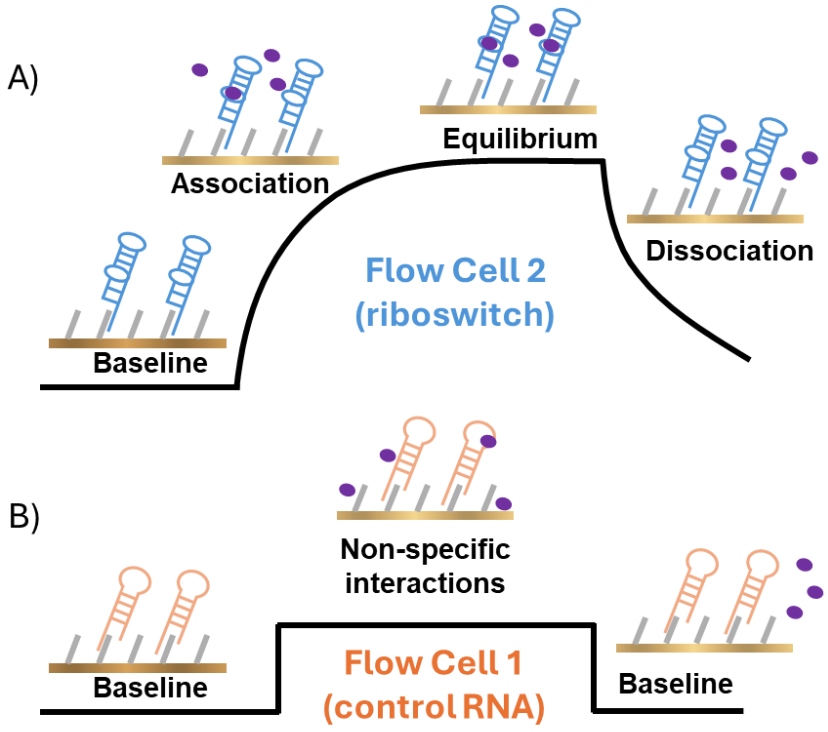

Figure 6: SPR-based ligand binding assay. Riboswitch is immobilized in active flow cell 2 (A), in parallel a scrambled control RNA is used in reference flow cell 1 (B) to account for any non-specific interactions.

As a benchmark for the screening, we first validated riboswitch binding via SPR using native ligands and known analogs. SPR sensorgrams for the TPP riboswitch revealed a K_D of 0.82 nM to native TPP ligand and 370 nM to thiamine (Figure S3), consistent with literature values obtained through SPR and in-line probing assays.[33, 34] It has previously been observed that as aptamer length increases, ligand binding affinities tend to decrease.[12, 13] Therefore, it is not unexpected that the K_D values determined for binding to the *thi*M aptamer domain are slightly lower than those reported in the literature for binding to the full length sequence (*thiM* aptamer domain + expression platform) and similar to previously observed differences.[32] Similarly, we previously found that this SPR method may be used to reproduce FMN, riboflavin and roseoflavin binding to the FMN riboswitch.[32]

The virtual screening of SAM analogues used the crystal structure of SAM-I/IV, env87 ΔU92G93 (PDB ID 4L81), which exhibits a reported affinity for SAM of 7.3 ± 0.4 μM measured by isothermal calorimetry (ITC).[35] Upon SPR analysis, equilibrium binding could not be reached (association time > 100 seconds), therefore we could not generated reliable fitting to obtain an experimental K_D (Figure S4B). Despite these challenges, we explored two alternative riboswitches for our experimental screening. The first, env87 ΔU92 which shares similarities with env87 ΔU92G93 and exhibits a superior reported affinity for SAM (K_D = 410 ± 50 nM[35] also presented similar kinetic challenges (Figure S4C). Consequently, we turned to the *yitJ* SAM-I riboswitch from *B. subtilis*, a riboswitch previously employed in our laboratory, which demonstrated a reported affinity of 4 ± 2 nM measured via SPR (Figure S4D).[32]

Experimental evaluation of virtually screened molecules. With the function of three model riboswitches validated via SPR, the selected compounds from virtual screening for each riboswitch were subjected to an initial binding assay screen. To facilitate high throughput analysis, each compound was tested at high concentrations, limited based on solubility for each individual molecule (Figure S5, S6, S7, S8, S9). From the 20 TPP (Figure S5), 18 FMN (Figure S6) and 14 SAM (Figure S7, S8, S9) compounds, one TPP compound (virtual hit **T12**), seven FMN compounds (virtual hits **F8**, **F10**, **F11**, **F12**, **F15**, **F16** and **F18**) and three SAM compounds (virtual hits **S2**, **S4** and **S11**) resulted in an interaction with the riboswitch immobilized to the SPR chip (Tables S1, S2 and S3, Figures S5, S6 and S7). This increase in

response compared to controls and background may possibly indicate binding. (Note that the SAM compounds were tested against all three SAM-I and SAM-I/IV riboswitch variants.) Following screening, binding isotherms were generated by testing a concentration series of the hits against each riboswitch compared to RNA control sequences. From this more detailed evaluation, Virtual hit **T12** against the TPP riboswitch aptamer domain demonstrated specific binding to the aptamer domain with a modest KD value of 170 ± 25 μM (Figure 7E). Excitingly, this hit was resynthesized, validated via NMR and its activity was confirmed (supporting information). From the FMN virtual hits, despite observing a notable response from seven small molecules against the FMN riboswitch aptamer domain, a reliable determination of the affinity constant, K_D, 220 ± 180 μM (Figure 7F) was achieved for only one FMN virtual hit (hit **F12**). Finally, no reliable binding isotherms were achieved with the virtual hits to the SAM-I/IV or SAM-I riboswitch. However, this is not surprising given that the native ligand did not result in a measurable response via SPR to the SAM-I/IV RNA, while the highly functional SAM-I riboswitch was not the structure used in our virtual screen. This highlights the uniqueness of each riboswitch structure and the need for additional high-resolution structures of riboswitches. Nevertheless, given that the experimental characterization was performed in the presence of a "Scrambled" RNA, we can confirm that the two newly discovered interactions in this work (T12 and F12) are not non-specifically interacting with nucleic acids. As such, the molecules identified (Figure 7A and 7B; Table s1 and S2) are specific to the TPP and FMN riboswitch structures, respectively.

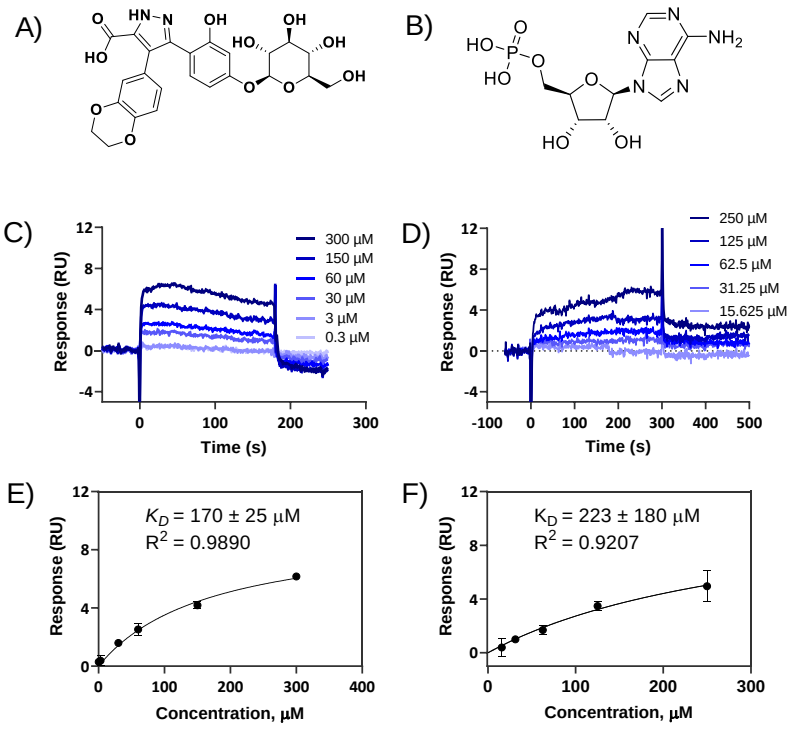

Figure 7. Two of the virtually screened molecules specifically bind to the natural riboswitches: TPP riboswitch (left panel) and FMN riboswitch (right panel). Chemical structure of TPP virtual hit **T12** (A) and FMN virtual hit **F12** (B). SPR responses with a series of concentrations of hits (C and D). Data obtained from C and D are plotted as a function of hit concentration and fitted to a simple 1:1 binding isotherm to determine the binding constant K_D (E and F respectively). Error bars represent the standard deviation of duplicate experiments. Both TPP and FMN experiments were performed at 5 mM MgCl2 with 0.5% DMSO.

CONCLUSION

Through this work, we have successfully demonstrated the potential of our docking program FITTED in the search for specific RNA-targeting small molecules. More explicitly, one molecule showed potential interaction with the TPP riboswitch, seven showed potential binding to the FMN riboswitch and three displayed features of binding to the SAM-I/IV riboswitch. Considering that only 51 molecules were selected by virtual screening and tested, this represents a success rate of over 20% for this class of macromolecules. Upon validation, two hits emerged with modest binding (micromolar range) but displayed specificity for the target riboswitch RNA compared to non-target RNA, suggesting a potential avenue for using our approach to develop new ligands to RNA structures. Interestingly, the identified compounds are all quite distinct from the native ligands of these riboswitches. While these molecules are only preliminary hits and would require significant optimization to be of interest as potential drugs, this work demonstrated that our docking method, despite its major development focusing on proteins, can be used to identify molecules binding to RNA motifs. As the growing resistance of pathogenic bacteria to antibiotics represents a major threat to human welfare, improving our ability to target new biomolecules, such as bacterial riboswitches,[10-12] is critical.

EXPERIMENTAL

Virtual library preparation. A virtual library of compounds was assembled for docking to the TPP riboswitch. The initial library included compounds with known antibacterial activity (from CHEMBL repository), RNA-binding motifs (from ChemSpace and Life Chemicals compound libraries), or metal-binding motifs (from ChemSpace compound library). This initial library was filtered based on drug-like properties to select compounds with fewer than 10 rotatable bonds, a molecular weight between 300 – 550 Da, no reactive groups (eg, aldehyde, acyl chloride, epoxide) using our programs SMART to compute molecular properties and REDUCE to select molecules fulfilling these criteria.[36] Based on the binding motif of TPP, the library was also filtered to select molecules that contain at least two metal coordinating groups that would be necessary for coordination to the magnesium ions of the pyrophosphate sensor helix. The library was also filtered to include molecules that contained at least one hydrogen bond donor that would be necessary to facilitate key hydrogen bond interactions in the pyrimidine sensor helix. Molecules containing functional groups that were too reactive to be included in potential drugs were removed and 6,000 diverse compounds were selected using our program SELECT and ECFP4 fingerprints. Libraries for docking against SAM and FMN were assembled following a similar protocol.

Riboswitch structures. The protein data bank was searched for all crystal structures of TPP, SAM and FMN riboswitches bound to ligands. Each crystal structure was self-docked and cross docking was also performed if more than one structure was available. Each crystal structure had the self-docking score, cross-docking score, organism, and resolution considered for the crystal structure that would be used for the ligand screening step. The crystal structures that were selected were 2GDI (TPP riboswitch), 4L81 (SAM-I riboswitch), and 3F2Y (FMN riboswitch).

Docking methods. Our program SMART was used to prepare the libraries for docking using default parameters. PREPARE and PROCESS were used to prepare the riboswitches for docking from crystal structures using default parameters. FITTED was used for docking. For each riboswitch, the FITTED Score was used to automatically select the top 250 scoring compounds. These compounds were visualized, and 20-30 compounds were selected for experimental testing based on manual evaluation of key binding interactions between the docked ligand and the riboswitches

Synthesis and Purification of DNA Templates. TPP, FMN and SAM-I riboswitches, along with scrambled control RNA, were produced through in vitro transcription from single-stranded DNA (ssDNA) template oligonucleotides (Table S4) purchased from Integrated DNA Technologies (IDT). These DNA templates contain a 24-mer poly(A) tail at the 3′ end for immobilization onto the SPR chip and 23-nucleotide

T7 promoter sequence at the 5′ end used for T7 transcription. The ThiM aptamer domain was designed using three separate oligos, while both the FMN and SAM-I aptamer domains were designed using four separate oligos to accommodate their longer length (Table S5).

PCR amplification of DNA templates was performed using Phusion High-Fidelity DNA Polymerase (New England Biolabs) enzyme. A 50 µL of reaction mix (RNase-free water: 32.5 µL; 5x Phusion Buffer: 10 µL; 10 mM dNTP solution: 1 µL; 10 µM forward oligo: 2 µL; 10 µM reverse oligo: 2 µL; ssDNA template: 2 µL; Phusion polymerase: 0.5 µL) was prepared as per the manufacturer guidelines. The PCR reaction was performed in a BioRad T100 Thermal Cycler starting at 98°C for 30 seconds. 35 cycles of 8 seconds at 98 degrees, 20 seconds for annealing at 52°C, and 10 seconds at 72 degrees for extension were performed, followed by a final 2 min extension at 72 degrees. PCR product and length was verified via 3% agarose gel electrophoresis with SYBR SafeDye in TAE buffer (Tris base, acetic acid, and EDTA) at pH 8.3. Size was confirmed by imaging with a Bio-Rad XR + Gel Documentation System Gel Imager by comparing PCR product to a 100-1500bp DNA Marker. PCR product was then purified using the New England Biolabs Monarch DNA Clean-Up Kit. DNA yield was determined via a Thermo Fisher NanoDrop Spectrophotometer and stored at -20◦C until use for T7 Transcription.

Synthesis of the RNA Riboswitches by *in vitro* T7 Transcription. The purified DNA templates were subjected to *in vitro* transcription using the MEGAshortscript T7 Polymerase kit (ThermoFisher) for small-scale transcriptions as described by the manufacturer. Reactions were performed for 16 hours at 37°C, treated with 1 µL of turbo DNase at 37°C for 30 minutes, then purified using the Zymo RNA Clean & Concentrator kit. All RNA were tested for appropriate length and purity on an 8% denaturing poly-acrylamide gel electrophoresis (PAGE). Finally, RNA was quantified by NanoDrop and stored in the freezer until use.

Preparation of Binding Buffer. All experiments were conducted using an HBS-N running buffer (Cytiva). To create a solution of 1X HBS-N with a pH of 7.4 and containing either 2 or 5 mM MgCl2, we added the appropriate amount of 1M MgCl2 to a 10X HBS-N stock, added DEPC-treated (RNAse-free water) to dilute the mixture and vacuum filtered. A similar buffer containing 0.1% or 0.5% Dimethyl sulfoxide (DMSO) for small molecules that required DMSO for dissolution was prepared. The exact same batch of buffer was used as the running buffer on the SPR and to prepare the small molecule fresh (same day).

Preparation of Small Molecules Solutions. The selected molecules were purchased from ChemSpace and used as is. To prepare solutions of small molecules for experiments, we took into consideration the solubility of the small molecules and used both aqueous solvent water and eventually the organic solvent Dimethyl sulfoxide (DMSO). For some small molecules such as TPP, Thiamine Hydrochloride and FMN we prepared 5 mM stock solutions. Stock solutions were then diluted as needed.

Many of the molecules identified by virtual screening had limited solubility in water, therefore dissolved in DMSO at concentrations ranging from 1 mM to 100 mM.. To be compatible with the SPR assay, these molecules later diluted to either 0.1% or 0.5% DMSO using running buffer without DMSO. To avoid any buffer mismatch, the same percentage of DMSO was maintained in both the running buffer and the sample solution throughout the assay.

SPR Sensor Surface Immobilization. A Biacore X100 instrument was used to conduct experiments at a temperature of 25°C. The DNA linker strand (AmMC6-TTTTTTTTTTTTTTTTTTTTTTTT) (IDT), was immobilized to a CM5 carboxymethylated dextran sensor chip. Before immobilization, a pre-concentration test is used to achieve higher immobilization levels of the DNA linker to maximize the sensor surface's capacity. A positively charged carrier 1.2 mM hexadecyltrimethylammonium bromide (CTAB) micelle prepared in 10 mM HEPES is used to overcome the electrostatic repulsion between the negatively charged DNA linker and sensor surface that is decorated with COOH group. Once the pre-concentration

assay showed enough interaction, an immobilization reaction to covalently attach the DNA linker onto both flow cells (FC1 and FC2) of the sensor chip was performed. The carboxylate group on the sensor chip was activated by injecting a 1:1 volume ratio of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and N-Hydroxysuccinimide (NHS) followed by injection of DNA linker and CTAB. An injection of 1M ethanolamine at pH 8.5 was used to block any remaining activated groups. Approximately 3,000 Response Units (RU) of the DNA strand were obtained by performing the immobilization reaction sequentially on both flow cells (FC1, FC2).

Riboswitch-Target Interaction Assay. The Biacore X100 instrument was prepared by priming with running buffer prior to all binding assays. To stabilize the sensorgram baseline, at least two startup cycles were performed for each assay. During each startup cycle, the riboswitch RNA was captured onto the sample flow cell FC2 and a scrambled sequence of the same riboswitch onto the flow cell (FC1) for 40 seconds at a flow rate of 5 µL/min. We aimed to have between 2000-5000 RU capture to ensure maximum binding response by the small molecule. To regenerate the sensor surface, 25 mM NaOH was injected for 30 seconds at a flow rate of 30 µL/min over both flow cells.

For each cycle, approximately 1µg of riboswitch was injected onto the flow cells for 40 seconds at a flow rate of 5 µL/min and the small molecule solution of different concentration that were prepared in running buffer was injected over both flow cells (FC1 and FC2) at a flow rate of 30 µL/min. Running buffer was also injected over both flow cells at a flow rate of 90 µL/min to monitor target dissociation. We adjusted the association and dissociation time, and flow rate according to the need to attain equilibrium. The riboswitch and small molecule solution was removed from the sensor surface by injecting 25 mM NaOH for 30 seconds at a flow rate of 30 µL/min over both flow cells to prepare the surface for the next cycle.

The Biacore X100 Evaluation Software version 2.0 (Cytiva) was used for processing and analyzing the data. To attain precise outcomes, a double-referencing method was employed to process all datasets. Initially, data from the sample flow cell (FC2) were referenced by subtracting the data from the reference flow cell (FC1). This step corrected for any bulk refractive index changes, nonspecific binding, injection noise, and baseline drift. Following this, the reference-subtracted data (FC2 − FC1) were double-referenced with a blank injection of running buffer to account for any systematic drift during the injection. Finally, the double-referenced data were fitted to a 1:1 binding model for a steady-state affinity model for thermodynamic analysis.

ADDITIONAL INFORMATION

Supporting Information. AUROC graph for testing the accuracy of docking to TPP, list and structure of compounds selected for testing, oligonucleotide sequences, sensorgrams, resynthesis of the TPP active hit. This material is available free of charge via the Internet at http://pubs.acs.org."

Corresponding Authors and author contributions

* Nicolas Moitessier – McGill University, Montreal, QC Canada. https://orcid.org/0000-0001-6933-2079

* Maureen McKeague – McGill University, Montreal, QC Canada. https://orcid.org/0000-0002-3750-6027

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. #These authors contributed equally.

Funding Sources

Genome Quebec (Programme d'intégration de la génomique); Fonds de Recherche du Quebec [Equipe grant 284708 to N.M.], Natural Sciences and Engineering Research Council of Canada [Discovery Grant

2019-04949 to M.M., Research Tools and Instrument Grant to M.M. for SPR and NSERC CREATE support to N.T.]; M.M. is supported by the Canada Research Chairs program.

ACKNOWLEDGMENT

We thank members of the Moitessier lab, McKeague lab and Molecular Forecaster for assistance and useful discussions. We thank Digital Research Alliance of Canada (formerly Compute Canada) for their generous CPU time allocations.

ABBREVIATIONS

SPR, Surface plasmon resonance; TPP, thiamine pyrophosphate; FMN, flavin mononucleotide; SAM, S-adenosyl methionine; SPR, surface plasmon resonance; K_D, equilibrium binding constant; DMSO, dimethyl sulfoxide; PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; ITC, isothermal calorimetry; RMSD, root mean square deviation; PDB, protein data bank.

REFERENCES

1. Mark Andrew, P.; Marisa, A. S.; Darby, L. W.; Zhong-Ru, X., Has Molecular Docking Ever Brought us a Medicine? In *Molecular Docking*, Dimitrios, P. V., Ed. IntechOpen: Rijeka, 2018; p Ch. 8.
2. Luo, J.; Wei, W.; Waldispühl, J.; Moitessier, N., Challenges and Current Status of Computational Methods for Docking Small Molecules to Nucleic Acids. *Eur. J. Med. Chem.* **2019,** *168*, 414-425.
3. Ruiz-Carmona, S.; Alvarez-Garcia, D.; Foloppe, N.; Garmendia-Doval, A. B.; Juhos, S.; Schmidtke, P.; Barril, X.; Hubbard, R. E.; Morley, S. D., rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Comput Biol* **2014,** *10* (4), e1003571.
4. Feng, Y.; Zhang, K.; Wu, Q.; Huang, S.-Y., NLDock: a Fast Nucleic Acid–Ligand Docking Algorithm for Modeling RNA/DNA–Ligand Complexes. *J. Chem. Inf. Model.* **2021,** *61* (9), 4771-4782.
5. Buckley, M. E.; Ndukwe, A. R. N.; Nair, P. C.; Rana, S.; Fairfull-Smith, K. E.; Gandhi, N. S., Comparative Assessment of Docking Programs for Docking and Virtual Screening of Ribosomal Oxazolidinone Antibacterial Agents. *Antibiotics* **2023,** *12* (3), 463.
6. Agarwal, R.; T, R. R.; Smith, J. C., Comparative Assessment of Pose Prediction Accuracy in RNA–Ligand Docking. *J. Chem. Inf. Model.* **2023,** *63* (23), 7444-7452.
7. Kallert, E.; Fischer, T. R.; Schneider, S.; Grimm, M.; Helm, M.; Kersten, C., Protein-Based Virtual Screening Tools Applied for RNA–Ligand Docking Identify New Binders of the preQ1-Riboswitch. *J. Chem. Inf. Model.* **2022,** *62* (17), 4134-4148.
8. Aslam, B.; Wang, W.; Arshad, M. I.; Khurshid, M.; Muzammil, S.; Rasool, M. H.; Nisar, M. A.; Alvi, R. F.; Aslam, M. A.; Qamar, M. U.; Salamat, M. K. F.; Baloch, Z., Antibiotic resistance: a rundown of a global crisis. *Infect. Drug Resist.* **2018,** *11*, 1645-1658.
9. Bartlett, J. G.; Gilbert, D. N.; Spellberg, B., Seven Ways to Preserve the Miracle of Antibiotics. *Clin. Infect. Dis.* **2013,** *56* (10), 1445-1450.
10. Kavita, K.; Breaker, R. R., Discovering riboswitches: the past and the future. *Trends Biochem. Sci.* **2022,** *48*, 119-141.
11. Panchal, V.; Brenk, R., Riboswitches as Drug Targets for Antibiotics. *Antibiotics* **2021,** *10* (1), 45.

12. Winkler, W.; Nahvi, A.; Breaker, R. R., Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **2002,** *419* (6910), 952-956.

13. Winkler, W. C.; Cohen-Chalamish, S.; Breaker, R. R., An mRNA structure that controls gene expression by binding FMN. *Proc. Natl Acad. Sci. USA* **2002,** *99* (25), 15908-15913.

14. Nahvi, A.; Sudarsan, N.; Ebert, M. S.; Zou, X.; Brown, K. L.; Breaker, R. R., Genetic Control by a Metabolite Binding mRNA. *Chem. Biol.* **2002,** *9* (9), 1043-1049.

15. Mironov, A. S.; Gusarov, I.; Rafikov, R.; Lopez, L. E.; Shatalin, K.; Kreneva, R. A.; Perumov, D. A.; Nudler, E., Sensing Small Molecules by Nascent RNA: A Mechanism to Control Transcription in Bacteria. *Cell* **2002,** *111* (5), 747-756.

16. Lim, J.; Winkler, W. C.; Nakamura, S.; Scott, V.; Breaker, R. R., Molecular-recognition characteristics of SAM-binding riboswitches. *Angew. Chem. Int. Ed. Engl.* **2006,** *45* (6), 964-968.

17. Mandal, M.; Boese, B.; Barrick, J. E.; Winkler, W. C.; Breaker, R. R., Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. *Cell* **2003,** *113* (5), 577-586.

18. Thore, S.; Leibundgut, M.; Ban, N., Structure of the eukaryotic thiamine pyrophosphate riboswitch with its regulatory ligand. *Science* **2006,** *312* (5777), 1208-1211.

19. McCown, P. J.; Corbino, K. A.; Stav, S.; Sherlock, M. E.; Breaker, R. R., Riboswitch diversity and distribution. *RNA* **2017,** *23* (7), 995-1011.

20. Pavlova, N.; Kaloudas, D.; Penchovsky, R., Riboswitch distribution, structure, and function in bacteria. *Gene* **2019,** *708*, 38-48.

21. Lee, E. R.; Blount, K. F.; Breaker, R. R., Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. *RNA Biol.* **2009,** *6* (2), 187-194.

22. Howe, J. A.; Wang, H.; Fischmann, T. O.; Balibar, C. J.; Xiao, L.; Galgoci, A. M.; Malinverni, J. C.; Mayhood, T.; Villafania, A.; Nahvi, A.; Murgolo, N.; Barbieri, C. M.; Mann, P. A.; Carr, D.; Xia, E.; Zuck, P.; Riley, D.; Painter, R. E.; Walker, S. S.; Sherborne, B.; de Jesus, R.; Pan, W.; Plotkin, M. A.; Wu, J.; Rindgen, D.; Cummings, J.; Garlisi, C. G.; Zhang, R.; Sheth, P. R.; Gill, C. J.; Tang, H.; Roemer, T., Selective small-molecule inhibition of an RNA structural element. *Nature* **2015,** *526* (7575), 672-677.

23. Daldrop, P.; Reyes, F. E.; Robinson, D. A.; Hammond, C. M.; Lilley, D. M.; Batey, R. T.; Brenk, R., Novel ligands for a purine riboswitch discovered by RNA-ligand docking. *Chem. Biol.* **2011,** *18* (3), 324-335.

24. Mariaule, G.; De Cesco, S.; Airaghi, F.; Kurian, J.; Schiavini, P.; Rocheleau, S.; Huskić, I.; Auclair, K.; Mittermaier, A.; Moitessier, N., 3-Oxo-hexahydro-1H-isoindole-4-carboxylic Acid as a Drug Chiral Bicyclic Scaffold: Structure-Based Design and Preparation of Conformationally Constrained Covalent and Noncovalent Prolyl Oligopeptidase Inhibitors. *J. Med. Chem.* **2016,** *59* (9), 4221-4234.

25. Plescia, J.; Dufresne, C.; Janmamode, N.; Wahba, A. S.; Mittermaier, A. K.; Moitessier, N., Discovery of covalent prolyl oligopeptidase boronic ester inhibitors. *Eur. J. Med. Chem.* **2020,** *185*, 111783.

26. Stille, J. K.; Tjutrins, J.; Wang, G.; Venegas, F. A.; Hennecker, C.; Rueda, A. M.; Sharon, I.; Blaine, N.; Miron, C. E.; Pinus, S.; Labarre, A.; Plescia, J.; Burai Patrascu, M.; Zhang, X.; Wahba, A. S.; Vlaho, D.; Huot, M. J.; Schmeing, T. M.; Mittermaier, A. K.; Moitessier, N., Design, synthesis and in vitro evaluation of novel SARS-CoV-2 3CLpro covalent inhibitors. *Eur. J. Med. Chem.* **2022**, 114046.

27. Moitessier, N.; Pottel, J.; Therrien, E.; Englebienne, P.; Liu, Z.; Tomberg, A.; Corbeil, C. R., Medicinal Chemistry Projects Requiring Imaginative Structure-Based Drug Design Methods. *Acc. Chem. Res.* **2016,** *49* (9), 1646-1657.

28. Moitessier, N.; Westhof, E.; Hanessian, S., Docking of aminoglycosides to hydrated and flexible RNA. *J. Med. Chem.* **2006,** *49* (3), 1023-1033

29. Wei, W.; Luo, J.; Waldispühl, J.; Moitessier, N., Predicting Positions of Bridging Water Molecules in Nucleic Acid–Ligand Complexes. *J. Chem. Inf. Model.* **2019,** *59* (6), 2941-2951.

30. Labarre, A.; Stille, J. K.; Patrascu, M. B.; Martins, A.; Pottel, J.; Moitessier, N., Docking Ligands into Flexible and Solvated Macromolecules. 8. Forming New Bonds─Challenges and Opportunities. *J. Chem. Inf. Model.* **2022,** *62*, 1061-1077.

31. Chang, A. L.; McKeague, M.; Liang, J. C.; Smolke, C. D., Kinetic and Equilibrium Binding Characterization of Aptamers to Small Molecules using a Label-Free, Sensitive, and Scalable Platform. *Anal. Chem.* **2014,** *86* (7), 3273-3278.

32. Rivera, M.; Ayon, O. S.; Diaconescu-Grabari, S.; Pottel, J.; Moitessier, N.; Mittermaier, A.; McKeague, M., A sensitive and scalable fluorescence anisotropy single stranded RNA targeting approach for monitoring riboswitch conformational states. *Nucl. Acids Res.* **2024,** *52* (6), 3164–3179.

33. Mehdizadeh Aghdam, E.; Sinn, M.; Tarhriz, V.; Barzegar, A.; Hartig, J. S.; Hejazi, M. S., TPP riboswitch characterization in Alishewanella tabrizica and Alishewanella aestuarii and comparison with other TPP riboswitches. *Microbiol. Res.* **2017,** *195*, 71-80.

34. Sudarsan, N.; Cohen-Chalamish, S.; Nakamura, S.; Emilsson, G. M.; Breaker, R. R., Thiamine Pyrophosphate Riboswitches Are Targets for the Antimicrobial Compound Pyrithiamine. *Chem. Biol.* **2005,** *12* (12), 1325-1335.

35. Trausch, J. J.; Xu, Z.; Edwards, A. L.; Reyes, F. E.; Ross, P. E.; Knight, R.; Batey, R. T., Structural basis for diversity in the SAM clan of riboswitches. *Proc. Natl. Acad. Sci. USA* **2014,** *111* (18), 6624-6629.

36. Therrien, E.; Englebienne, P.; Arrowsmith, A. G.; Mendoza-Sanchez, R.; Corbeil, C. R.; Weill, N.; Campagna-Slater, V.; Moitessier, N., Integrating medicinal chemistry, organic/combinatorial chemistry, and computational chemistry for the discovery of selective estrogen receptor modulators with FORECASTER, a novel platform for drug discovery. *J. Chem. Inf. Model.* **2012,** *52* (1), 210-224.